

A fitness-maximalizáló genomredukció lehetőségei *Escherichia coli* baktériumban

PhD értekezés

Vernyik Viktor

Témavezető:

Dr. Pósfai György

tudományos tanácsadó



Biológia Doktori Iskola

Szegedi Tudományegyetem Természettudományi és Informatikai Kar

Szegedi Biológiai Kutatóközpont, Biokémiai Intézet

Eötvös Loránd Kutatási Hálózat (ELKH)

2022

Szeged

Tartalomjegyzék

Rövidítések jegyzéke	4
I. Szakirodalmi áttekintés	5
1. Genomredukció a természetben	5
2. A mesterséges genomredukció koncepciója	6
3. Csökkentett genomú baktériumtörzsek	7
4. A genomcsökkentés hatása a baktériumtörzsek növekedésére	9
5. Célzott és véletlenszerű genomredukciós módszerek.....	10
II. Célkitűzések	12
III. Anyagok és módszerek	13
1. Felhasznált törzsek, plazmidok és táptalajok	13
2. Módosított Tn5 transzpozon készítése és transzposzómák előállítása	15
3. Random deléciós ciklusok	15
4. Célzott deléciók előállítása	17
5. Genomi pontmutáció létrehozása	18
6. Teljes genom szekvenálás	18
7. Növekedési paraméterek meghatározása	18
8. Kompetíciós kísérletek	19
9. Baktériumsejtek mozgásképességének vizsgálata	20
10. Biomassza hozam meghatározása	20
11. A deletálódott genomi régiók maggenom-elemzése	21
IV. Eredmények	22
1. A random, fitness-optimalizáló deléciós módszer követelményei.....	22
2. Előzetes kísérletek.....	22
3. A RANDEL módszer	23
4. A RANDEL módszer működési jellemzői	25
5. A RANDEL alkalmazásának eredményei.....	26
a) Deléciók és deléciós mechanizmusok	26
b) Pontmutációk és egyéb változások.....	32
c) A szelekció hatásának jelei.....	34

6.	A deléciók és pontmutációk fenotipikus hatásainak elkülönítése.....	35
a)	A rekonstruált deléciók hatása	35
b)	A pontmutációk általános hatása.....	38
c)	Egy kompenzáló pontmutáció.....	40
7.	A célzott és random módszerekkel létrehozott deléciók összehasonlítása	41
8.	A deletálódott genomi régiók és a maggenom viszonya	42
V.	Az eredmények megvitatása	44
VI.	Következtetések.....	49
	Összefoglalás.....	50
	Summary	54
	Köszönetnyilvánítás	58
	Saját közlemények jegyzéke	59
	Irodalomjegyzék.....	60
	Függelék.....	68

Rövidítések jegyzéke

bp/kbp/Mbp	bázispár/kilobázispár/megabázispár
LB	Luria-Bertani (tápközeg)
MS	mineral salts (tápközeg)
aTc	anhidrotetraciklin
dP	6-(β -D-2-dezoxiribofuranozil)-3,4-dihidro-8H-pirimido-[4,5-c][1,2]oxazin-7-on (mesterséges nukleozid analóg)
ap	ampicillin
km	kanamicin
cm	klóramfenikol
HsvTK	<i>Herpes simplex</i> vírus timidin-kináz
OD	optikai denzitás
RANDEL	a dolgozat tárgyát képező random deléciós módszer
DSB	double-stranded DNA break (kettős szálú DNS-törés)
AEJ	alternative end-joining
IS	inszerciós szekvencia
SNV	single nucleotide variant (egyedi nukleotidváltozat)
REP	repetitive extragenic palindrome
TMDR	transposon mediated random deletion

I. Szakirodalmi áttekintés

1. Genomredukció a természetben

A genomredukció, azaz az ősi génkészlet csökkenése egy természetes evolúciós folyamat¹, amely számos baktériumtörzsben megfigyelhető²⁻⁴. A baktériumok általában 2000-8000 gént tartalmaznak, a legkisebb genomúak mindössze 130-600 gént hordoznak³. Tipikusan gazdaszervezethez kötött szimbionták és paraziták mutatják a genomredukció jelenségét⁵. Mivel a gazda stabil tápanyagellátást és védelmet biztosít a környezeti hatásokkal szemben, a sejten belüli niche-ben szükségtelenné váló gének elvesznek az evolúció során. A levéltetvek speciális sejtjeiben élő *Buchnera* fajok szabadon élő, *E. coli*-val közös őse a genom 75%-át elvesztette, ami így 600 kbp méretűre csökkent, amikor obligát endoszimbionta életmódra váltott⁶. Tekintve, hogy az *E. coli* genom mérete 4,5-5,5 Mbp⁷, ez igen jelentős redukciót jelent. A legkisebb ismert, önálló replikációra képes genommal egy kabóca szimbionta, a *Nasuia deltocephalinicola* rendelkezik, mindössze 112 kbp méretű és 137 fehérjekódoló génje van⁸. Jelentős genomcsökkenésen mentek keresztül az intracelluláris parazita baktériumok^{2,9}, közülük a legkisebb a *Mycoplasma genitalium* genomja (580 kbp, ~500 gén)¹⁰. Azonban nem csak a gazdaszervezethez kötődő organizmusokban, hanem a szabadon élő baktériumok között is találunk redukált genomú szervezeteket. Ezek tápanyagszegény tengervizekben rendkívül nagy számban előforduló baktériumok (például *Prochlorococcus*¹¹ és *Pelagibacter*⁴ fajok, valamint az OM43 klád obligát metilotróf tagjai¹²), melyek 1,3-2,5 Mbp méretű genomokkal rendelkeznek.

A genomméret csökkenésében vélhetőleg neutrális és adaptív folyamatok egyaránt szerepet játszanak. Az endoszimbionta és intracelluláris parazita genomok nagymértékű zsugorodását a deléciók irányába eltolódó mutációs aránytalansággal⁶ és genetikai sodródással¹³ magyarázzák, ami a kis effektív populációméretnek köszönhető. A nagy populációmérettel rendelkező, szabadon élő tengeri mikroorganizmusok felfedezése vezetett az ún. adaptív genom „áramvonalasítás” hipotézisének¹⁴ bevezetéséhez. Az elmélet szerint a szelekció közvetlenül a genomméret csökkentésére irányul, hogy a sejt megszabaduljon a metabolikus tehertől, amit az adaptív értékkel nem rendelkező DNS replikációja okoz, és ezáltal hatékonyabban tudja hasznosítani a

korlátozott mennyiségben elérhető tápanyagokat. Mivel a DNS-szintézis a sejtek teljes energiafelhasználásának mindössze 2-3%-át jelenti¹⁵, nem feltétlenül a replikációs költségminimalizálás áll a jelenség mögött. A DNS összmennyiségének csökkenése a sejtek zsugorodásához és a felszín-térfogat arány növekedéséhez vezet, ami hozzájárulhat a hatékonyabb tápanyagtranszportoz¹⁶. Arra is találunk bizonyítékokat, hogy a nagyméretű baktériumpopulációkban a szelekció nem önmagában a genom méretére hat, hanem olyan génfunkciók elvesztésére irányul, amelyek bizonyos környezeti körülmények között fitnesselőnyhöz juttatják a baktériumot¹⁷⁻²⁴. A közelmúlt filogenetikai vizsgálatai pedig arra mutattak rá, hogy mind az endoszimbionta, mind a szabadon élő baktériumokban a megemelkedett mutációs ráta lehet a genomredukció fő hajtóereje²⁵.

2. A mesterséges genomredukció koncepciója

A természetben megfigyelhető evolúciós folyamatok mintául szolgálhatnak a laboratóriumi körülmények között történő, mesterséges genomredukciós törekvésekhez. Ezeket a projekteket a minimális élet és az ahhoz szükséges genom keresése, valamint biotechnológiai célok motiválják. Biotechnológiai szempontból az az alapfeltevés, hogy a laboratóriumban vagy az iparban általánosan használt törzsek (pl. *E. coli*, *B. subtilis*, *C. glutamicum*) széles génrepertoárral rendelkeznek, biztosítva a változatos körülmények között történő túlélést a természetben. Azonban ezek közül sok gén szükségtelen egy szűkebb, viszonylag állandó környezetben, például lombikban vagy fermentorban történő tenyésztés során. A felesleges gének fitnessköltséget jelenthetnek a sejtek számára, ezért az eltávolításuk a sejtek gazdaságosabb működéséhez, erőforrásainak hatékonyabb kihasználásához vezethet^{24,26}. A várható előnyök között szerepel a magasabb növekedési ráta és biomassza hozam, nagyobb genetikai stabilitás és a potenciális patogének „megszelídítése”, azaz a biztonságosabb kezelés.

A kísérleti evolúciós munkák nem jártak nagymértékű genomredukcióval. *Salmonella enterica* baktériumon végzett sorozatos átoltásos kísérletben a deléciós rátát 0,05 bp–2,5 bp per genom per generáció értékre becsülték²⁷. Hasonlóképpen, egy hosszú távú laboratóriumi evolúciós kísérletben *E. coli* sejtek 20 000 generáción

keresztül történő átoltása után csak néhány deléciót kaptak, amelyek összesített mérete mindössze 38 kbp volt²⁸. Az ilyen módon elérhető genomredukció mértéke a gyakorlati alkalmazhatósághoz túlságosan alacsony és jelenleg nem ismert olyan rendszer, amelynek segítségével közvetlenül a kisebb genomokra lehetne szelektálni. Ennélfogva a genomcsökkentéssel foglalkozó projektek tipikusan racionálisan tervezett, specifikus deléciók egymásra építésén alapulnak.

Az eltávolításra kijelölt célszekvenciákat főként összehasonlító genomikai módszerekkel azonosították, de felhasználták gén-esszencialitási és genomszerkezeti vizsgálatok eredményeit, valamint *in silico* modellek előrejelzéseit is. A deléciós célszekvenciák többnyire evolúciós léptékben a közelmúltban szerzett, horizontális transzferrel szerzett genomi szakaszok, mint például profágok, inszerciós szekvenciák (IS-ek). Emellett gyakori célpontok a repetitív szekvenciák, virulencia faktorok és sejtfelszíni struktúrák, a sejt védekező rendszerének elemei és genetikai diverzitásképző faktorok^{29–31}.

3. Csökkentett genomú baktériumtörzsek

A legtöbb genomredukciós munka az *E. coli* baktérium egyszerűsítését célozta^{32–37}, de más fajokban is végeztek genom áramvonalasítást, így számottevő genomcsökkenést értek el *Bacillus subtilis*-ben^{38–40}, *Corynebacterium glutamicum*-ban⁴¹, valamint több *Streptomyces*^{42,43} és *Pseudomonas* fajban^{44–47}. A 4,6 MBp méretű *E. coli* K-12 MG1655 genom számos genomredukciós projekt kiindulási pontja. Kutatócsoportunk, valamint több japán laboratórium jelentős eredményeket értek el az *E. coli* genom méretének csökkentésében. A leginkább figyelemre méltó genomcsökkentett törzsek közé tartozik a $\Delta 16$ ³⁴ nevű (1,38 MBp deléció, 29,7%-os redukció), az MS56⁴⁸ (1068 kbp, 23%), az MDS42⁴⁹ (663 kb, 14,3%), és az előző folytatásának tekinthető MDS69⁵⁰ (939 kbp, 20,3%). Az *E. coli* K-12 W3110 törzs szolgált alapul az MGF-01³⁶ törzsnek (1,03 Mbp, 22%) és a további deléciókat, valamint kisebb módosításokat tartalmazó DGF-298³⁷ törzsnek (1,67 Mbp, 35,8%).

A célzott deléciók tervezését a génfunkciók és az epiztatikus kölcsönhatások hiányos ismerete, valamint genomszerkezeti korlátok nehezítik. További probléma, hogy a célzott deléciók létrehozásának munkaigényes folyamata limitált számú egyedi deléció

létrehozását engedi meg, illetve a deléciók útvonalak (deléciók kombinációja, sorrendje) hatásának tesztelése is erősen korlátozott. Mindezek a nehézségek tehát jelentősen szűkítik a genomredukciós kísérleti terepet. Valószínűleg ezek a korlátok is felelőssé tehetők azért, hogy a genom áramvonalasítással foglalkozó munkák során a várt előnyök gyakran csak részben mutatkoztak meg.

A fenti nehézségek ellenére célzott genomdeléció révén számos előnyös tulajdonságú törzset sikerült előállítani. A mobilis genetikai elemek és a hibázó DNS-polimerázok kiiktatásával sikeresen megnövelték egy organizmus genetikai stabilitását^{51,52}, és ennek jelentőségét demonstrálták *E. coli*-ban⁵³. Stabilizált genomú törzseket alkalmaztak hosszú direkt repeateket tartalmazó lentivirális vektorok fenntartására⁵⁴, IS-mentes P1 fágok készítésére⁵⁵, virális^{56,57} és humán⁵⁸ fehérjék expressziójára. Ezen felül esetenként magasabb szintű rekombináns fehérje expressziót és kismolekula termelést értek el. Egy redukált genomú *Pseudomonas putida* gazdasejtben 40%-os hozamnövekedést tapasztaltak a GFP (zölden fluoreszkáló fehérje) termelésében⁵⁹. *Pseudomonas mendocina* NKU421 sejtekben az intracelluláris ATP/ADP arány 11-szeresére nőtt, és a PHA (polihidroxialkanoát) termelés 115%-kal emelkedett a szülői törzshöz képest⁴⁷. *Bacillus subtilis* MGB874 sejtekben egy extracelluláris celluláz és egy proteáz expressziójában mutattak ki növekedést³⁸, a P10 törzsben pedig olyan fehérjét sikerült termeltetni (pl. staphylococcalis antigéneket), melyeket más *B. subtilis* törzsekben nem volt lehetséges^{39,40}. *E. coli* MDS42 gazdában pedig jelentős javulást értek el egy iparilag fontos izoamiláz enzim expressziójában⁶⁰. Mindemellett számottevő hozamnövekedést írtak le az L-treonin aminosav produkciójában módosított *E. coli* MDS42⁶¹ és MGF-01³⁶ törzsekben. Megjegyezzük, hogy talán a legjelentősebb alkalmazása a redukált genomú törzseknek, konkrétan a laboratóriumunkban készített MDS42-nek, hogy a genetikai kódrendszer egyszerűsítésére irányuló kísérletekben kiindulási törzsnek használták fel⁶².

Az eddig tárgyalt „top-down” megközelítés, azaz a meglévő genomok redukciója mellett érdekes alternatívát nyújtanak a teljes genomok szintézisén alapuló „bottom-up” módszerek. Egy úttörő munkában a *Mycoplasma mycoides* 1,08 Mbp méretű genomjának darabjait *in vitro* szintetizálták, élesztőben összeállították, majd a szintetikus genomot (JCVI-syn1.0) *Mycoplasma capricolum* recipiens sejtekbe

transzplantálták, amelyeket aztán teljes egészében a szintetikus kromoszóma kontrollált⁶³. Néhány évvel később, Hutchison és munkatársai ennek a genomnak egy jelentősen redukált változatát hozták léte⁶⁴. Transzpozon mutagenézisen alapuló gén-esszencialitás vizsgálatok eredményeire támaszkodva megtervezték, szintetizálták, majd sikeresen transzplantálták a *M. mycoides* genom 531 kbp méretű változatát (JCVI-syn3.0), amely mindössze 473 gént tartalmaz.

4. A genomcsökkentés hatása a baktériumtörzsek növekedésére

A számos biotechnológiai alkalmazás ellenére alig találunk példát a csökkentett genomú baktériumsejtek gazdaságosabb működésére, azaz a tápanyagok hatékonyabb felhasználására vonatkozóan. Sőt, úgy tűnik, a genomredukációs folyamat előrehaladása gyakran a növekedési ráta csökkenésével jár^{34,50,65}. Az MDS-sorozat esetén bizonyos genomi szegmensek deléciója növelte a fitnesszt, ugyanakkor nem találtak pozitív lineáris kapcsolatot a fitnessz és a genomcsökkenés mértéke között⁵⁰. Az MGF-sorozat⁶⁵ és a $\Delta 16$ törzs³⁴ leszármazottjaiban korrelációt mutattak ki a genomméret és a növekedési ráta csökkenése között. A MGF-származékok esetén a hatás különösen a tápanyagokban szegényebb táptalajokban (MAA, M63) volt jelentős⁶⁵. A 22%-os redukciós állapotot képviselő MGF-01 viszont M9 minimál táptalajban a kiindulási törzshöz hasonló növekedési rátát mutatott³⁶. Ennek az lehet az oka, hogy a készítésénél figyelembe vették a növekedési képességet: az egyedi deléciókat hordozó törzsek növekedését ellenőrizték M9 táptalajban, és csak a jól növekvő konstrukciókkal folytatták a deléciós sort. További deléciók generálásával, bizonyos deletált szakaszok visszahelyezésével (a proVXW régió visszaállítása megszüntette a magas ozmolaritással szembeni érzékenységet) és kisebb genetikai módosításokkal (kijavították a valin érzékenységet felelős *ilvG* mutációt és a részleges pirimidin éhezésért felelős *rph-1* mutációt) létrehozták a 2,98 Mbp genomméretű DGF-298 törzset. A baktérium M9 minimál táptalajban jobban nőtt a szülői W3110 törzsnél és magasabb sejtsűrűséget mutatott CSL fermentációs médiumban³⁷. A törzs transzkriptom analízise alapján feltételezik, hogy egyes hőszokk chaperon fehérjék és abnormális fehérjék lebontásáért felelős proteázok csökkent kifejeződése lehet a kedvező növekedési sajátságok egyik magyarázata³⁷.

5. Célzott és véletlenszerű genomredukciós módszerek

A korábbi genomredukciós módszerek elsősorban tervezett, előre meghatározott deléciók precíz létrehozását célozták. Ezt homológ rekombinációt alkalmazó, plazmid beépülésen/kivágódáson alapuló, illetve változatos, a λ Red rendszert használó „recombineering” módszerek igénybevételével érték el. Ehhez első lépésként egy mesterséges DNS-darabot rekombináltatnak a genomba, amely a létrehozni kívánt delécióval határos genomi szakaszokkal homológ szekvenciákat és egy szelekciós markert hordoz. A második rekombinációs lépés során az exogén szekvencia kivágódik, és egy markermentes deléció keletkezik. A lépések ismétlésével a deléciók felhalmozhatók a sejtekben²⁹.

Részlegesen véletlenszerű, „szemi-random” módszernek tekinthető az egyik első genomredukciós munka, amelynek alkalmazásával 6,7%-kal csökkentették az MG1655 genomot³². Először egy térképezett transzpozon könyvtárat hoztak létre, majd P1 fágtranszdukcióval két transzpozon inszerciót egy sejtbe vittek át, végül Cre/loxP helyspecifikus rekombinációs rendszer segítségével a köztes szakaszt kivágták. A kapott deléciókat P1 transzdukcióval egyesítették egyetlen genomra. A módszer alkalmazása azonban komplikált és a rekombináz felismerő szekvenciájának visszamaradásával jár, ami genomi instabilitáshoz vezethet.

Egy másik korai munkában egy kompozit transzpozon származékot használtak sorozatos véletlenszerű deléciók generálására *E. coli*-ban³³. A transzpozon inszercióját követően indukálták a „belső” transzpozon intramolekuláris transzpozícióját, ami a szomszédos kromoszómális régió delécióját vagy inverzióját okozta. A folyamatot ciklikusan ismételve 5,6%-os redukciót értek el. A módszer azonban számos hátránnyal bír: a transzpozon inszerciós sejtek csak egynegyed részében megy végbe megfelelő átrendeződés, viszonylag kisméretű deléciók képződésének kedvez, valamint egy 64 bp-os exogén szekvenciát hagy hátra a genomban.

Ebben a munkában egy új megoldásokon alapuló, véletlenszerű, szelekciót is magába foglaló eljárást mutatunk be, melyhez az inspirációt Koskiniemi és munkatársainak tanulmánya²¹ jelentette, amelyben megmutatták, hogy a szelekció a génvesztés hajtóereje lehet baktériumokban. *Salmonella enterica* kromoszóma véletlenszerű helyeire transzpozonokat integráltak, majd mérték ezek spontán

delécióinak rátáját és a fitnessre gyakorolt hatását. Megállapították, hogy a különböző genomi lókuszok deléciós rátái között több, mint 200-szoros különbség lehet. A deléciókat a kiindulási törzsben rekonstruálva megmutatták, hogy a vizsgált deléciók kb. 25%-a fitnesselőnyvel jár, melyek a sorozatos átoltások folyamán fixálódnak a populációban.

II. Célkitűzések

Munkánk alapvető célja a lehető legnagyobb fitnessszel rendelkező, redukált genomú *E. coli* törzsek előállítása.

Ehhez kapcsolódóan a következő kérdésekre szeretnénk választ kapni:

- Vannak-e olyan deléciók, melyek adott körülmények között növelik a fitnesszt, illetve elősegítik a hatékonyabb tápanyag-felhasználást, azaz gyorsabb növekedést és nagyobb biomassza termelést okoznak?
- Ha igen, konkrétan melyek ezek a deléciók?
- Létrejönnek-e párhuzamos, különböző redukációs útvonalak? Van-e jelentősége a deléciók sorrendiségének?
- Szelektálódnak-e a genomredukció közben kompenzáló mutációk?

Mindezen kérdések megválaszolásához célul tűztük egy olyan módszer kidolgozását, mely véletlenszerűen generál deléciókat, és szelektációs lépéseket tartalmaz a legjobb fitnesszú sejtek kiválogatására.

III. Anyagok és módszerek

1. Felhasznált törzsek, plazmidok és táptalajok

Kísérleteinkben a vad típusnak tekintett *Escherichia coli* K-12 MG1655 törzset⁶⁶ (NC_000913.3), ennek multideléciós származékát, az MDS42-t⁴⁹ (NC_20518.1) és az MDS42 π -t (Pir fehérjét expresszál) használtuk. Az MDS42 π törzset laboratóriumunkban korábban úgy állítottuk elő, hogy DH5 α pir⁶⁷-ből P1 fágtranszdukcióval átvittük a Pir fehérje génjét MDS42 törzsbe. A DNS konstruktok MDS42-ben készültek, a random deléciós ciklusok kiindulási törzsei az MG1655 és az MDS42 voltak. Az MDS42 π törzset az R6K γ replikációs origóval rendelkező plazmidok fenntartására használtuk (pl. pSG76-A). A pHKH⁶⁸ és pSG76-A⁶⁹ plazmidok a transzpozon elkészítése során PCR templátként szolgáltak. Az I-SceI enzim expresszióját a pSTAST vagy pSTCST nevű plazmidról indukáltuk; ezek a pSTKST⁷⁰ plazmid ampicillin, illetve klóramfenikol rezisztencia gént hordozó változatai. pST76-A⁶⁹ alapú plazmidokat a célzott deléciók és egy genomi pontmutáció elkészítéséhez használtunk fel. A felhasznált törzsek és plazmidok tulajdonságait az 1. táblázat foglalja össze.

A baktériumkultúrák növesztéséhez Luria-Bertani (LB) tápközeget (1% tripton, 0,5% élesztőkivonat, 0,5% NaCl) vagy 0,2% glükózzal kiegészített mineral salts (MS) minimál tápoldatot⁷¹ használtunk. Az elektroporáció utáni regenerációs lépéshez SOC médiumot (2% tripton, 0,5% élesztőkivonat, 10 mM NaCl, 2,5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glükóz) alkalmaztunk. A kompetíciós kísérletekhez MacConkey agar (BD Difco) lemezeket használtunk a kompetítor sejtípusok megkülönböztetéséhez.

Az anhidrotetraciklin (aTc) indukálószer 100 ng/ml, a 6-(β -D-2-dezoxiribofuranozil)-3,4-dihidro-8H-pirimido-[4,5-c][1,2]oxazin-7-on (dP) nevű nukelozid analógot 1 μ M végkoncentrációban használtuk. Az antibiotikumokat a következő végkoncentrációban alkalmaztuk: 100 μ g/ml ampicillin (ap), 25 μ g/ml kanamicin (km), 24 μ g/ml klóramfeninikol (cm).

1. táblázat. Felhasznált törzsek és plazmidok.

Törzs vagy plazmid	Releváns jellemzők	Funkció	Forrás vagy hivatkozás
<i>E. coli</i> K-12 törzsek:			
MG1655	vad típus (genotípusa: F-lambda- <i>ilvG</i> - <i>rfb</i> -50 <i>rph</i> -1)	kiindulási törzs random deléciók generálásához	66
MG1655 (LacZ)	a vad típus LacZ deficiens változata	törzsek elkülönítése kompetíciós kísérletekben	laboratóriumi gyűjtemény
MDS42	a vad típusú MG1655 multideléciós származéka	kiindulási törzs random deléciók generálásához; pST76-A plazmid származékok gazdatörzse	49
MDS42 π	Pir fehérjét kifejező MDS42 törzs, ami lehetővé teszi az R6K γ replikációs origóval rendelkező plazmidok fenntartását	gazdatörzs a pSG76-A és a pSG76-A-Tn5-SHK plazmidok számára	laboratóriumi gyűjtemény; ⁶⁷
Plazmidok:			
pHKH	két ellenszelekciós gén (<i>hsvTK</i>) között egy szelekciós markert hordoz (<i>km^R</i>)	az EZ-Tn5(SHK) transzpozon szelekciós-ellenszelekciós kazettájának forrása	68
pSG76-A	<i>ap^R</i> gént és R6K γ origót tartalmaz, replikációjához Pir fehérjét igényel	plazmid váz az EZ-Tn5(SHK) készítéséhez	69
pSG76-A-Tn5-SHK	pSG76-A alapú plazmid, egy <i>ap^R</i> gént és az EZ-Tn5(SHK) transzpozont hordozza	az EZ-Tn5(SHK) transzpozon DNS forrása	jelen dolgozat
pSTAST	hőmérséklet-érzékeny replikont tartalmazó alacsony kópiaszámú plazmid; I-SceI-t expresszál a tet promóter szabályozása alatt; a pSTKST származéka, amelyben a <i>km^R</i> gént <i>ap^R</i> génnel cserélték	anhidrotetraciklin (aTc) indukció után biztosítja az I-SceI enzim termelődését	laboratóriumi gyűjtemény; ⁷⁰
pSTCST	a pSTKST származéka, amelyben az <i>ap^R</i> gént <i>cm^R</i> génnel cserélték	anhidrotetraciklin (aTc) indukció után biztosítja az I-SceI enzim termelődését	laboratóriumi gyűjtemény; ⁷⁰
pST76-A származékok	hőmérséklet-érzékeny replikonnal és <i>ap^R</i> génnel rendelkező „öngyilkos” plazmid	célzott, rekonstruált deléciók és egy genomi pontmutáció készítése	69

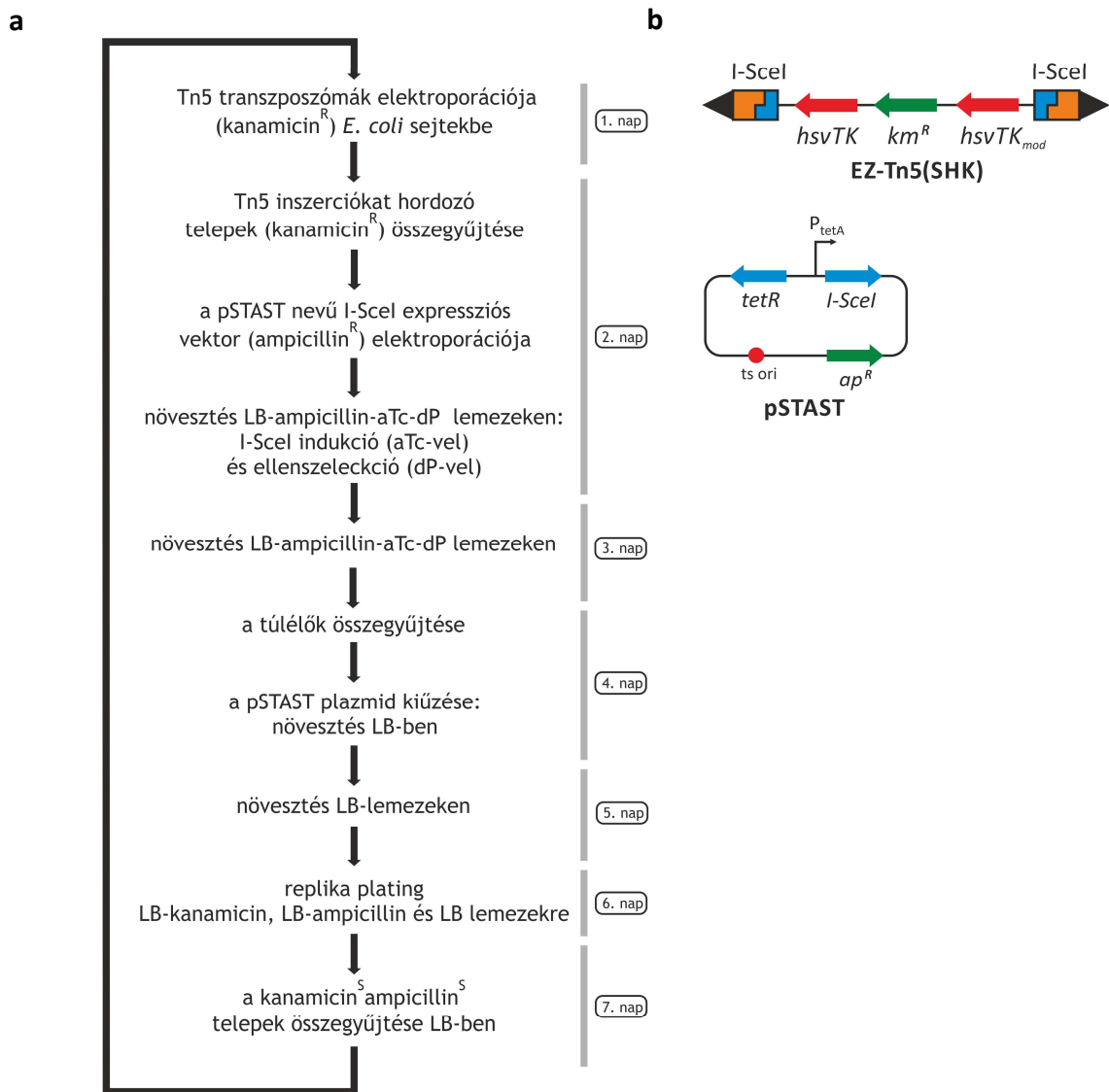
2. Módosított Tn5 transzpozon készítése és transzposzómák előállítása

A pSG76-A plazmidot PCR segítségével linearizáltuk. Az amplifikációhoz olyan primereket használtuk, amelyek 5' túlnyúló végei tartalmazták a hiperaktív Tn5 transzpozon fordított ismétlődéseit (mozaik végeket) és ellentétes orientációban az I-SceI felismerő helyeket. A pHKH plazmidról PCR segítségével felszorzoztuk a duplikált *Herpes simplex* vírus timidin-kináz (hsvTK) géneket tartalmazó szelektációs kazettát (hsvTK-km^R-hsvTK_{mod})⁶⁸, majd ligáltuk azzal a plazmid vázzal, amelyre előzetesen integráltuk a mozaik végeket és az I-SceI felismerő szekvenciákat. Az így kapott plazmid a pSG76-A-Tn5-SHK nevet kapta. Az EZ-Tn5(SHK) transzpozont (Függelék 1. ábra) ebből a plazmidból vágtuk ki PvuII-emésztéssel. A transzposzómák⁷² létrehozásához a kihalított DNS-fragmentumot gélből izoláltuk, majd a gyártó utasításait követve inkubáltuk a transzpozáz enzimmel (EZ-Tn5 Transposase, Lucigen). A pSG76-A plazmid vázra azért esett a választásunk, mert R6K γ replikációs origóval rendelkezik, ezáltal nem képes a célsejtekben replikálódni. Így elkerülhető, hogy a transzpozon DNS kivágása után az esetleges cirkuláris plazmid maradék a transzposzómák bejuttatása után a sejtekbe kerüljön, és álpozitív háttérrel adjon. A felhasznált PCR primerek listáját lásd a Függelék 1. táblázatában.

3. Random deléziós ciklusok

Egy random deléziós ciklus egyes lépéseit az 1. ábrán szemléltetjük. Elektrokompetens sejteket készítettünk a szülői törzsből, illetve a második ciklustól kezdve a megelőző ciklus végén nyert sejtekből. Egy tipikus kísérletben 0,5 μ l EZ-Tn5(SHK) transzposzómát elektroporáltunk 40 μ l elektrokompetens sejtbe, majd 1 ml SOC médiumban, 1 órán keresztül, 37 °C-on rázatva regeneráltattuk a sejteket. Ezután LB-km lemezekre szélesztettük azokat, és a táplemezeket egy éjszakán át, 37 °C-on inkubáltuk. Az inszerciós mutáns telepeket LB-km tápfolyadékban lemostuk a lemezekről, 10 ml LB-km médiumban OD₅₅₀~1 (OD=optikai denzitás) sűrűségűre hígítottuk, majd 2-3 órán keresztül növesztettük 37 °C-on, és starterként használtuk fel elektrokompetens sejtek készítéséhez. 40 μ l elektrokompetens sejtbe ~50 ng pSTAST helper plazmidot transzformáltunk, majd a regenerációt követően (30°C-on 1 óra rázatás 1 ml SOC

táptalajban) LB-ap-aTc-dP lemezekre szélesztettük azokat (az aTc az I-SceI enzimet kódoló gén expressziójának indukciószerként, a dP ellenszelekciós ágensként szolgált). A lemezeket 30 °C tartottuk 35-40 órán keresztül. A helper plazmid kiűzéséhez a kinőtt telepeket folyékony LB-ben összemostuk, 10 ml antibiotikum-mentes LB-ben $OD_{550} \sim 0,1$ értékig hígítottuk, majd egy éjszakán át növesztettük 37 °C-on. Ebből különböző hígításokat szélesztettünk LB-lemezekre, és egy éjszakán át növesztettük 37 °C-on. Végül a sejteket LB és LB-ap/km lemezekre oltva kiválasztottuk azokat a telepeket, amelyek km-re és ap-re is érzékenyek voltak (azaz elvesztették a transzpozont és a helper plazmidot is). A kiválogatott telepeket 10 ml LB-be oltottuk, és starterként használtuk a következő ciklus első lépéséhez, az elektrokompetens sejtek készítéséhez.



1. ábra. Egy random deléciós ciklus egyes lépései (a) és a helper DNS konstruktok (b) vázlatos rajza. Az EZ-Tn5(SHK) transzpozon hordoz egy, az integrációs mutánsok

szelekciójára szolgáló markert (km^R), egy ellenszelekciós gént ($hsvTK$) és annak módosított kópiáját, valamint két ellentétes orientációban elhelyezkedő I-SceI hasító helyet. A pSTAST egy hőmérséklet-érzékeny replikonnal rendelkező plazmid, amely indukálható módon expresszálja az I-SceI enzimet. Minden egyes ciklus növesztési lépéseket tartalmaz (összesen ~ 90 generáció/ciklus), amely során szelektálódnak a magas fitnessű genotípusok.

A protokoll kevésbé mutagén változatához a következő módosításokat vezettük be: a pSTAST elektroporációja után a regenerációt (30°C -on 1 óra rázatás 1 ml SOC táptalajban) követően a sejteket 10 ml LB-ap-aTc folyadékban 5 órán keresztül, 30°C -on növesztettük az I-SceI indukciója érdekében. A kultúrából LB-ap-aTc-dP lemezekre szélesztettünk, majd 30°C -on kb. 16 órán keresztül inkubáltuk. A ciklus ezután következő szakasza megegyezik a fent leírtakkal.

4. Célzott deléciók előállítása

A vad típusú háttéren rekonstruált deléciókat az ún. öngyilkos plazmid alapú módszer segítségével hoztuk létre⁷³. Röviden, ehhez először a tervezett deléciókat határoló ~ 1 kbp méretű DNS-szegmentumokat (homológia karokat) egy hőérzékeny replikációs origóval rendelkező „öngyilkos” plazmidba (pST76-A⁶⁹) klónoztunk MDS42⁴⁹ gazdában. Ezután a plazmid replikációjára nézve restriktív hőmérsékleten, antibiotikum szelekció mellett az egyik homológia karon keresztül a genom megfelelő részére rekombináltattuk. Az így kapott deléciós intermedierekben indukáltuk az I-SceI meganukleáz expresszióját a pSTKST nevű helper plazmidról⁷⁰, ami a kromoszómába integrált plazmidon található egyedi I-SceI felismerő helyen kettős szálú DNS-törést idéz elő. A DNS-törés RecA-mediált intramolekuláris rekombinációs javító mechanizmus segítségével, a homológ szegmensek felhasználásával javítódik, ami a vad típusú kromoszóma reverziójával vagy egy markermentes deléció képződésével jár. A tervezett deléciókat hordozó sejteket PCR-amplifikáció segítségével azonosítottuk, majd Sanger-szekvenálással ellenőriztük a deléciós kapcsolódási pontokat. Többszörös deléciók létrehozásához P1 fágtranszdukcióval⁵⁵ vittük át a deléciós intermediereket, amelyet a fent ismertetett módon I-SceI által stimulált intramolekuláris rekombináció, marker

eltávolítás és ellenőrzés követett. A felhasznált PCR primerek listáját lásd a Függelék 1. táblázatában.

5. Genomi pontmutáció létrehozása

A Rac profág deléciója mellett rekonstruáltuk a *yehB* génben található pontmutációt is – annak fitnesscsökkenést kompenzáló hatásának bizonyításához –, az egyéb pontmutációktól mentes törzsben. Az allélcicserélés a 4. pontban leírt módon történt, azzal a különbséggel, hogy a deléciókat határoló homológia karok helyett a pontmutációt hordozó genomi szegmentumot klónoztuk az „öngyilkos” plazmidba, majd ezt a konstrukciót rekombináltattuk a rekonstruált Rac deléciót hordozó MG1655 törzsbe. A mutáns allélt az egyik random deléciós ciklus során kapott c1-15 nevű törzsről amplifikáltuk PCR segítségével. A felhasznált PCR primerek listáját lásd a Függelék 1. táblázatában.

6. Teljes genom szekvenálás

A random deléciós ciklusok során kapott vonalak teljes genom szekvenciáját Ion Torrent platform segítségével határoztuk meg. Az Ion Torrent könyvtárakat a már korábban leírt módon készítettük⁷⁴. A szekvenálási adatok kiértékelését (jelfeldolgozás és bázisazonosítás) Ion Torrent Suite v5.2 program segítségével végeztük. A nyers leolvasási adatokat CLC Genomics Workbench v11.0 (Qiagen) programba importáltuk, trimmeltük (azaz levágtuk az adaptereket), majd az *E. coli* K-12 MG1655 (NC_000913.3) vagy MDS42 (NC_020518.1) genomi szekvenciákra térképeztük. A referenciától eltérő variáns régiókat manuálisan ellenőriztük. A nagyméretű genomi átrendeződéseket (tipikusan az 1 kbp feletti deléciókat) CLC Genomics Workbench Tool segítségével manuálisan azonosítottuk.

7. Növekedési paraméterek meghatározása

A baktériumtörzsek növekedési paramétereinek meghatározásához a sejteket Synergy 2 (BioTek) nevű automatizált mikrolemez-olvasó készülékben növesztettük. A 96-lyukú lemez egy-egy mélyedésébe 99 µl friss LB médiumot vagy 0,2% glükózzal kiegészített

MS-mininál tápoldatot tettünk, majd ebbe 1 µl éjszakán át növesztett starter kultúrát oltottunk. A sejteket 24 órán keresztül, 37 °C-on növesztettük folyamatos rázatás mellett. A kultúrák 600 nm-en mért optikai denzitás (OD) értékét 5 percenként határoztuk meg. Minden egyes törzs esetén 4-18 replikátumot használtunk. A növekedési paramétereket (növekedési ráta, OD növekmény) korábban publikált módszerek⁷⁵ felhasználásával készített R programozási nyelven írt szkript segítségével határoztuk meg. Először minden egyes mért OD értékből kivontuk a háttérét, azaz a tápfolyadék OD-ját. Mivel magas sejtsűrűség esetén a sejtdenzitás és az OD között nemlineáris korreláció áll fenn, ezért a mért OD értékeket egy kalibrációs formula segítségével korrigáltuk ($OD_{\text{korrigált}} = OD + 0,49312 \cdot OD^3$). A mérési zaj hatásainak csökkentéséhez a növekedési görbéket átlagolással simítottuk, és eltávolítottuk a negatív meredekséget eredményező adatpontokat. A simított görbéket log-transzformáltuk, majd kiszámítottuk a görbe meredekségét az egymástól 20 perc távolságra eső adatpárok között. A maximális növekedési rátát úgy határoztuk meg, hogy a hét legmagasabb érték közül a két legmagasabbat figyelmen kívül hagytuk és kiszámoltuk a maradék öt átlagát. Az OD növekményt a simított görbe utolsó OD értékének és az első öt OD érték átlagának különbségeként határoztuk meg.

8. Kompetíciós kísérletek

Az egyes törzsek relatív fitnessét kompetíciós kísérletek segítségével határoztuk meg. Éjszakán át növesztett starter kultúrákat készítettünk a vad típusú, szülői MG1655-ből és a kompetíciós partner törzsből. Ezeket azonos OD-ra hígítottuk, 1:1 térfogatarányban összekevertünk, majd ebből a starter keverékből 40 µl-t 10 ml LB médiumba oltottuk, és 100 ml-es Erlenmeyer lombikban 37 °C-on rázattuk 225 rpm fordulatszámon, 24 órán keresztül. Ebből a kultúrából 40 µl-t naponta friss tápfolyadékba oltottunk. Az átoltásokat legfeljebb 3 napon át végeztük, a kísérleteket minden vizsgált törzs esetében 3-4 alkalommal ismételtük meg. A kompetíciós partnerek arányának nyomon követéséhez közvetlenül a leoltást követően és minden egyes újraoltás során mintát vettünk a kultúrákból. Ezekből különböző hígításokat szélesztettünk MacConkey agar lemezekre, és meghatároztuk a kompetitorok arányát. A vizsgált törzseket a laktóz fermentációs képességre utaló telepszín alapján különítettük el a vad típustól, mivel az

MG1655-nek egy LacZ-deficiens változatát használtuk, amelyben a LacZ funkciója egy stop kodon beépítésével inaktíválva lett⁷⁶. A relatív fitness értékeket egy korábban leírt módszerrel határoztuk meg⁷⁷. Ehhez vettük a vizsgált és a szülői törzs sejtszám arányának természetes alapú logaritmusát, és ábrázoltuk a felnőtt generációk függvényében. Az egyes adatpontokra illesztett lineáris regressziós egyenes meredeksége adja az adott törzs szelekciós együtthatóját (s). A szülői törzshöz viszonyított kompetitív fitness (w) a $w = 1 + s$ összefüggéssel adtuk meg. Az 1 relatív fitness érték jelentése, hogy a vizsgált törzs rátermettsége megegyezik a szülői törzsével.

9. Baktériumsejtek mozgáskéességének vizsgálata

A flagelláris genomi régió jelenlétét úgy ellenőriztük, hogy teszteltük a sejtek mozgási képességét lágy agarban. Ehhez egy 96-lyukú mikrolemez egyes mélyedéseibe 200 μ l lágy agart (0,3% agart tartalmazó LB médium) öntöttünk, amelybe a vizsgálandó telepkekből egy kacsnyi mennyiséget szűrtünk. A lemezeket 37 °C-on, egy éjszakán át inkubáltuk, majd a szaporodás diffúz vagy lokalizált jellegét értékeltük. A flagellummal rendelkező, mozgékony baktériumok a szűrés mentén diffúz növekedési zónát hoznak létre, míg a mozgásképtelenek a szűrés vonala mellett növekednek.

10. Biomassza hozam meghatározása

A biomassza hozamot a deléciós ciklusok során is alkalmazott LB médiumban határoztuk meg. Exponenciális növekedési fázisban lévő starter kultúrát 20 ml LB médiumba oltottunk úgy, hogy az 550 nm-en mért OD kiindulási értéke 0,01 legyen. A sejteket 100 ml-es Erlenmeyer lombikban, 160 rpm fordulatszámon történő rázatás mellett egy éjszakán át növesztettük 37 °C-on. Lombikonként 8 ml kultúrából mikrocentrifuga csövekbe centrifugáltuk a sejteket, majd a pelletet ScanSpeed 40 típusú vákuumcentrifugában szárítottuk (1000 rpm, <0.3 mbar) 20 órán keresztül. A száraz biomassza meghatározásához analitikai mérleggen megmértük a kiszárított pelletek tömegét. A méréseket törzsenként 5 biológiai, kultúránként 2 technikai replikátummal végeztük.

11. A deletálódott genomi régiók maggenom-elemzése

A D1-D12 jelölésű deletált régiókon belül található géneket két osztályba, ún. mag (core) és járulékos gének csoportjába soroltuk. A csoportosítást egy korábbi publikáció⁷⁸ alapján végeztük, amelyben meghatározták 60 *E. coli* törzs közösen előforduló génjeit, és ezáltal kb. 2000 core gént azonosítottak. Egyes besorolással nem rendelkező géneket az UniProt adatbázisból vett funkcionális információk alapján osztályoztunk. Az inszerciós elemeket, profágokat és az ismeretlen funkciójú géneket a járulékos kategóriába soroltuk, míg a többit kihagytuk az elemzésből.

IV. Eredmények

1. A random, fitness-optimalizáló deléciók módszer követelményei

A célunk egy olyan módszer kifejlesztése volt, amelynek segítségével hatékonyan generálhatók random deléciókat többféle kombinációban tartalmazó *E. coli* sejtkönyvtárak, amelyekben aztán a legjobban növekedő klónok dúsulnak fel. Ehhez a következő feltételeknek kell teljesülnie: a deléciók kialakulásának helye és a deléciók mérete legyen véletlenszerű, ciklusonként minden egyes sejtben keletkezzen lehetőleg egy deléció, az eljárás nagy sejtpopuláción legyen alkalmazható. A rendszer elemei legyenek robusztusak az esetleges inaktíváló mutációkkal szemben, a folyamat pedig tartalmazzon növekedési ciklusokat, biztosítva a gyorsabb növekedésű klónok szelekcióját. A hatékonyabb tápanyag-felhasználásra nincsen jó szelekciós módszer, azonban feltételeztük, hogy a növesztési körülmények között felesleges gének eltávolítása erre is esélyt ad.

2. Előzetes kísérletek

A módszer kialakításához előzetes kísérletekben teszteltünk egy sor szóba jövő megoldást. A részletek ismertetése nélkül:

1. Elsőként egy célzott deléciók előállítását szolgáló módszert módosítottunk úgy, hogy véletlenszerűen generáljon génkiejtéseket. A λ Red rendszer segítségével, véletlenszerű kihagyással összefűzött genomi homológia boxokat – szelekciós markerrel együtt – rekombináltattunk a kromoszómába, majd indukáltuk a fragmentum delécióval járó, homológ rekombináció általi kivágódását, hasonlóan a csoportunkban korábban leírt módszerhez⁷⁰. A Keio törzsgyűjtemény⁷⁹ segítségével létrehozott, a homológia boxokat alkalmas sorrendben hordozó, fúziós klónkönyvtár azonban számos klónozási műterméket tartalmazott, így nem volt alkalmas a megbízható deléció-előállítására.

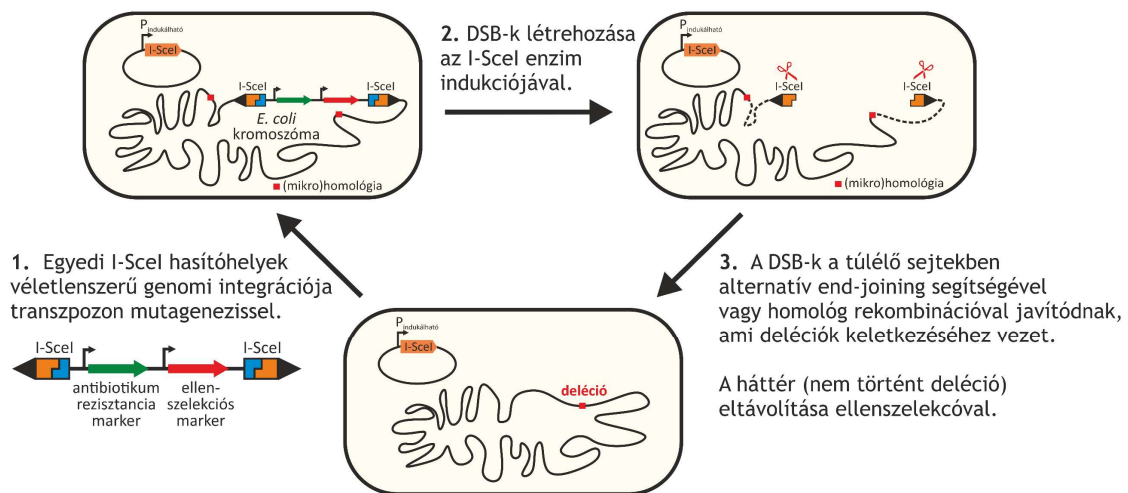
2. Másik rendszerre váltva, a Tn5 transzpozon véletlenszerű integrációján és ezt követő kivágódásán alapuló deléciós rendszer elemeit teszteltük számos variációban. Egy megvalósíthatósági kísérletben megállapítottuk, hogy a megfelelően módosított transzpozon random integrációt majd I-SceI hasítást követő kivágódása – ellenszelekció

alkalmazásával – hatékony eszköz lehet. Azt is megállapítottuk, hogy a genomi hasítás javításánál működő mechanizmusok (homológ rekombináció, illetve alternatív end-joining) nem részesítik előnyben a transzpozon végein található, rövid homológiákat (ez kizáró ok lenne), és a javítás eredményeként többféle méretű, a genom különböző szakaszai között létrejövő deléció keletkezhet.

3. Az előzetes kísérletek alapján a transzpozonos, genomhasításon alapuló megoldás potenciális nehézségeire is fény derült: (i) kellő számú transzpozon integráns sejt létrehozásához optimalizálni kell a folyamatot, (ii) hatékony, gyakorlatilag meghibásodásmentes ellenszelekciós módszerre van szükség, (iii) különös gondot kell fordítani a rendszer mutációkkal szembeni robusztusságára és a hamis pozitív sejtek kiküszöbölésére.

3. A RANDEL módszer

Az előzetes kísérletek tanulságai alapján alakítottuk ki azt a módszert (RANDEL: random deléciós séma), amely ciklikusan ismételtető formában alkalmas véletlenszerű deléciók készítésére *E. coli* sejtekben. A folyamat általános működését a 2. ábrán mutatjuk be. A módszer alkalmazása során először egy szelektálható markert építünk be véletlenszerű módon a genom egy pontjára. Ezután stimuláljuk az inszertált marker eltávolítását, és a ciklusok végén szelektálunk azokra a sejtekre, amelyek elveszítették a beépült DNS-darabot. Az inszertált fragmentum kiesése olyan DNS-hibajavító folyamatok beindulásával jár, amely az esetek többségében a beépülés közelében lévő genomi szakaszok delécióját eredményezik idegen szekvenciák visszamaradása nélkül. A vázolt beépülési és kiesési folyamat mindkét eleme szelektálható, amely lehetővé teszi, hogy az egyes ciklusok végén kapott sejtek mindegyike hordozzon egy-egy deléciót.



2. ábra. A random deléciós módszer általános sémája. DSB: double-stranded DNA break, kettős szálú DNS-törés. További részletekért lásd az 1. ábrát.

A ciklikus folyamat egyes lépései részleteiben a 1. ábrán láthatóak. A deléciók kialakulási helyének véletlenszerűségét biztosítja, hogy az első lépésben egy olyan transzpozont ugrasztunk be a kromoszómába, amely gyenge inszerciós hely preferenciával rendelkezik, így a beépülése véletlenszerűnek tekinthető⁸⁰. A beépített módosított Tn5 transzpozon szelekciós és ellenszelekciós markerek mellett a transzpozon végek közelében hordozza az I-SceI meganukleáz enzim felismerő helyeit. A transzpozon-inszerciós sejteket antibiotikum rezisztenciájuk alapján szelektáljuk. A következő lépésben egy plazmidról indukáljuk az I-SceI enzim kifejeződését. A termelődő endonukleáz a kromoszómára integrált transzpozon végein található felismerő helyeknél hasítja a kromoszómát. A keletkező dupla szálú DNS-törések (DSB: double-stranded DNA break) a sejt saját hibajavító mechanizmusai révén javítódnak. Ez végső soron a transzpozon elvesztéséhez vezet azáltal, hogy a DSB-k közelében található homológ szekvenciák között rekombináció történik, mivel a DNS templátot nem igénylő nem homológ végösszekapcsolás (non-homologous end joining) összetevői *E. coli*-ban nincsenek jelen⁸¹. Rekombináció történhet hosszabb homológ szakaszok között (>50 bp) a sejt saját RecA enzimének közvetítésével⁸², vagy néhány bp hosszúságú mikrohomológiák mentén, az ún. alternatív end-joining (AEJ) mechanizmus révén⁸³. Homológ szekvenciák, különösen a mikrohomológiák szerte a genomon megtalálhatóak, így egy adott deléció kiterjedésének mértéke véletlenszerűnek tekinthető. Az AEJ hibajavító mechanizmus hatékonysága viszonylag alacsony (kb. 100 000 DSB-t hordozó

sejt közül 1-ben történik javítás⁸³), ezért nagyon erős ellenszelekcióra van szükség ahhoz, hogy eltávolítsuk a populációból azokat a sejteket, amelyekből nem tűnt el a transzpozon. Emiatt kettős ellenszelekciót alkalmazunk a ciklusok során: I-SceI hasítást és a dP-hsvTK rendszert⁸⁴ használjuk azon sejtek kiválogatására, amelyek elvesztették a transzpozont. Itt az utóbbi ellenszelekciós rendszernek egy olyan meghibásodásbiztosnak tartott változatát alkalmaztuk, amelyben a *Herpes simplex* vírus timidin-kináz (HsvTK) egy duplikált génről expresszálódik, foszforilálja a kívülről adott dP nevű mesterséges nukleozid analógot, ezáltal letális mutagenézist okoz⁶⁸. Ezután a tenyésztési hőmérséklet emelésével eltávolítjuk az I-SceI enzimet kifejező hőmérséklet-érzékeny plazmidot, végül replika plating segítségével azonosítjuk azokat a sejteket, amelyek nem hordozzák sem a transzpozont, sem a segítő plazmidot.

A módszer felskálázhatóságának határait gyakorlati megfontolások szabják meg, különösen a replika plating lépés munkaigényes volta. A ciklusok során a „kimeneti” (deléciót hordozó) és a „bemeneti” (transzpozon-inszerciós) sejtek aránya nagyon alacsony, ezért számolnunk kell a rendszer elemeinek (szelekciós és ellenszelekciós gének, I-SceI gén, I-SceI felismerőhelyek) mutációs inaktivációjával. Ennek megakadályozására a rendszer komponenseit minden egyes ciklus elején frissen juttatjuk be, és a ciklusok végén replika plating segítségével megbizonyosodunk a segítő elemek sikeres eltávolítása felől. A módszer fontos sajátossága, hogy a beépített növesztési periódusok elősegítik a legnagyobb rátermettséggel rendelkező sejtek feldúsulását (folyékony táptalajban 5, szilárd táptalajon 4 növesztési periódus/ciklus, ami kb. 90 generáció növekedését jelenti ciklusonként).

4. A RANDEL módszer működési jellemzői

A deléciós módszer egyes lépéseinek hatékonyságát 5 egymást követő ciklus végigvitele során teszteltük. Ciklusonként tízezres nagyságrendben kaptunk számú transzpozon inszerciós mutáns sejtet. Replika plating segítségével megmutattuk, hogy az ellenszelekciót és a plazmid kiűzést követően az együttes transzpozon- és a plazmidvesztés aránya átlagosan kb. 60% volt. Minden egyes ciklus végén 100-1100 olyan egyedi sejtet nyertünk, amely nem tartalmazta sem a transzpozont, sem a helper plazmidot (2. táblázat). Az 1. ciklusban a pSTCST nevű I-SceI expressziós vektort

használtuk, de annak alacsony kiűzési hatékonysága miatt a 2. ciklustól kezdve a pSTAST plazmidra tértünk át.

2. táblázat. A RANDEL egyes lépéseinek hatékonysága MG1655 kiindulási sejteken végrehajtott 5 ciklus alkalmazása során.

Ciklus	Tn5 inszertánsok száma	I-SceI expressziós plazmid	Tn5-vesztés aránya	plazmid-vesztés aránya	Tn5- és plazmidvesztés együttes aránya	Tn5- és plazmidmentes sejtek száma
1.	13500	pSTCST	99,9%	48,4%	48,4%	810
2.	2900	pSTAST	68,6%	98,9%	67,6%	935
3.	6000	pSTAST	96,2%	99,9%	96,2%	1153
4.	21000	pSTAST	15,4%	96,3%	13,4%	132
5.	16000	pSTAST	64,3%	87,8%	53,5%	234

Ismert, hogy az ellenszelekcióhoz használt dP nevű nukleozid analóg mutagén hatású. A kezdeti protokollunk alkalmazása során egy időben történt az I-SceI hasítást követő DSB javítása és a dP közvetítette ellenszelekció. Ez változó számú pontmutáció (1-912/ciklus) kialakuláshoz vezetett a túlélő sejtek genomjában (3.a táblázat). A mutációs háttér lecsökkentése érdekében a későbbiekben módosítottuk a protokollt: az I-SceI vágást/javítást és a dP-mediált ellenszelekciót nem egyszerre, hanem egymást követően alkalmaztuk. Az I-SceI vágás/javítás után egy 5 órás időablakot illesztettünk be, ami lehetőséget biztosított a HsvTK (a dP-mutagenézisben szerepet játszó enzim) kiürülésére a dP-ellenszelekciós lépés előtt. Ez a módosítás a háttér szintjére csökkentette a pontmutációs rátát (0-2 mutáció/ciklus) (3.b táblázat). A genomok és fenotipikus hatások alább bemutatott elemzését mind a magas, mind az alacsony mutagenitású protokoll alkalmazása során nyert sejteken elvégeztük.

5. A RANDEL alkalmazásának eredményei

a) *Deléciók és deléciós mechanizmusok*

A ciklusok során bekövetkező genetikai változások azonosítása érdekében több kísérletsorozatban összességében 60 teljes genom szekvenálását végeztük el. (A

szekvenálásokat Nagy István és munkatársai hajtották végre a SeqOmics Kft.-nél.) Először olyan törzsvonalakat vizsgáltunk, amelyek növekedési paraméterei az előzetes mérések során nem bizonyultak rosszabbnak a vad típusnál. Ebből a csoportból 2 törzset szekvenáltunk a kontroll ciklusból (1 ciklus transzpozon-inszerció nélkül), 11-et az 1. ciklusból (c1), 10-et a 2. ciklusból (c2) és 2-t az 5. ciklusból (c5). Ezen kívül szekvenálásra került 11 véletlenszerűen, előzetes növekedésmérés nélkül kiválasztott törzs a c1-ből (3.a táblázat). Továbbá meghatároztuk 12 olyan egyetlen cikluson végigvitt vonal teljes genom szekvenciáját, amelyek esetében a kiindulási törzs a genomredukció egy előrehaladott állapotát képviselő, már számos szükségtelen genomszakaszt elvesztett MDS42 volt (4. táblázat, 6. táblázat). Mindemellett 12 olyan egy cikluson keresztülment törzset is szekvenáltunk, amelyek esetében a protokoll csökkentett mutációs rátájú változatát használtuk (3.b táblázat). A szekvenálásra kijelölt törzsek növekedési paramétereit ismételtelen lemérve megerősítettük (Függelék 2. táblázat).

3. táblázat. Az erősen mutagén (a) és a gyengén mutagén (b) protokollok felhasználásával kapott deléciók, valamint háttér SNV (single nucleotide variant, egyedi nukleotidváltozat) számok MG1655 alapú törzsekben. A szekvenálásra került törzseket jó növekedési sajátságaik alapján (a növekedési paramétereik nem rosszabbak, mint a vad típusé) vagy véletlenszerűen választottuk ki. A piros téglalapok az adott deléció jelenlétét mutatják.

a

Ciklusok száma	Kiválasztás alapja	Törzs	Deléció												SNV-szám	
			D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12		
1 kontroll ciklus	jó növekedés	contr-56														2
		contr-59														2
1 ciklus	jó növekedés	c1-16														5
		c1-17														54
		c1-53														16
		c1-30														79
		c1-21														57
		c1-23														97
		c1-24														64
		c1-48														56
		c1-72														53
		c1-29														25
	c1-32														37	
	random	c1-5														795
		c1-9														228
		c1-10														856
		c1-11														51
		c1-12														845
		c1-15														1
		c1-6														912
c1-14															849	
c1-7														226		
c1-8														80		
c1-13														267		
2 ciklus	jó növekedés	c2-31													123	
		c2-17													138	
		c2-36													107	
		c2-23													62	
		c2-11													70	
		c2-8													60	
		c2-12													71	
		c2-20													79	
		c2-26													63	
c2-32													63			
5 ciklus	jó növekedés	c5-158												418		
		c5-170												325		

b

Ciklusok száma	Kiválasztás alapja	Törzs	Deléció												SNV-szám
			D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	
1 ciklus	jó növekedés	c1_lowmut-80													0
		c1_lowmut-87													2
	random	c1_lowmut-51													0
		c1_lowmut-52													0
		c1_lowmut-53													0
		c1_lowmut-54													1
		c1_lowmut-56													0
		c1_lowmut-57													2
		c1_lowmut-58													2
		c1_lowmut-59													1
		c1_lowmut-55													1
		c1_lowmut-60													1

4. táblázat. Egy ciklus után kapott deléciók és háttér SNV-számok MDS42 alapú törzsekben. A piros téglalapok az adott deléció jelenlétét mutatják.

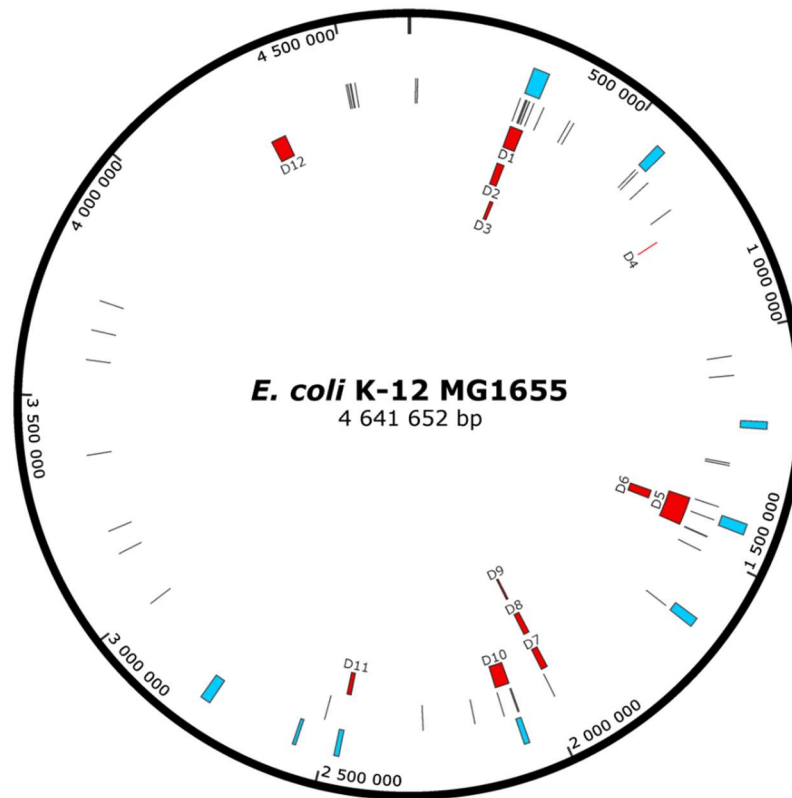
Törzs	Deléció			SNV-szám
	D13	D14	D15	
c1_MDS42_1				0
c1_MDS42_2				0
c1_MDS42_3				0
c1_MDS42_4				0
c1_MDS42_5				1
c1_MDS42_6				0
c1_MDS42_8				1
c1_MDS42_10				0
c1_MDS42_18				0
c1_MDS42_7				60
c1_MDS42_14				13
c1_MDS42_16				42

A megfigyelt deléciók száma nem minden esetben egyezett az elvégzett ciklusok számával (egy deléció/ciklus). Az MG155 törzsből származtatott vonalak 13%-a (6/48) a vártnál eggyel kevesebb deléciót tartalmazott. Ez a megfigyelés a transzpozon nyomtalan kivágódásával magyarázható: transzpozon-inszerció során a beépülés helyén a transzpozonvégek mellett egy 9 bp méretű duplikáció keletkezik⁸⁵, így előfordulhat, hogy az AEJ javító mechanizmus ezeknek a rövid homológiáknak a segítségével kapcsolja össze a DNS-végeket. Ezt a felvetést alátámasztja, hogy az MDS42 törzsből kiinduló vonalakban magasabb a feltételezhetően nyom nélkül végbemenő deléciók aránya (75%; 9/12), mivel a redukált genomú szülői törzsből el lettek távolítva azok a hosszú ismétlődő szekvenciák, például az IS elemek, amelyek elősegítik a DNS-törés javítását.

A genomok elemzése során arra is fény derült, hogy az egy cikluson keresztülment törzsek jelentős része (29%; 10/34) a vártnál eggyel több deléciót hordoz, amelyet tovább is visznek a következő ciklusokba. Ennek oka, hogy a deléció két változata (D7 és D8) a flagelláris génklasztert érinti, amely spontán módon, a DSB/javítás hozzájárulása nélkül is gyakran elveszik (elemzését lásd még lentebb).

Mindent egybevetve az egy- és többciklusos törzsekben 12 különböző deléciót figyeltünk meg, ezek mérete a 1 871 bp és 70 678 bp közötti tartományba esik (5. táblázat, 3. ábra). A deléciók határainak vizsgálata a kiesés mechanizmusáról is információt ad. A DSB-k javítódásának szerkezeti követelményei szélesek: a deléciók

határai lehetnek inszerciós szekvenciák (IS-ek), fág attachment helyek, REP (repetitive extragenic palindrome) elemek, gén-pseudogén párok, különböző gének homológ régiói. A homológiák mérete a 11 bp-os mikrohomológiáktól az 1 kbp feletti IS régióig terjed (5. táblázat).



3. ábra. A munka során kapott random deléciók helye a vad típusú *E. coli* MG1655 genom térképén. A kék téglalapok a kriptikus profágok pozícióját, az egyenes vonalak az inszerciós szekvenciákat, a piros téglalapok pedig a random deléciókat jelölik (D1-től D12-ig). A deléciók magyarázata az 5. táblázatban található.

A kezdeti, első lépésben (34 szekvenált egyciklusos genom) 6 különböző deléciót figyeltünk meg. Ezek közül kettő kriptikus profágokat érintett (Rac, CPS-53); ezek deléciója többször is előfordul az elemzett törzsekben. Érdekes módon a flagelláris gének kiesése több alkalommal is megfigyelhető az első ciklus után, sőt, még a kontroll ciklus (nincs transzpozon inzerció és I-SceI hasítás) során is létrejött.

A flagelláris deléció feltűnő gyakorisága miatt az 5 egymást követő ciklus kiindulási törzseként használt MG1655 stock-ban is megvizsgáltuk a flagelláris régió jelenlétét. Motilitási teszt alapján azt találtuk, hogy a sejtek 25%-a már eleve elvesztette

ezt a genomi régiót. A további kísérletek azt mutatták, hogy ezek a sejtek kompetitív előnnyel bírnak, és ezért egy ciklus után 25%-ról 75%-ra dúsultak fel a populációban.

5. táblázat. Az MG1655 szülői törzsön végrehajtott random deléciós ciklusok során kapott deléciók tulajdonságai. A genomi koordináták az MG1655 genomra vonatkoznak (NCBI azonosító: NC_000913.3).

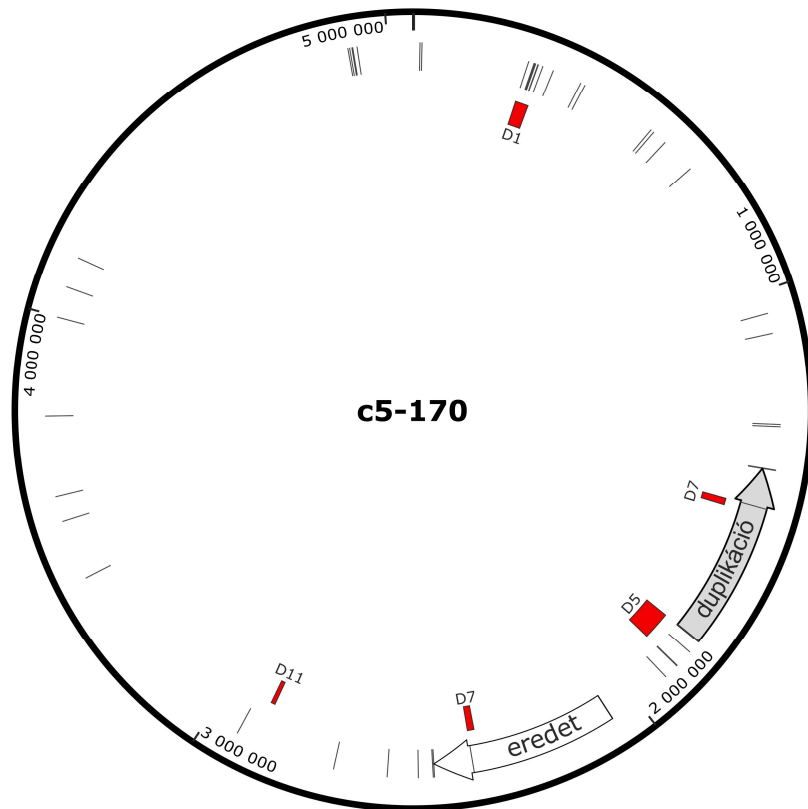
Deléció	Bal oldali határ	Jobb oldali határ	Méret (bp)	Leírás	A kialakulás feltételezett mechanizmusa
D1	258 167	290 892	32 726	két IS1 között található genomi szegmens (magában foglalja a prolin bioszintézis géneket és a CP4-6 profág részleges delécióját)	rekombináció az IS1I és az IS1C homológ régiói között
D2	258 607	279 861	21 255	két IS1 között található genomi szegmens (magában foglalja a CP4-6 profág részleges delécióját)	rekombináció az IS1I és az IS1B homológ régiói között
D3	279 876	291 346	11 471	genomi szegmens az <i>rhsC</i> és az <i>rhsO</i> gének között	rekombináció az IS1B és az IS1C homológ régiói között
D4	732 605	734 475	1 871	két IS2 között található genomi szegmens (magában foglalja a Rac profág delécióját)	rekombináció az <i>rhsC</i> és az <i>rhsO</i> gének homológ régiói között
D5	1 397 236	1 467 913	70 678	Rac profág	rekombináció az IS2E (előfordulása ebben a genomi lókuszban egy duplikáció eredménye) és az IS2D homológ régiói között
D6	1 411 899	1 434 958	23 060	flagelláris és kemotaxis gének (az IS1 szekvenciát határolják)	kivágódás a Rac profág attachment helyénél
D7	1 962 083	1 978 502	16 420	két IS5 között található genomi szegmens (magában foglalja a hisztidin bioszintézis géneket és a CP4-44 profágot)	rekombináció az IS1H szekvencia replikatív transzpozícióját követően
D8	1 962 204	1 978 502	16 299		
D9	1 972 841	1 978 502	5 662		
D10	2 066 704	2 102 294	35 591	REP elemek közötti genomi szegmens	rekombináció az IS5H és az IS5I homológ régiói között
D11	2 466 369	2 476 583	10 215	CPS-53 profág	rekombináció az <i>argW-pawZ</i> gén-pszudogén pár homológ régiói között a CPS-53 profág attachment helyénél
D12	4 285 317	4 323 260	37 944		rekombináció a REP320 és a REP324 homológ régiói között

6. táblázat. Az MDS42 szülői törzsön végrehajtott egyetlen random deléziós ciklus során kapott deléciók tulajdonságai. A genomi koordináták az MDS42 genomra vonatkoznak (NCBI azonosító: NC_20518.1).

Deléció	Bal oldali határ	Jobb oldali határ	Méret (bp)	Leírás	A kialakulás feltételezett mechanizmusa
D13	266 028	301 308	35 281	REP elemek között található genomi szegmens (magában foglalja a lac operon delécióját)	rekombináció a REP23 és a REP31 homológ régiói között
D14	266 090	283 922	17 833	REP elemek között található genomi szegmens	rekombináció a REP23 és a REP28 homológ régiói között
D15	688 716	695 120	6 405	genomi szegmens a <i>ybhR</i> gén és a RIP75 elem között	nem ismert, a deléció közelében nem található homológ szekvenciák

b) Pontmutációk és egyéb változások

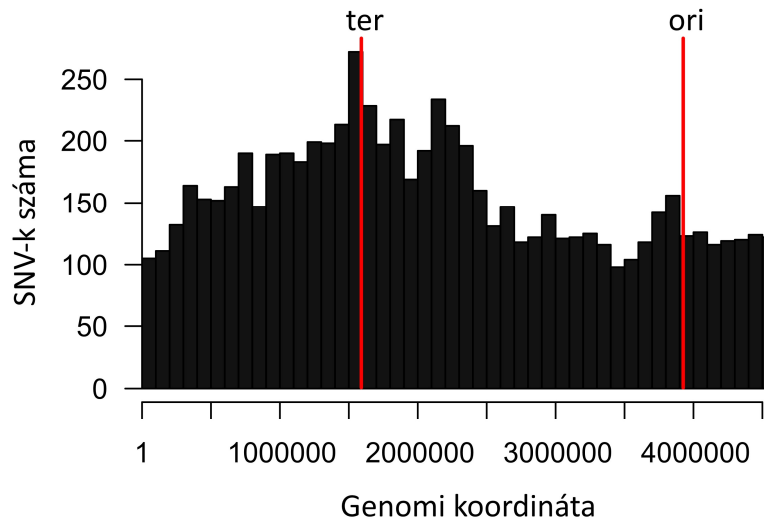
A kezdeti, erősen mutagén protokoll használatának mellékhatásaként minden ciklusban megfigyeltünk genomként átlagosan néhány tíz egyedi nukleotid változatot (single nucleotide variant, SNV). A két 5. ciklus végén kapott törzsben az SNV-k száma százak nagyságrendű (418 és 325), az egyikben egy nagyméretű duplikációt is azonosítottunk a pontmutációk mellett (4. ábra).



4. ábra. A duplikáció pozíciója az egyik 5. ciklus után kapott random deléziós törzsben (c5-170). A vad típusú MG1655 genom térképét kiegészítettük a 398 kb hosszúságú duplikációval, amit szürke nyíl mutat. A duplikálódott régió forrását fehér nyíl jelzi. A vékony egyenes vonalak az inszerciós szekvenciákat, a piros téglalapok a random deléciókat reprezentálják (D1, D5, D7, D11). A deléciók magyarázata a 6. táblázatban található.

A pontmutációk nagy száma váratlan volt, az irodalmi adatok alapján kevesebb változásra számítottunk. Tominaga és munkatársai szerint a dP nukleozid analóg alkalmazása az általunk használt koncentrációban csak tízszeresére emeli a spontán mutációs gyakoriságot⁶⁸.

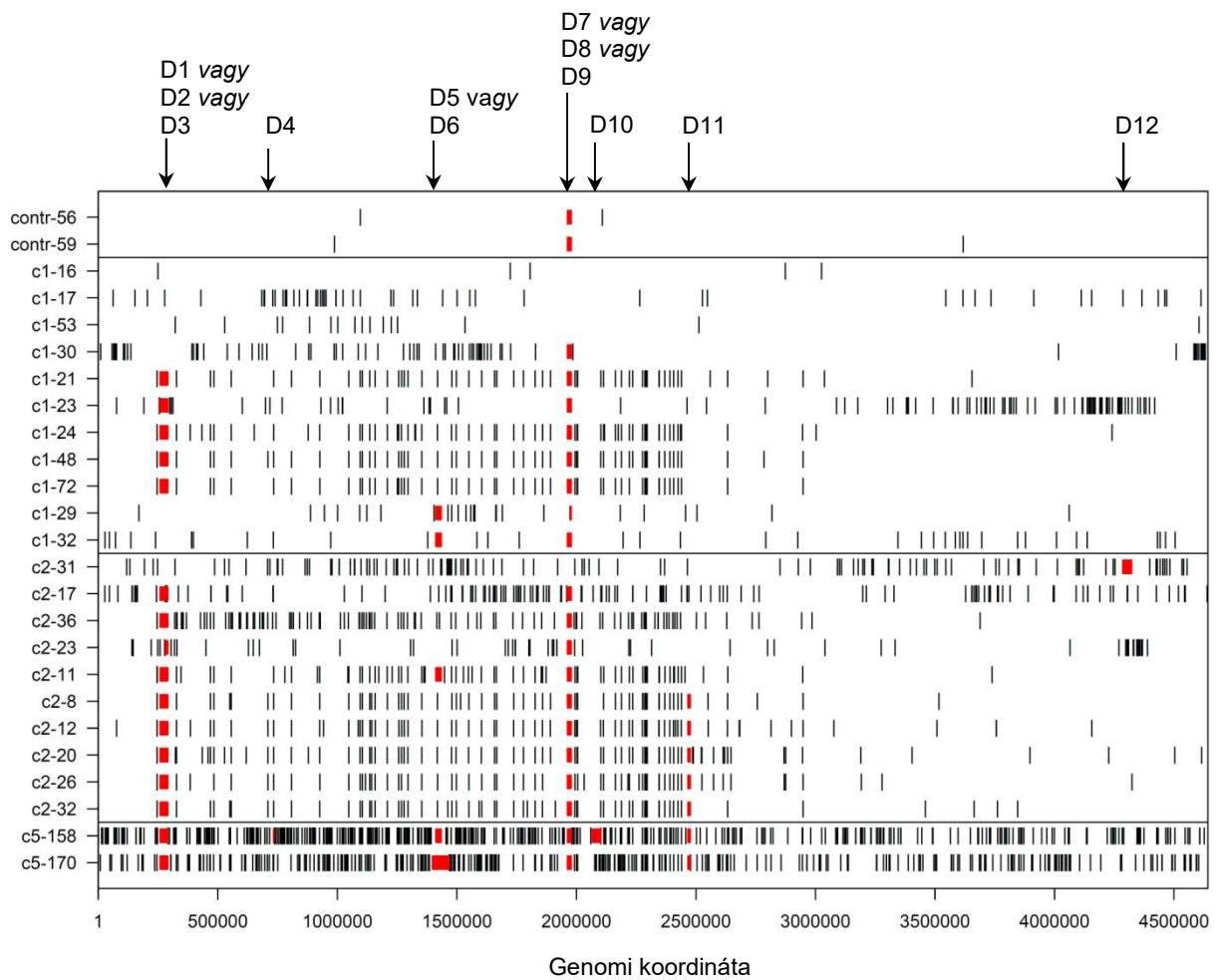
Az SNV-k nem azonos gyakorisággal fordultak elő a kromoszóma különböző pontjain: a replikációs terminusnál mintegy kétszer annyi pontmutációt találtunk, mint a replikációs origó közelében. Emellett az SNV-szám hullámszerűnek tűnő sűrűsödése és ritkulása figyelhető meg a kromoszóma mentén (5. ábra).



5. ábra. Az SNV-k eloszlása a genomon belül az erősen mutagén protokoll alkalmazása során kapott törzsekben. Az ábrán a 36 szekvenált vonalban előforduló, összesen 7237 db SNV eloszlását tüntettük fel. Az oszlopok a ~100 kbp méretű genomi szakaszra eső SNV-k számát mutatják. Piros egyenesekkel a replikációs origót (ori) és terminust (ter) jelöltük.

c) A szelekció hatásának jelei

Az SNV-k eloszlásából (6. ábra) következtethetünk arra, hogy a sejtvonalak közös eredetűek vagy egymástól függetlenül alakultak ki. Hasonló SNV-mintázatot láttunk 4 első ciklusos és 5 második ciklusos genomban a „jól növvő” (a szülői törzsnél nem rosszabb növekedési paraméterekkel rendelkező) törzsek között, ami az előnyös deléciós változatok szelekciójára utal. Az eltérő SNV-mintázatok viszont azt jelenthetik, hogy a több törzsben is megfigyelt deléciók, mint a Rac profág (D6), CPS-53 profág (D11) és különösen a flagelláris régió (D7, D8, D9) kiesése párhuzamosan, függetlenül is megtörténhet az egyes vonalakban. A bizonyítékok tehát azt mutatják, hogy a szelekció eredményeképp bizonyos genomi régiók delécióját nagyobb gyakorisággal figyelhetjük meg, és az előnyös deléció-pontmutáció kombinációk gyorsan fixálódnak a populációban.



6. ábra. Az SNV-k eloszlása a kontroll törzsek (egy cikluson mentek keresztül transzpozon inszerció nélkül) és a – mutagén eljárással készült – deléciós törzsek genomjában. A piros téglalapok a deléciókat (D1-től D12-ig), a függőleges vonalak az SNV-ket jelölik.

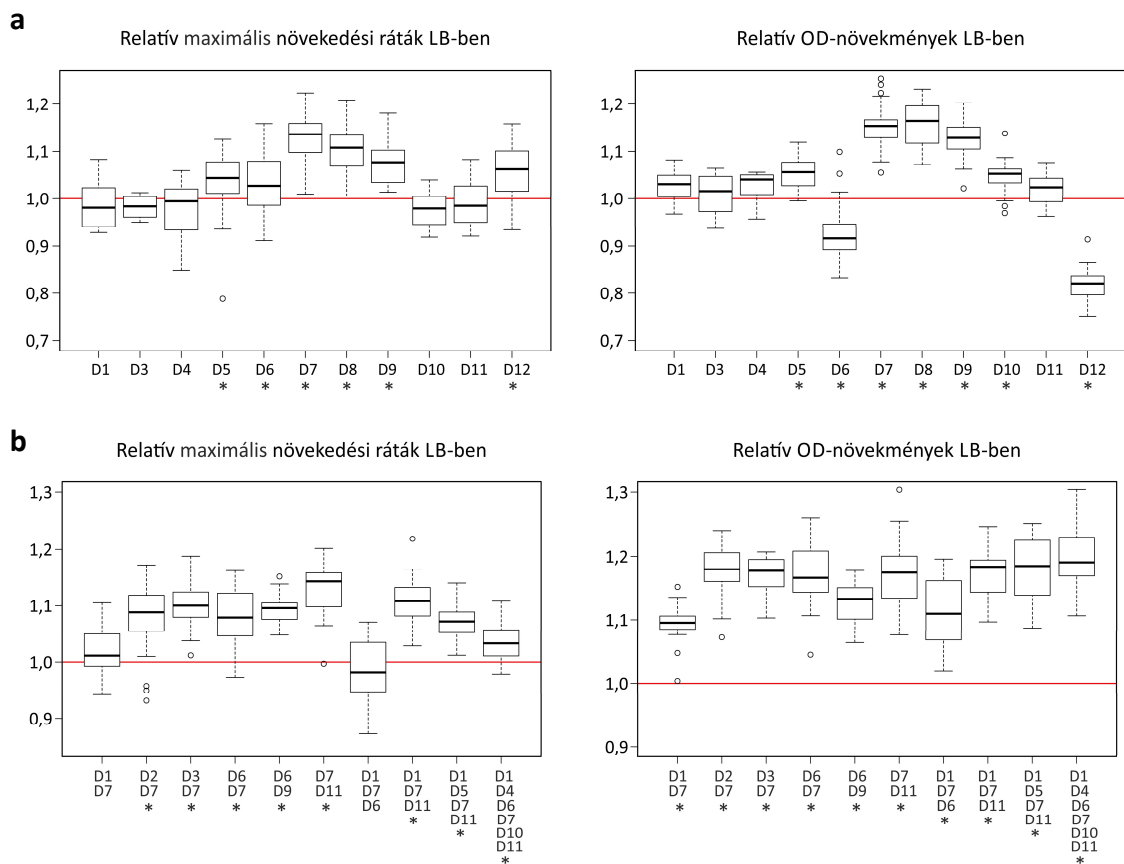
6. A deléciók és pontmutációk fenotipikus hatásainak elkülönítése

a) A rekonstruált deléciók hatása

Ahhoz, hogy a deléciók funkcionális következményeit a pontmutációk zavaró hatása nélkül tudjuk tanulmányozni, mutációktól mentes szülői háttéren rekonstruáltuk a ciklusok során kapott egyedi deléciókat és azok kombinációit (5. táblázat). (A törzsek rekonstrukciója Karcagi Ildikó jelentős hozzájárulásával készült.) Ehhez egyesével, illetve

gyakorlatilag az összes tapasztalt kombinációban előállítottuk a deléciókat precíz, „öngyilkos” plazmid alapú genommodosító módszerrel az eredeti, MG1655 gazdában.

A rekonstruált, pontmutáció-mentes háttérű törzsek növekedési ráta és optikai denzitás (OD) növekmény (a biomassa hozam közelítő mérőszáma) vad típushoz viszonyított értékeit a 7. ábrán foglaltuk össze. Bizonyos deléciók nem befolyásolták a növekedési paramétereket, viszont 6 deléció (az azonosított deléciók 50%-a), amely lényegében három különböző típust képvisel (flagelláris régió, Rac, D12), a növekedési ráta szignifikáns emelkedését okozta. Különösen erős hatása volt annak a deléciónak, amely működésképtelenné tette a flagelláris apparátust. Abban az esetben is magasabb növekedési rátát és OD-növekményt mértünk a vad típushoz képest, amikor 5 másik deléció volt mellette. A lágy agarban elvégzett motilitási tesztek megerősítették a flagelláris deléciókkal járó úszásképtelen fenotípust.



7. ábra. Vad típusú háttéren rekonstruált deléciókat tartalmazó egyszeres (a) és többszörös (b) deléciós törzsek növekedési paramétereit. A törzseket az általuk hordozott deléciók szerint neveztük el (D1-től D12-ig). A maximális növekedési ráta és OD-növekmény értékeit a vad típusú M1655 törzshöz hasonlítottuk. A méréseket gazdag táptalajban (LB) végeztük. A feltüntetett értékek legalább 18 párhuzamos mérés

eredményei. A dobozábrákon a vastag vízszintes vonalak a mediánt mutatják, a dobozokat a 25. és 75. percentilisek határolják, a függőleges vonalak a kvartilisek közötti távolság 1,5-szereséig nyúlnak ki, az ezen kívül eső értékeket körök reprezentálják. A csillagok a vad típus értékeihez viszonyított szignifikáns különbséget jelzik (Welch-próba, $P < 0,05$).

Néhány kiválasztott egyedi és többszörös deléciós vonal MG1655 szülői törzshöz viszonyított fitnessét kompetíciós kísérletek segítségével határoztuk meg (7. táblázat). Fontos eredmény, hogy a legtöbb deléciót hordozó (6 deléció, beleértve a flagelláris régió kiesését is), 2,5%-os genomcsökkenést mutató törzs LB médiumban kiszorította a vad típusú törzset (7. táblázat). Ennek a törzsnek nemcsak a fitnessze, hanem a biomassza hozama is előnyös változást mutatott. A törzs száraz tömeg hozama 6%-kal bizonyult magasabbnak a szülői MG1655 értékénél (8. táblázat).

7. táblázat. Deléciós törzsek kompetitív fitnessze a vad típusú MG1655 szülői törzshöz viszonyítva.

Törzs	Relatív fitnessz (a szülői típushoz képest)	P érték
D6	1,00 (0,008)	0,5031
D7	1,11 (0,030)	0,0243
D1_D4_D6_D7_D10_D11	1,03 (0,006)	0,0028
D1_D4_D6_D7_D10_D11 (c5-158)	0,92 (0,009)	0,0043

A törzseket az általuk hordozott deléciók nevével azonosítottuk. A D6, D7 és D1_D4_D6_D7_D10_D11 tiszta szülői genetikai háttéren rekonstruált deléciós törzsek. A D1_D4_D6_D7_D10_D11 (c5-158) törzset a mutagén random deléciós ciklusok során kaptuk. A feltüntetett átlag és szórás (zárójelben) értékek legalább 3 független kísérlet eredményein alapulnak. A kompetíciós kísérleteket gazdag táptalajban (LB) végeztük. Az egymintás t-próbák megmutatták, hogy a c5-158 törzs szignifikáns kompetitív hátrányban, míg a D7 és D1_D4_D6_D7_D10_D11 szignifikáns kompetitív előnyben van a vad típushoz képest ($P < 0,05$).

8. táblázat. A rekonstruált 6-deléciós törzs és a vad típusú MG1655 biomassza hozama.

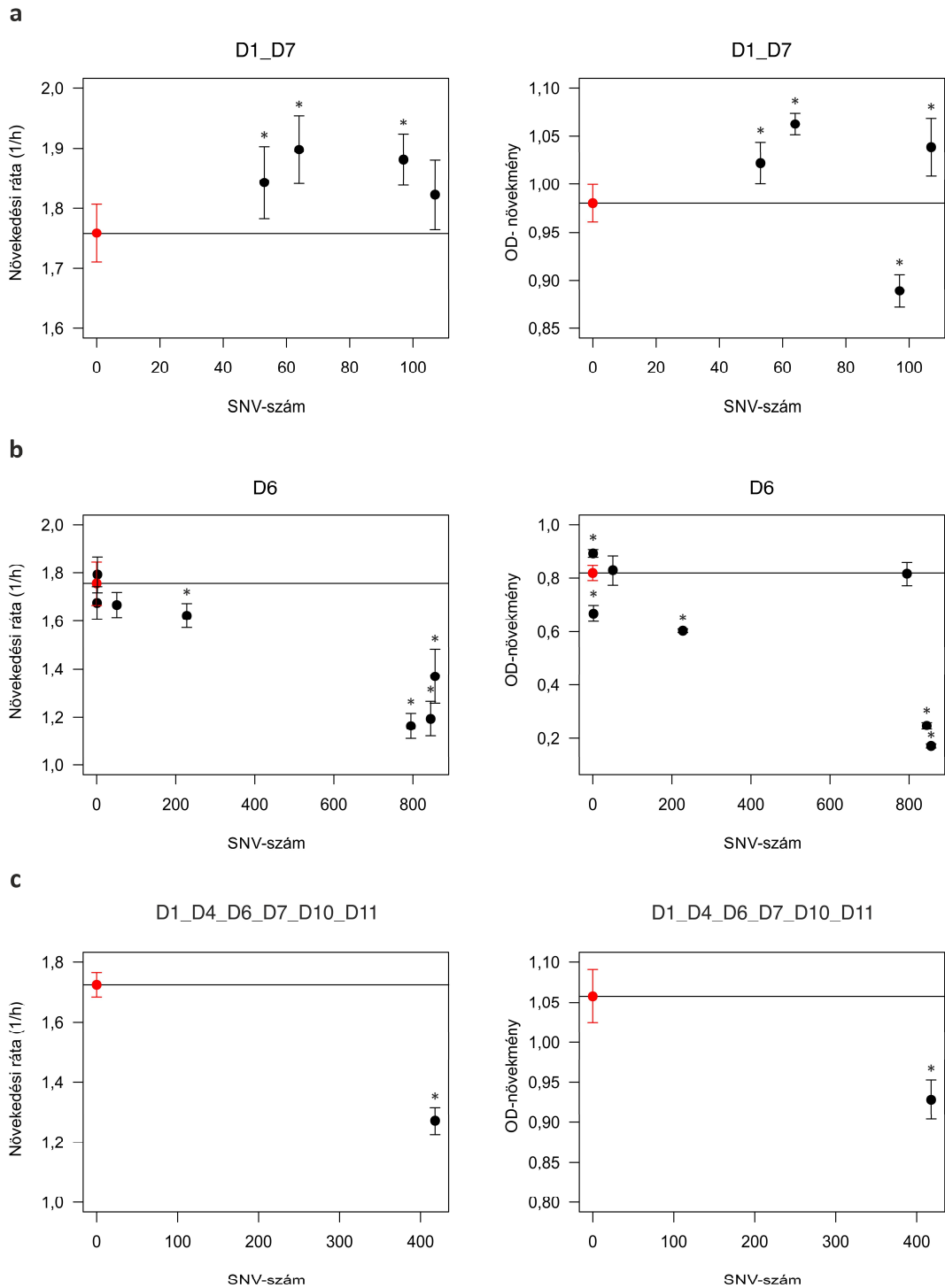
Törzs	Száraz tömeg (g/l)
MG1655	1,83 (0,034)
D1_D4_D6_D7_D10_D11	1,95 (0,064)

Gazdag táptalajban (LB) meghatároztuk az egy éjszakán át, 37 °C-on növesztett lombikos kultúrák száraz tömegét. A feltüntetett átlag és szórás (zárójelben) értékek 5 ismételt mérés eredményén alapulnak. A Welch-próba megmutatta, hogy a 6-deléciós rekonstruált D1_D4_D6_D7_D10_D11 törzs biomassza hozama szignifikánsan magasabb a vad típushoz képest ($P=0,011$).

A genomredukciót eredményező random deléciós ciklusokat gazdag táptalajban végeztük. A folyamat során létrejött előnyös deléciók szelektálódása eltérő körülmények között ellentétes irányú hatásokkal járhat: a D1, D2 és D10 deléciókat hordozó sejtek minimál médiumban életképtelenek az aminosav bioszintézisben részt vevő gének eliminációja miatt. A D1 és D2 delécióval a prolin, a D10 delécióval pedig a hisztidin szintéziséért felelős gének estek ki (5. táblázat).

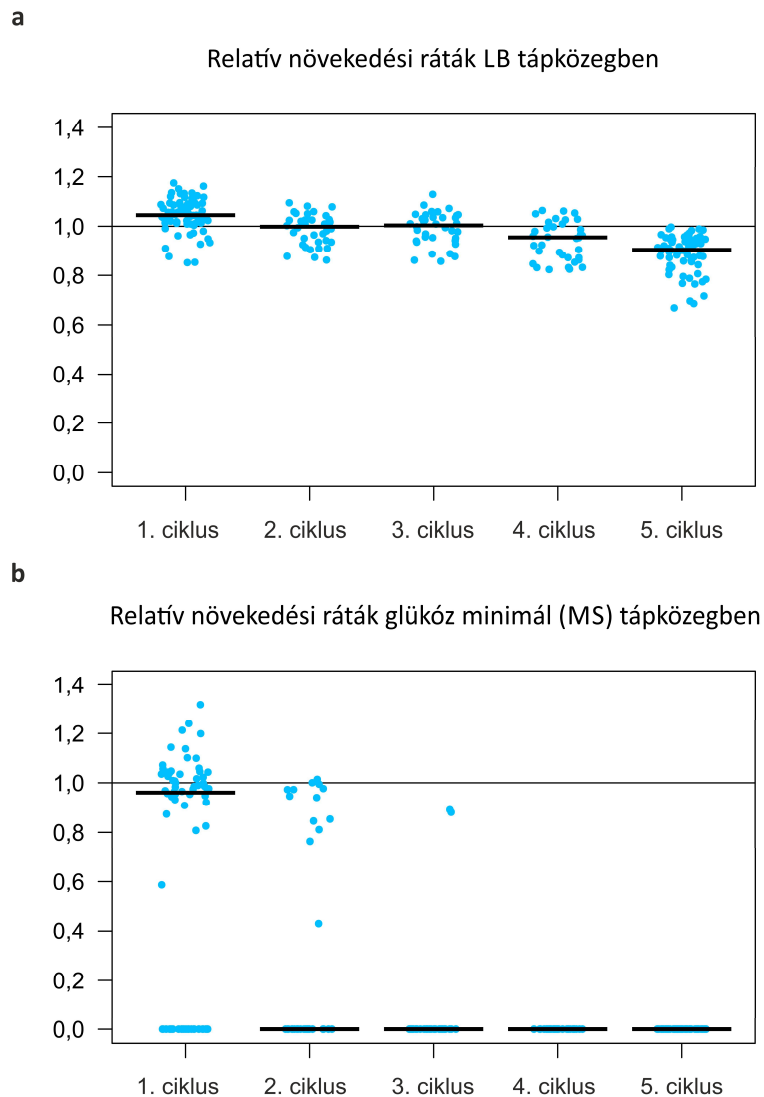
b) A pontmutációk általános hatása

A szekvenált genomokban azonosított SNV-k száma tág határok között változott (0-912/ciklus, 5 ciklus után legfeljebb 418 SNV). Ha összehasonlítjuk egy adott deléciós változat különböző számú SNV-vel rendelkező változatait (pl. nulla SNV a rekonstruált genomokban, változó számú SNV az ugyanazon deléciót hordozó, függetlenül létrejött vonalakban), akkor a következő, nem meglepő tendencia rajzolódik ki: alacsony mutációs ráta mellett keletkezhetnek a fitnesszt esetlegesen növelő pontmutációk, ugyanakkor a nagy pontmutációs háttér káros hatású, ami a növekedési paraméterek jelentős romlásában mutatkozott meg (8. ábra). Ezenfelül az erősen mutagén protokoll alkalmazása során kapott törzsek életképessége jelentősen csökken egy eltérő niche-ben, glükóz minimál médiumban. (9. ábra).



8. ábra. Egy 2-deléciós (a), egy 1-deléciós (b) és egy 6-deléciós (c) törzs LB táptalajban mért növekedési ráta és OD-növekmény adatai. A random deléciós ciklusok során kapott, különböző számú SNV-t tartalmazó vonalak növekedését a rekonstruált, pontmutáció-mentes párjaikhoz hasonlítottuk. A törzseket az általuk hordozott deléciók alapján neveztük el (D1-től D12-ig). A pontok az adott növekedési paraméter 5 ismételt mérés alapján számított átlagát jelölik. Piros szín mutatja a pontmutációkat nem tartalmazó rekonstruált törzs értékeit. A hibavonalak 95%-os konfidencia-intervallumot

jelentenek. A csillagok a növekedési paraméter értékben szignifikáns különbséget jelölnek az adott törzs pontmutációktól mentes változatához képest (Welch-próba, $P < 0,05$).

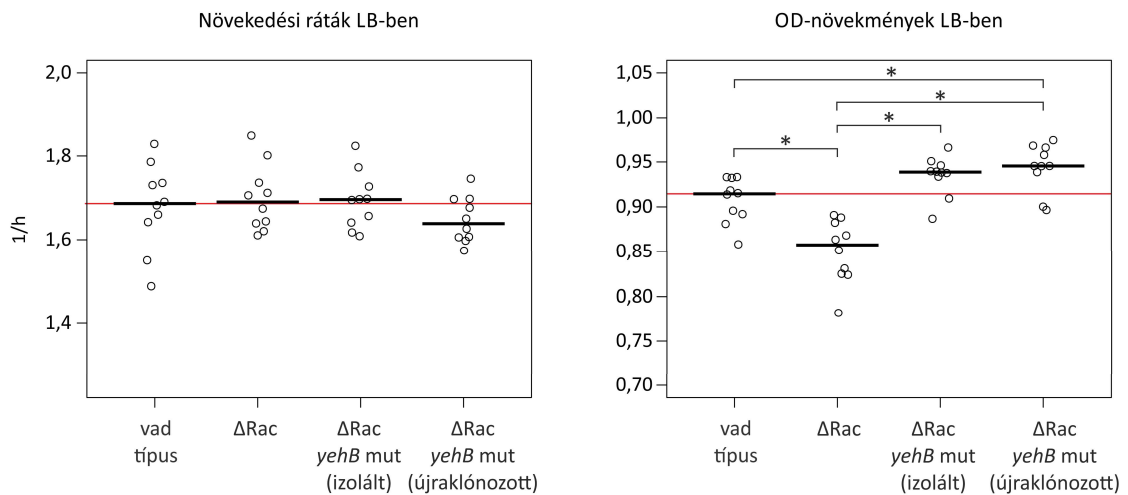


9. ábra. Az erősen mutagén protokoll alkalmazása során kapott törzsek növekedési rátái. A méréseket (a) gazdag (LB) és (b) glükóz minimál (MS) táptalajban egyaránt elvégeztük. Ciklusonként 30 véletlenszerűen kiválasztott telep növekedését teszteltük. A növekedési rátákat minden egyes ciklus sejtjeinek mérése során a szülői MG1655 törzs mért értékének átlagához viszonyítottuk. A pontok az egyedi törzsek előbbieik szerint számított relatív növekedési rátáját mutatják.

c) Egy kompenzáló pontmutáció

Érdekes módon a Rac profág deléciójának rekonstruált, mutációmentes háttéren létrehozott változata a vad típusénál alacsonyabb OD értéket mutatott, míg a random

deléciós ciklusok során izolált változat – amely a Rac delécióon kívül egyetlenegy aminosavcserét okozó pontmutációt tartalmazott, a *yehB* génben – nem mutatott csökkenést. A pontmutációt a „tisztá” háttérű Rac deléció mellett rekonstruálva bizonyítottuk ennek a *yehB* pontmutációnak a fitnesscsökkenést kompenzáló hatását (10. ábra). (A mutáns *yehB* allél létrehozását Tímár Edit végezte.)

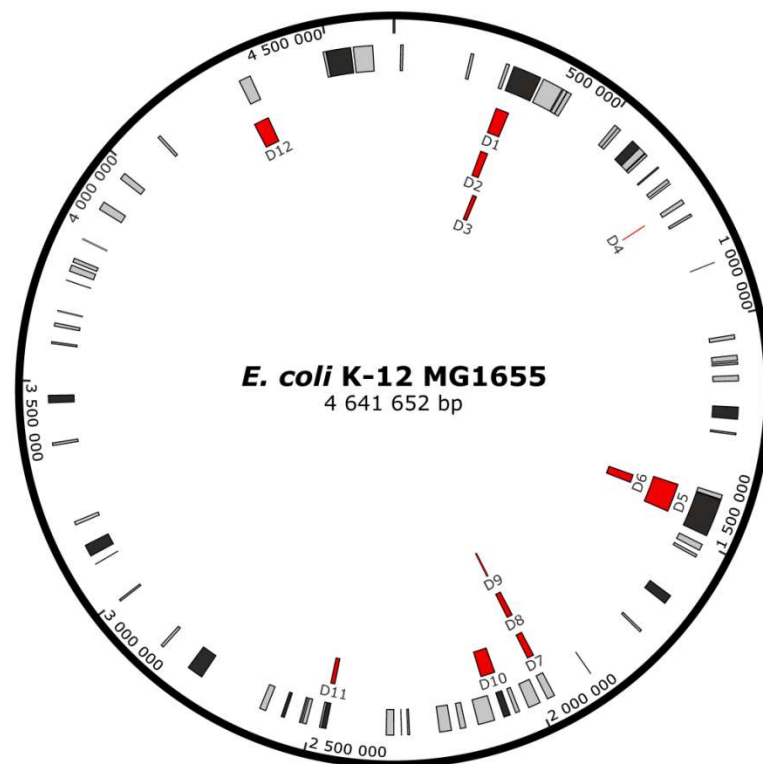


10. ábra. Egy kompenzáló pontmutáció hatása a Rac deléciós törzsben: a vad típusú és a Rac deléciós törzs növekedési paraméterei vad típusú és egy pontmutációt tartalmazó *yehB* allállal. A vízszintes fekete vonalak 10 mérés mediánját mutatják. A csillagok szignifikáns különbséget jelölnek a törzsek között ($P < 0,05$, egyutas ANOVA post hoc Tukey-teszttel).

7. A célzott és random módszerekkel létrehozott deléciók összehasonlítása

A csoportunk által korábban publikált genomredukciós munka alapja a racionális tervezés volt^{49,50,70}. Az eltávolított genomi szakaszok kijelölése komparatív genomikai analízisek (az *E. coli* törzscsoporton belüli hasonlóságok, illetve a jól körvonalazódó, egyedi genomi szigetek, horizontális géntranszferrel bekerült szakaszok, profágok, IS elemek azonosítása), gén esszencialitás (gazdag és minimál táptalajban nélkülözhetetlen gének) és funkcionális genomikai vizsgálatok (irodalmi adatok alapján a fitnesszt jelentősen nem befolyásoló, illetve ismeretlen funkciójú gének) eredményei alapján történt. Ha összehasonlítjuk a tudatos tervezéssel létrehozott MDS12⁷⁰ mérföldkő törzs 12 célzott delécióját (összesen 423 gén/376 kbp) a jelen munkában kapott 12 random delécióval (összesen 218 gén/206 kbp), akkor érdekes módon csak 7 egyezést találunk a deléciók között (11. ábra). Ez a megfigyelés alátámasztja a

különbségeket a racionális tervezés és a véletlenszerű, szelekció által hajtott genom „áramvonalasítás” között. Az MDS12 törzshöz viszonyítva kiterjedtebb deléciók készlettel rendelkező MDS69⁵⁰ mérőldkő törzs (69 célzott deléció, összesen 965 gén/942 kbp) deléciói viszont nagymértékben átfednek az általunk kapott 12 random delécióval (11. ábra). Bizonyos deléciók meghosszabbításánál láthatunk néhány kisebb eltérést, ami összességében 43 új gén megjelenését jelenti a random deléciók készletben az MDS69 törzshöz viszonyítva.

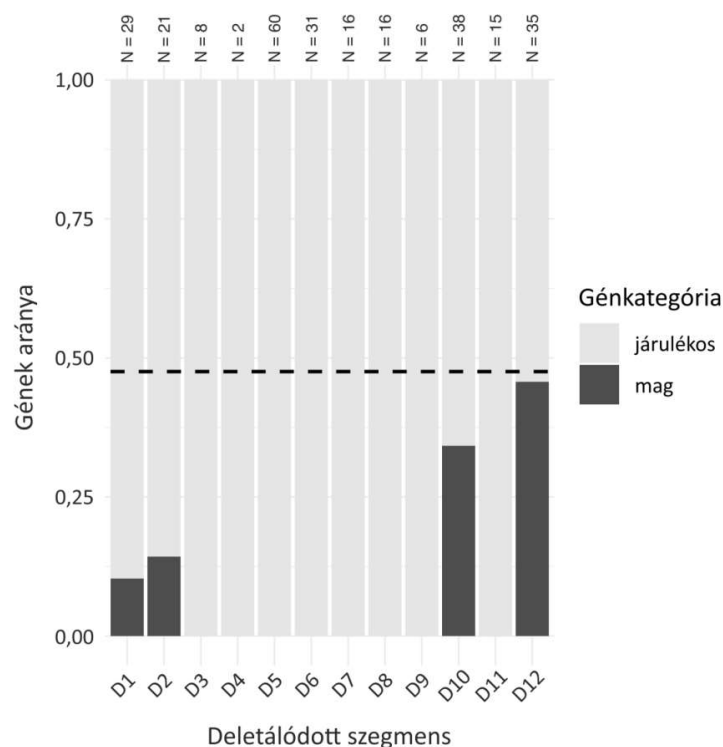


11. ábra. A célzott és random módszerekkel kapott deléciók összehasonlítása. A piros téglalapok az ebben a munkában azonosított random deléciókat mutatják (D1-től D12-ig), a fekete téglalapok az MDS12 nevű multideléciók törzs célzott delécióit jelölik, a szürke téglalapok pedig az MDS69 törzs előbbiei mellett jelen lévő delécióit reprezentálják.

8. A deletálódott genomi régiók és a maggenom viszonya

Végezetül pedig azt a kérdést tettük fel, hogy a véletlenszerűen deletálódott genomi szakaszok nélkülöznek-e az olyan géneket, amelyek az *E. coli* ún. maggenomjához (core genom) tartoznak (azaz az olyan géneket, amelyek az *E. coli* klád összes törzsében

megtalálhatóak, lásd még az Anyagok és módszerek fejezetben). (Az elemzést Papp Balázs és Györkei Ádám végezte.) A várttal összhangban a legtöbb deletálódott szegmens egyáltalán nem, vagy csak igen kis számban tartalmaz core géneket (12. ábra), ami arra utal, hogy a módszerünk alkalmazása többnyire a járulékos gének kiesését eredményezi. Azonban meglepő módon találtunk két olyan régiót, a D10 és a D12 jelűt, amelyben a core gének gyakorisága megközelíti a genomi átlag értékét (12. ábra). Ez a megállapítás azt mutatja, hogy a szelekción alapuló genomredukciós módszerünk célpontjai olyan genomi szakaszok is lehetnek, amelyek összehasonlító genomikai megközelítéssel nem lennének előrejelezhetőek.



12. ábra. A deletálódott genomi szegmenseken belül található mag- (core) és járulékos gének gyakorisága. Az egyes deléciók (D1-től D12-ig) az 5. táblázatban megadott genomi pozíció alapján azonosíthatók be. Az oszlopok tetején feltüntettük az adott régióban található összes olyan gén darabszámát, amely besorolható core és járulékos kategóriákba. A sötét színű sávok a core gének számának arányát mutatják az adott deletált régióban található összes génhez képest. A szaggatott vízszintes vonal a teljes genomban található core gének számának hányadát mutatja.

V. Az eredmények megvitatása

A RANDEL módszer lehetővé teszi véletlenszerű deléciók sorozatos generálását az *E. coli* genomban idegen szekvenciák visszamaradása nélkül, és a növesztési periódusok révén együtt jár a populáció legfittebb változatainak szelekciójával. Az eljárás idő- és munkaigénye, valamint a genomcsökkentés mértéke (az átlagos delécióméret 24 kbp) hasonló a genomredukciós munkákban általában alkalmazott, célzott deléciós technikákhoz³⁰. Bár törekedtünk az egyes lépések optimalizációjára, úgy gondoljuk, hogy további próbálkozásokkal a ciklusok időtartama jelentősen lerövidíthető (jelenleg kb. egy hét/deléció), és a folyamat egyszerűsíthető. Ugyancsak további optimalizálást igényel a módszer áteresztőképességének (jelenleg mintegy 10 000 sejt/ciklus), a kezelt populációméretnek a növelése a változatos deléciós spektrum és útvonal kialakulási lehetőségének biztosítására.

A módszer sajátossága, hogy – alkalmas metódus híján – nem direkt a kisebb genomra szelektál, hanem mesterségesen bevitt genommodosításokkal (antibiotikum rezisztencia marker random beépítése, genomhasítással stimulált törésjavítás, beépített gén elleni szelekció a deléciós utódok kiválogatására) „kényszeríti ki” az 1 deléció/ciklus folyamatot.

Kísérleteink feltárták a random genomredukció mesterséges előidézésének nehézségeit is. A deléciót eredményező genomtörés-javítás mechanizmusai – egyes specifikus folyamatok, homológ rekombináció és különösen az AEJ – olyan alacsony hatásfokúak, hogy a deléció gyakorisága a populációban a spontán mutációs ráta gyakoriságával van összehasonlítható szinten. Ez azt jelenti, hogy a többlépcsős folyamatban az alkalmazott segédgének és konstrukciók (pl. antibiotikum-rezisztencia-marker, ellenszelekciós gén, plazmidok) a deléciókhoz hasonló gyakorisággal mutációt szenvednek, és hamis pozitív sejteket eredményeznek. Ezek kiküszöbölése a segédkonstrukciók ciklusonkénti eliminációját és újra bevitelét, illetve replika plating alkalmazását teszik szükségessé. Valószínűleg nem véletlen, hogy a random mesterséges genomredukció viszonylag kézenfekvő ötletét eddig nem valósították meg sehol. Anekdotikus hírek és konferencia-poszterek alapján más laboratóriumokban is kísérleteztek ilyen irányban, azonban ezekből publikáció nem született. A RANDEL módszer koncepcióját használták fel Ma és munkatársai⁸⁶, akik az eljárást annyiban

módosították, hogy a genomi kettős szálú DNS törések létrehozására meganukleáz helyett a CRISPR/Cas9 rendszert használták. Az általuk publikált TMDR (transposon mediated random deletion) módszer alkalmazásával 5 random deléciós ciklust hajtottak végre *E. coli*-ban, ahol 5,5%-os genomredukciót értek el úgy, hogy a deléciós sejtek növekedési paraméterei nem változtak jelentős mértékben a szülői törzshöz képest. A deletálódott régiók csak néhány esetben mutattak átfedést a RANDEL módszerrel kapott deléciókkal (prolin bioszintézis gének, flagelláris gének). A deléciók mérete (25-50 kbp/ciklus) a RANDEL-hez hasonló volt, viszont a TMDR robusztusabbnak bizonyult a redukált genomú kiindulási törzsekben és kisebb pontmutációs terheléssel járt (0-20 SNV/ciklus). Hátránya viszont, hogy nem tartalmaz negatív szelekciós lépést a transzpozon elvesztésére, ezért a manuális munkaigénye miatt korlátozott az áteresztőképessége.

Az általunk kidolgozott módszer alkalmazása során hátrányosnak bizonyult, hogy a dP nevű mutagén ellenszelekciós ágens használatának következtében a kezdetben alkalmazott protokoll a deléciók mellett SNV-k kialakulását is előidézte. Eredményeink némileg váratlanok voltak, mivel korábban mindössze tízszeres mutációs gyakoriság emelkedést írtak le dP hatására az általunk használt koncentrációban⁶⁸. Ez a megfigyelés azonban olyan mérésekből származik, ahol a dP-kezelés időtartama folyékony médiumban csak 6 óra volt, mi pedig 24-48 órán keresztül tartó növesztést alkalmaztunk dP tartalmú szilárd táptalajon.

További érdekességet szolgáltatott a pontmutációk eloszlásának mintázata a genomokban: pontmutációk váltakozó sűrűsödését és ritkulását figyeltük meg a kromoszóma mentén; legnagyobb mértékű sűrűsödést a replikáció terminációs régiójában. Forster és munkatársai⁸⁷ hasonló jelenséget tapasztaltak mismatch repair-defektív háttérű mutáció akkumulációs törzsekben: a bázispár szubsztitúciók hullámszerű térbeli mintázatát írták le, amely szimmetrikus a replikációs origóra. Egy 2019-ben megjelent átfogó tanulmányban⁸⁸ megállapították, hogy a jellegzetes mintázatot a DNS replikációt, a hibajavítást és a kromoszóma-szerkezetet befolyásoló tényezők együttes hatása alakítja ki.

A protokoll módosítása a mutációs rátát jelentős mértékben, a háttér szintjére csökkentette. A háttérmutációk létrejöttének megakadályozására további alternatívát

jelenthet a hsvTK-dP rendszer lecserélése egy másik nagy hatékonyságú, ugyanakkor nem mutagén ellenszelekciós módszerre^{89,90}.

Megjegyezzük, a kezdeti protokoll gyakorlati szempontból hátrányos, magas pontmutációs rátája bizonyos tekintetben előnyt is jelenthet. Ha nem csupán a deléciókkal létrehozott genomi áramvonalasítás, illetve a génkészlet-csökkenés hatásainak vizsgálata a cél, hanem egy hosszú távú genomevolúciós pálya vizsgálata, abba a génvesztés mellett a pontmutációk létrejötte is beletartozik. Az irodalmi adatokkal összevetve protokollunk mind a deléciók, mind a pontmutációk akkumulációja tekintetében hasonló, a természetes ütemhez képest mintegy 10^4 - 10^5 -szeres gyorsulást jelenthet, azaz a lombikban tulajdonképpen egy felgyorsított természetes evolúció zajlik. Ennek átfogó tanulmányozása azonban szélesebb körű vizsgálatokat kívánna, olyanokat, amelyek a természetes evolúció többi aspektusát is modellezik (pl. változó körülmények, populációs expanziók és szűk keresztmetszetek).

A felhasználás gyakorlati szempontjait előtérbe helyező, ugyanakkor elméleti kérdésekre is választ kereső munkánk során a szekvenált 60 genomban mindössze 12 különböző deléciót azonosítottunk (közülük néhány többször is előfordul). A kezdeti ciklusokban kapott deléciók korlátozott változatossága három tényezővel magyarázható: (1) specifikus mechanizmusok, amelyek hozzájárulnak egy bizonyos genomi szegmens viszonylag gyakori deléciójához (pl. a profágok saját rendszerük által végrehajtott, stresszindukált kivágódása), (2) szintén a relatív gyakoriságot előidéző, deléciós forró pontokat jelentő kiterjedt homológ szekvenciák jelenléte, és (3) egy adott deléció gyors szelekciója/eliminációja az általa nyújtott kompetitív előny/hátrány következtében. Mind az irodalmi adatok, mind a saját eredményeink alátámasztják az előbb ismertetett mechanizmusokat. Már korábban leírták, hogy a kriptikus *E. coli* profágok kis gyakorisággal bár, de kivágódhatnak a genomból⁹¹. Az esetünkben a dP-kezelés jelentette stressz stimulálhatja a profágok kivágódását. Ezenfelül ismert, hogy különböző genomi lókuszok spontán deléciós gyakoriságaiban 200-szoros különbség is lehet²¹. Szintén leírták a flagelláris gének nagy frekvenciájú spontán delécióját. A flagelláris deléciók egy IS-szekvenciák által közvetített mechanizmus révén jöhetnek létre, és gyorsan fixálódnak a populációban⁹².

A szelekció által hajtott génvesztés jelenségét *E. coli*-ban korábban már leírták, az erről tanúskodó adatok azonban meglehetősen gyérek. Mindössze két olyan genomi

régiót találtak, amelyről kimutatták, hogy a deléciójuk kompetitív előnyt biztosít²¹. Ebben a munkában, gazdag LB táptalaj alkalmazása mellett mi is találtunk fitnessnövelő deléciókat, de szintén csak korlátozott számban. Ezek a flagelláris génklasztert, a Rac profágot (a kapcsolódó szomszédos régiókkal vagy anélkül), és egy 38 kbp-os genomi szegmenst (D12) érintenek. A többi deléció neutrálisnak tűnik; ezek az előnyös deléciók mellett maradtak fent a populációban, legalábbis a tesztelt növekedési körülmények között. Jövőbeli feladat annak kiderítése, hogy a RANDEL módszer segítségével azonosíthatók-e további fitnessnövelő deléciók abban az esetben, ha a növesztési lépéseket különböző környezeti feltételek mellett hajtjuk végre²⁴.

A flagelláris apparátus gyakori elvesztésének oka magától értetődő: lombikban rázatott kultúrákban nincs rá szükség, és a működtetése igen költséges: értékét a sejt energiafelhasználásának 4,5%-ára becsülik²¹. Ennek az energiaigényes struktúrának az elvesztése emeli a növekedési rátát és a biomassza hozam értékét. Az ilyen törzsek lombikban növesztve szignifikáns kompetitív fitnesselőnyvel bírnak, és eredményeink tanúsága szerint ezt az előnyt akkor sem veszítik el, amikor további öt, többnyire neutrális hatású deléciót hordoznak mellettük. A többi fitnessnövelő deléció előnye nem nyilvánvaló. Mivel a DNS elvesztésének energetikai nyeresége elhanyagolható¹⁵, az ok a szükségtelen fehérjék szintézisének abbamaradásában valószínűsíthető²⁶. Módszerünk a növekedés sebességére szelektál, magára a biomassza hozamra nem. A tapasztalt hozamnövekedés minden bizonnyal ennek az energetikai nyereségnek, a csökkent fehérjeszintézis igénynek köszönhető.

Érdekes módon a deletálódott genomi szegmensek többsége az *E. coli* klád járulékos genomjához tartozik. Az itt található gének specifikus körülmények között játszhatnak szerepet (pl. flagelláris gének) vagy horizontális géntranszfer révén válhattak a genom részévé (pl. profágok), és ezáltal hajlamosabbak elveszni laboratóriumi körülmények között. Másfelől néhány olyan deléciót is találtunk, amelyek jelentős hányada a maggenom részének tekinthető. Érdeemes megjegyezni, hogy ezeknek a régióknak a nélkülözhetőségét összehasonlító genomikai megközelítéssel nem lehetne előre jelezni.

A redukált genomú törzsek csökkent fitnessének növelésére megoldásként kínálkozhat az utólagos laboratóriumi evolúció, amelyet sikerrel alkalmaztak egy *E. coli* törzs esetében⁹³. Az MS56 törzs (23%-os genomredukció) növekedése gazdag LB táptalajban

a vad típuséhoz hasonló, M9 minimál médiumban viszont lassan nő. Az adaptív laboratóriumi evolúciós kísérlet során az MS56-ot LB-vel kiegészített M9 táptalajban tenyésztették, miközben annak LB-tartalmát az átoltások során fokozatosan csökkentették. A létrejött mutációk hatására a törzs fitnessze minimál táptalajban jelentősen javult. Multiomikai vizsgálatok tanúsága szerint ennek hátterében a lassabb növekedést okozó anyagcsere-perturbáció transzkriptom és transzlátom szintű visszarendeződése áll. Ez a célzott genomredukciót követő, utólagos, korlátozott korrekció azonban nem alkalmas arra, hogy a szelekció által támogatott, potenciálisan előnyös deléciós útvonalak kialakuljanak. Ezzel szemben a RANDEL módszer kísérletes lehetőséget biztosít többféle genomredukciós evolúciós pálya kihasználására; ezeket ráadásul adaptív és kisebb kompenzáló mutációk kísérhetik. A módszer alkalmazásával kapott multideléciós törzs (6 deléción, a genom 2,5%-át jelentő csökkenés) megnövekedett kompetitív fitnesszt és biomassza hozamot mutatott, ezzel demonstrálva a megközelítésben rejlő lehetőségeket. A fitnessznövelő deléciók száma ugyan limitálnak tűnik, azonban a többségében neutrális deléciók iteratív felhalmozódásával hosszabb távon – a praktikus alkalmazást megkönnyítő további módszer-optimalizációval – jelentős genomegyszerűsödés valósítható meg, még hozzá olyan módon, hogy szelektálódnak a legnagyobb rátermettséggel rendelkező változatok. A legfittebb variánsok szelekciójára gazdag táptalajt használtunk, de specifikus céloknak megfelelő körülmények között történő növesztési periódusok is beilleszthetők az eljárásba, amelyek így evolúciós szűrőként szolgálhatnak. Ennek egy közelmúltban publikált gyakorlati példája, hogy egy biotechnológiai szempontból jelentős vegyület, a polihidroxibutirát (PHB) bioszintézisének optimalizálásra a RANDEL módszer továbbfejlesztésén alapuló véletlenszerű genomredukciós technikát⁸⁶ használták fel. A random deléciós ciklusok és a PHB termelés egyidejű alkalmazásával olyan csökkentett genomú mutánsokat szelektáltak, amelyek a kiindulási törzshöz képest kb. 1,5-szeres biomassza hozammal és 2-szeres PHB termelési képességgel rendelkeztek⁹⁴.

VI. Következtetések

A célzott módszerekkel szemben az itt bemutatott random deléciók eljárás különféle genomredukciós útvonalak kihasználását teszi lehetővé, és együtt jár a populáció legfittebb variánsainak automatikus szelekciójával.

- Megmutattuk, hogy bizonyos genomi szegmensek eliminációja valóban képes megnövelni mind a kompetitív fitneszt, mind a biomassza hozamot, ugyanakkor az igazán előnyös deléciók csak néhány genomi régióra korlátozódnak.
- Eredményeink alapján megállapítható, hogy létezhetnek különféle deléciók útvonalak, azonban ezek jelentőségének felméréséhez, valamint a redukciós módszer gyakorlati alkalmazásához további módszer-optimalizációra van szükség.
- Kimutattuk, hogy a genomredukció során a deléciók mellett kompenzáló, fitnesznövelő pontmutáció is fixálódhat.
- Az eljárás lehetővé teszi a fitnesznövelő és neutrális deléciók előnyös kombinációban történő felhalmozódását a génhálózatok előzetes ismerete nélkül.

Összefoglalás

A genomredukció, vagyis az ősi génkészlet csökkenése egy természetes evolúciós folyamat. Jellemzően endoszimbionta és intracelluláris parazita baktériumok mutatják ezt a jelenséget, de szabadon élő tengeri mikroorganizmusok között is előfordulnak redukált genomú szervezetek. A genomcsökkenéshez neutrális és adaptív folyamatok is hozzájárulhatnak. Kisméretű populációkban a genetikai sodródás lehet a meghatározó tényező, a tengervizekben nagy tömegben jelen lévő mikrobákban viszont szerepe lehet a szelekciónak, amely valószínűleg közvetlenül a genom csökkentésére hat. A jelenség mögött a replikációs költségminimalizálás vagy a DNS mennyiségének csökkenésével járó felszín-térfogat arány növekedése állhat, ami a tápanyagok hatékonyabb transzportjához járulhat hozzá. Mindemellett bizonyos génszakaszok elvesztése bizonyos körülmények között növekedési előnyhöz juttathatja a baktériumot. Továbbá fitnesselőnyt jelenthet a csökkent fehérjeszintézis-igényből eredő energetikai nyereség.

A természetben megfigyelhető folyamatok mintául szolgálnak a mesterséges, laboratóriumi körülmények között történő genomredukciós törekvések számára. A csoportunk úttörő munkáját követően számos genomcsökkentéssel foglalkozó tanulmány született *E. coli*, de más mikroorganizmusok felhasználásával is. A genomcsökkentéssel foglalkozó projekteket alap kutatási és biotechnológiai célok motiválják. Alapfeltevésük, hogy laboratóriumi és ipari körülmények mellett szükségtelenek azok a gének, amelyek csak bizonyos, a természetben előforduló körülmények közötti túlélést szolgálják. Így eltávolításuk csökkent komplexitású, magasabb genetikai stabilitású és gazdaságosabb működésű sejtek létrejöttét eredményezheti. Az említett előnyök közül a gazdaságosabb növekedésre nem találunk meggyőző irodalmi példát, sőt, esetenként a redukciós folyamatot követő fitnesscsökkenés figyelhető meg. Ennek hátterében valószínűleg az áll, hogy az eddigi munkák főként célzott redukciós megközelítést alkalmaztak, amelynek során a génfunkciók és kölcsönhatások korlátozott ismeretére támaszkodva jelölték ki az eltávolítandó genomi szakaszokat. Ezenfelül célzott módszerekkel csak kisszámú deléció-kombináció hozható létre, tehát a lehetséges redukciós útvonalak száma is limitált.

Munkánk alapvető célja a lehető legnagyobb fitnesszel rendelkező redukált genomú *E. coli* törzsek előállítása. Ehhez kapcsolódó fő kérdésünk, hogy léteznek-e olyan deléciók, illetve deléció-kombinációk, amelyek adott körülmények között növelik a fitnesszt és elősegítik a jobb tápanyag-felhasználást. A kérdések megválaszolásához a célzott módszer helyett kifejlesztettünk egy szelekciót is alkalmazó rendszert, amely véletlenszerűen generál deléciókat *E. coli*-ban. Az eljárás sokféle deléció-kombináció létrejöttét eredményezi és az előnyös fitnesszű variánsok feldúsulásával jár.

A módszerünk egy szelekciós marker random inszercióján és stimulált elvesztésén alapul, ami a sejt saját hibajavító mechanizmusainak közreműködésével deléció képződéséhez vezet idegen szekvenciák visszamaradása nélkül. A ciklikus folyamat első lépése egy módosított Tn5 transzpozon random inszerciója, amely egy antibiotikum rezisztencia gént, egy ellenszelekciós markert és az *E. coli* genomban nem előforduló I-SceI nukleáz felismerő helyeket tartalmaz. A második lépés során kettős szálú DNS-töréseket generálunk az I-SceI enzim indukciójával. A túlélő sejtekben a DNS-törés mikro- és makrohomológiák mentén javítódik, ami az inszertált markerek elvesztésével és deléció képződésével jár. A hsvTK-dP ellenszelekciós rendszer biztosítja, hogy a ciklus végén olyan sejteket kapjunk, amelyekben deléció alakult ki. A módszer minden lépése kontrollálható, nagy sejtpopuláción használható, és egy ciklus során sejtenként általában egy új deléciót eredményez. A folyamat során növekedési periódusok biztosítják a legfittebb variánsok szelekcióját.

Az új módszer teszteléséhez 5 egymást követő random deléciós ciklust hajtottunk végre vad típusú *E. coli* MG1655 sejteken, gazdag táptalajban. A kapott deléciók átlagos hossza 24 kbp volt. A leginkább redukált törzs 5 ciklus után 6 deléciót tartalmazott, ami 2,5%-os genomcsökkenést jelent. A kapott genomok szekvenálása során fény derült arra, hogy a deléciók csupán néhány genomi régióra korlátozódnak. Mindössze 12 különböző deléciót figyeltünk meg 60 szekvenált genomban. Bizonyos esetekben eltérő számú deléciót kaptunk, mint amennyit a deléciós ciklusok alapján vártunk. Néhány deléció több törzsben is szerepelt, különösen felülreprezentáltak voltak a flagelláris apparátus elvesztésével járó deléciók. Ennek nyilvánvaló oka, hogy lombikban való tenyésztés során nincs rá szükség, ráadásul a működtetése különösen energiaigényes, így a flagelláris deléciók könnyen szelektálódnak a populációban. Szintén gyakori volt a horizontális géntranszferrel szerzett kriptikus profágok deléciója.

Létrejöttük mechanizmusát tekintve a deléciók változatos méretű genomi ismétlődések mentén alakultak ki: gén-pszudogén párok, inszerciós szekvenciák homológ régiói, ún. REP elemek vagy mikrohomológiák között.

A delécióknak a kísérleteinkben tapasztalt korlátozott változatossága többféleképpen magyarázható. Először is, lehetnek specifikus mechanizmusok, amelyek elősegítik egy bizonyos régió delécióját (profágok stresszindukált kivágódása, flagellumok spontán deléciója). Másodrészt, a hosszú DNS repeat-ek deléciós forró pontként szolgálhatnak; harmadrészt, bizonyos deléciók szelekciós előnnyel vagy hátránnyal járnak, így gyorsan feldúsulhatnak vagy elvesznek a populációból. Az is valószínűsíthető, hogy valójában kevés olyan genomi régió van a hosszú evolúcióval csiszolódott természetes genomban, amely teljesen felesleges, netán elvesztése előnnyel is jár az adott tenyésztési körülmények között.

A módszer különösen hatékony ellenszelekciós rendszer alkalmazását igényli. Az általunk használt hsvTK-dP rendszer megfelelt a célnak, de a vártnál nagyobb mutagenitásúnak bizonyult, így mellékhatásként nagyszámú pontmutációt figyeltünk meg a szekvenált törzsekben. Később, a protokoll módosításával a mutációs rátát sikerült a háttér szintjére csökkentenünk. A pontmutációk eloszlásának elemzése megmutatta, hogy bizonyos előnyös deléciós változatok szelektálódhatnak, ugyanakkor egymástól függetlenül is kialakulhatnak a különböző vonalakban. Egyes pontmutációk előnyt biztosító hatására is találtunk példákat (pl. fitneszcsökkenést kompenzáló pontmutáció a *yehB* géneben).

A deléciók és pontmutációk fenotipikus hatásának elkülönítéséhez vad típusú, háttérmutációktól mentes háttéren rekonstruáltuk az egyedi deléciókat, illetve azok kombinációit. A deléciók egy része nem befolyásolta a növekedési paramétereket. Bizonyos deléciók – amelyek lényegében három különböző típust képviselnek (flagelláris apparátus, Rac profág, D12 deléció) – viszont a növekedési ráta emelkedését okozták. A legkifejezettebb hatása a flagelláris régió kiesésének volt. Abban az esetben is magasabb növekedési rátát és optikai denzitás növekményt (a biomassa hozam mérőszáma) mértünk, amikor a flagelláris deléciót 5 másik delécióval kombináltuk. A 6-deléciós törzs előnyét kompetíciós kísérletek és biomassa hozam mérések is megerősítették.

Megállapítottuk, hogy a legtöbb deletálódott szegmens az *E. coli* klád járulékos genomjához tartozik. Azonban meglepő módon találtunk két olyan deletált régiót,

amelyek jelentős hányada a maggenom részének tekinthető. Ennek az a jelentősége, hogy olyan régiók is a módszer célpontjai lehetnek, amelyek nélkülözhetőségét összehasonlító genomikai megközelítéssel nem lehetne előre jelezni.

Eredményeinknek két fontos aspektusa van. (1) Kifejlesztettünk egy szelekción alapuló, iteratív genomredukációs módszert, melynek segítségével véletlenszerű deléciók generálhatók nagy *E. coli* sejtpopulációkban, és ahol az előnyös deléció-kombinációk automatikusan szelektálódnak. A módszer lehetővé teszi számos előnyös genomredukációs pálya felderítését a génfunkciók és kölcsönhatások előzetes ismerete nélkül. (2) Az eljárás *E. coli*-ban történő alkalmazásával és a redukált genomok elemzésével megmutattuk, hogy bizonyos genomi szegmensek elvesztése növeli a fitnesszt és a biomassza hozamot, ugyanakkor a valóban előnyös deléciók köre csupán néhány genomi régióra korlátozódik.

A távlatokat tekintve: a módszer gyakorlati alkalmazását megkönnyítő további optimalizáció esetén lehetőség nyílik a neutrális és fitnessznövelő deléciók előnyös kombinációban történő felhalmozására, jelentős, fitnessz-maximalizáló genomredukációra.

Summary

Genome reduction, the shrinkage of the ancestral gene set, is a natural evolutionary process typically seen in endosymbiotic and intracellular parasitic bacteria, but also occurs in free-living marine microorganisms. Both neutral and adaptive processes are thought to contribute to genome reduction. In small populations, genetic drift may be the primary driving force, while in microbes abundant in seawater, selection may play a role, likely having a direct effect on the reduction of genome size. The reason behind it might be the minimization of replication costs, or attaining better nutrient transport by the increased surface-to-volume ratio of shrunken cells having less amount of DNA. Furthermore, loss of certain gene segments can provide growth advantage for bacteria under certain conditions. Additionally, fitness advantage may stem from energetic gain derived from reduced protein synthesis demand.

The natural evolutionary processes inspire artificial genome reduction efforts in laboratory settings. Following the pioneering work of our group, many genome reduction studies were published, applying mostly *E. coli*, but also in other microorganisms. These projects are motivated by fundamental research aims, and also by biotechnological interests. Their basic assumption is that many genes are unnecessary under laboratory or industrial conditions, because they are needed only for the survival in their natural environment. Consequently, the removal of this genetic material would result in cells with reduced complexity, higher genetic stability and improved cellular economy. Convincing examples of more economical growth are not found in the literature. Moreover, in some cases, a gradual decline of fitness accompanying the genome reduction is seen. The possible reason behind is that these studies mainly use targeted reduction approach, where the basis of individual deletion target selection, gene functions and interactions, is often not well known. In addition, using targeted methods, only a restricted number of deletion combinations can be generated, therefore the number of possible reduction pathways is limited.

The main goal of our work was to generate reduced genome *E. coli* strains with the highest possible fitness. The main related question is whether there are deletions, or combinations of deletions that increase fitness, and promote better utilization of nutrients under certain conditions. To answer the questions, instead of a targeted

method, we developed a selection-driven system that randomly generates deletions in *E. coli*. The process results in many deletion combinations, and in the enrichment of variants with a fitness advantage.

Our method is based on random insertion and stimulated loss of a selectable marker, which leads to formation of a deletion due to intrinsic DNA repair mechanisms without leaving exogenous sequences in the genome. The first step of the cyclic process is the random insertion of a modified Tn5 transposon carrying an antibiotic resistance gene, a counter-selectable marker, and unique (not present in the *E. coli* genome) I-SceI nuclease recognition sites. The second step is the induction the I-SceI nuclease enzyme to create double-stranded DNA breaks. In the survivor cells, the DNA break is repaired at micro- or macrohomologies, resulting in the loss of the inserted markers, and the formation of a deletion. The hsvTK-dP counter-selection system ensures that cells will carry a deletion at the end of the cycle. In this process, every step is controllable, can be used on large population of cells, and generally results in one deletion per cycle per cell. The growth periods inherent in the procedure allow the selection of the fittest variants.

To test the new method, we performed five consecutive random deletion cycles on wild-type *E. coli* MG1655 cells in rich medium. The average deletion size was around 24 kbp. The most reduced strain obtained in five deletion cycles carried 6 deletions, having a genome reduced by 2.5%. Sequencing of multiple genomes revealed that deletions were limited to a few genomic regions. We observed only 12 different deletions among 60 sequenced genomes. In some cases, we observed different number of deletions than expected from the number of cycles. Some of the deletions were present in multiple lines; e. g. deletions affecting the flagellar apparatus were particularly overrepresented. The reason behind is obvious: they are not needed in the shake flask, and the estimated cost of running flagella is remarkably high, hence flagellar deletions are easily selected in the population. Deletions of the horizontally acquired cryptic prophages were also frequent. In terms of mechanism, deletions were formed along genomic repeats of various sizes: between gene-pseudogene pairs, homologous regions of insertion sequences, between REP elements or microhomologies.

The limited variability of the observed deletions can be explained in multiple ways. First, there are specific mechanisms that promote the deletion of a particular segment (stress-induced excision of prophages, spontaneous deletion of flagella).

Second, long DNA repeats can serve as deletion hotspots; third, certain deletions have a selection advantage or disadvantage, so they are rapidly selected or eliminated from the population. As a matter of fact, it is also likely that the naturally evolved genome comprises only a few genomic regions that are completely unnecessary, or their loss would confer benefits in certain culturing conditions.

The method requires the use of an especially efficient counter-selection system. The applied hsvTK-dP system was fit for the purpose, but proved to be more mutagenic than expected, therefore a large number of point mutations were observed in the sequenced strains as a side effect. Later, by modifying the protocol, we were able to reduce the mutation rate to the background level. Analysis of the distribution of point mutations showed that certain preferred deletion variants may be selected, but could also emerge independently in different lines. We also found examples of the advantageous effect of some point mutations (e. g., a point mutation in *yehB* gene compensating for fitness loss).

To delimit the phenotypic effect of deletions and point mutations, we reconstructed individual deletions and their combinations in the clean background of the non-evolved, mutation-free parental strain. While some of the deletions did not change the growth parameters, six deletions, essentially representing three types (flagellar apparatus, Rac, D12 deletion) caused a significant elevation of the growth rate, compared to wild-type. The most pronounced effect was observed when the flagellar region was deleted. When the flagellar deletion was combined with five additional deletions, the multideletion strain still showed elevated growth rate and optical density increment (a proxy for biomass yield). The advantage of the 6-deletion strain over wild-type was also confirmed by competition experiments and biomass yield measurements.

We found that most of the deleted segments belong to the accessory genome of the *E. coli* clade. Surprisingly, however, we found two deleted regions, in which a significant fraction of genes are part of the core genome. This finding suggests that our method can also target regions that would not be predicted by comparative genomic approaches.

Our results have two important aspects. (1) We developed a selection-driven, iterative genome reduction method that generates successive random deletions in large populations of *E. coli* cells, where the beneficial deletion-combinations are automatically

selected. The method allows the exploration of large number of advantageous genome streamlining trajectories without prior knowledge of gene functions and interactions. (2) By applying the method to *E. coli* and by analysing the reduced genomes, we showed that elimination of certain genomic segments can improve competitive fitness and biomass yield, but the choice of the truly beneficial deletions are limited to a few genomic regions.

Concerning the future perspectives, further optimisation facilitating the practical application of the method will enable the accumulation of neutral and fitness increasing deletions in favourable combinations, and will allow substantial, fitness-maximizing genome reduction.

Köszönetnyilvánítás

Köszönöm témavezetőm, Pósfai Gyuri kitartó bizalmát, ötleteit és javaslatait, az írásban nyújtott segítségét. Külön köszönöm a részvételi lehetőséget a konferenciákon és egy nyári iskolában.

Hálás vagyok a nem hivatalos témavezetőmnek, Györfy Zsuzsának. Köszönöm, hogy a kezdetektől fogva türelemmel mentorált, és hogy hihetetlenül jó érzékkel adagolt empatikus konfrontációval segített át a nehézségeken.

Köszönöm Fehér Tamásnak a kifogyhatatlan érdeklődését, építő hozzászólásait. Köszönettel tartozom Nyerges Ákosnak, hogy projekt kezdeti szakaszában részt vett az ötletelésekben.

Nagyon köszönöm Karcagi Ildikónak a törzsrekonstrukciókat, Tímár Editnek a genomi pontmutáció létrehozását, Nagy Istvánnak a genomszekvenálásokat, Papp Baláznak és Györkei Ádámnak pedig a maggenom-elemzésben nyújtott segítségét. Baláznak külön köszönöm, hogy észrevételeivel növelte az elsőszerezős cikkem színvonalát.

Hálás vagyok a házi védésem bírálóinak, Dr. Papp Tamásnak és Dr. Bálint Baláznak, hogy a dolgozatom aprólékos és értő átnézésével emelték az értekezés minőségét.

Köszönöm a 448-as és 450-es laborban megfordult munkatársak társaságát. Köszönöm, hogy mindig számíthattam rájuk, és hogy teret adtak a szóvicceimnek.

Mindenekfelett köszönöm a családomnak és a barátaimnak a támaszt és a biztatást.

Köszönöm az OTKA K116455 és GINOP-2.3.2-15-2016-00001 pályázatoknak, hogy biztosították a munkám anyagi feltételeit.

Saját közlemények jegyzéke

MTMT azonosító: 10048443

Összesített impakt faktor (I.F.): 18,77

Tudományos folyóiratcikkek

Vernyik V, Karcagi I, Tímár E, Nagy I, Györkei Á, Papp B, Gyorfy Z, Posfai G. Exploring the fitness benefits of genome reduction in *Escherichia coli* by a selection-driven approach. *Scientific Reports*. 2020 Apr 30;10(1):7345. (I.F.: 3,998) *

Lipinszki Z, Vernyik V, Farago N, Sari T, Puskas LG, Blattner FR, Posfai G, Gyorfy Z. Enhancing the Translational Capacity of *E. coli* by Resolving the Codon Bias. *ACS Synth Biol*. 2018 Nov 16;7(11):2656–64. (I.F.: 5,571)

Gyorfy Z, Draskovits G, Vernyik V, Blattner FF, Gaal T, Posfai G. Engineered ribosomal RNA operon copy-number variants of *E. coli* reveal the evolutionary trade-offs shaping rRNA operon number. *Nucleic Acids Res*. 2015 Feb 18;43(3):1783–94. (I.F.: 9,202)

*a disszertáció alapját képező publikáció

Könyvfejezet

Balikó G, Vernyik V, Karcagi I, Györffy Z, Draskovits G, Fehér T, Posfai G. Rational Efforts to Streamline the *Escherichia coli* Genome. In: *Synthetic Biology*. John Wiley & Sons, Ltd; 2018. p. 49–80.

Irodalomjegyzék

1. Wolf, Y. I. & Koonin, E. V. Genome reduction as the dominant mode of evolution. *Bioessays* **35**, 829–837 (2013).
2. Moran, N. A. Microbial Minimalism: Genome Reduction in Bacterial Pathogens. *Cell* **108**, 583–586 (2002).
3. McCutcheon, J. P. & Moran, N. A. Extreme genome reduction in symbiotic bacteria. *Nat Rev Microbiol* **10**, 13–26 (2012).
4. Giovannoni, S. J. *et al.* Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**, 1242–1245 (2005).
5. Klasson, L. & Andersson, S. G. E. Evolution of minimal-gene-sets in host-dependent bacteria. *Trends in Microbiology* **12**, 37–43 (2004).
6. Mira, A., Ochman, H. & Moran, N. A. Deletional bias and the evolution of bacterial genomes. *Trends in Genetics* **17**, 589–596 (2001).
7. Lukjancenko, O., Wassenaar, T. M. & Ussery, D. W. Comparison of 61 Sequenced *Escherichia coli* Genomes. *Microb Ecol* **60**, 708–720 (2010).
8. Bennett, G. M. & Moran, N. A. Small, Smaller, Smallest: The Origins and Evolution of Ancient Dual Symbioses in a Phloem-Feeding Insect. *Genome Biol Evol* **5**, 1675–1688 (2013).
9. Sällström, B. & Andersson, S. G. Genome reduction in the α -Proteobacteria. *Current Opinion in Microbiology* **8**, 579–585 (2005).
10. Fraser, C. M. *et al.* The minimal gene complement of *Mycoplasma genitalium*. *Science* **270**, 397–403 (1995).
11. Dufresne, A., Garczarek, L. & Partensky, F. Accelerated evolution associated with genome reduction in a free-living prokaryote. *Genome Biol* **6**, R14 (2005).
12. Giovannoni, S. J. *et al.* The small genome of an abundant coastal ocean methylotroph. *Environ Microbiol* **10**, 1771–1782 (2008).
13. Kuo, C.-H., Moran, N. A. & Ochman, H. The consequences of genetic drift for bacterial genome complexity. *Genome Res* **19**, 1450–1454 (2009).
14. Giovannoni, S. J., Cameron Thrash, J. & Temperton, B. Implications of streamlining theory for microbial ecology. *ISME J* **8**, 1553–1565 (2014).

15. Kim, B. H. & Gadd, G. M. Bacterial Physiology and Metabolism by Byung Hong Kim. *Cambridge Core* /core/books/bacterial-physiology-and-metabolism/C18A2AA9FE946BC7EAE72D93510FA6D3 (2008) doi:10.1017/CBO9780511790461.
16. Button, D. K. Nutrient uptake by microorganisms according to kinetic parameters from theory as related to cytoarchitecture. *Microbiol Mol Biol Rev* **62**, 636–645 (1998).
17. Maurelli, A. T., Fernández, R. E., Bloch, C. A., Rode, C. K. & Fasano, A. ‘Black holes’ and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Proc Natl Acad Sci U S A* **95**, 3943–3948 (1998).
18. Cooper, V. S., Schneider, D., Blot, M. & Lenski, R. E. Mechanisms causing rapid and parallel losses of ribose catabolism in evolving populations of *Escherichia coli* B. *J Bacteriol* **183**, 2834–2841 (2001).
19. Notley-McRobb, L., King, T. & Ferenci, T. rpoS mutations and loss of general stress resistance in *Escherichia coli* populations as a consequence of conflict between competing stress responses. *J Bacteriol* **184**, 806–811 (2002).
20. Girgis, H. S., Hottes, A. K. & Tavazoie, S. Genetic architecture of intrinsic antibiotic susceptibility. *PLoS One* **4**, e5629 (2009).
21. Koskiniemi, S., Sun, S., Berg, O. G. & Andersson, D. I. Selection-Driven Gene Loss in Bacteria. *PLoS Genet* **8**, (2012).
22. Lee, M.-C. & Marx, C. J. Repeated, selection-driven genome reduction of accessory genes in experimental populations. *PLoS Genet* **8**, e1002651 (2012).
23. Morris, J. J., Lenski, R. E. & Zinser, E. R. The Black Queen Hypothesis: Evolution of Dependencies through Adaptive Gene Loss. *mBio* **3**, (2012).
24. D’Souza, G. *et al.* Less Is More: Selective Advantages Can Explain the Prevalent Loss of Biosynthetic Genes in Bacteria. *Evolution* **68**, 2559–2570 (2014).
25. Bourguignon, T. *et al.* Increased Mutation Rate Is Linked to Genome Reduction in Prokaryotes. *Current Biology* **30**, 3848–3855.e4 (2020).
26. O’Brien, E. J., Utrilla, J. & Palsson, B. O. Quantification and Classification of *E. coli* Proteome Utilization and Unused Protein Costs across Environments. *PLoS Computational Biology* **12**, e1004998 (2016).

27. Nilsson, A. I. *et al.* Bacterial genome size reduction by experimental evolution. *Proc Natl Acad Sci U S A* **102**, 12112–12116 (2005).
28. Barrick, J. E. *et al.* Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243–1247 (2009).
29. Fehér, T., Papp, B., Pal, C. & Pósfai, G. Systematic genome reductions: theoretical and experimental approaches. *Chem Rev* **107**, 3498–3513 (2007).
30. Martínez-García, E. & de Lorenzo, V. The quest for the minimal bacterial genome. *Current Opinion in Biotechnology* **42**, 216–224 (2016).
31. LeBlanc, N. & Charles, T. C. Bacterial genome reductions: Tools, applications, and challenges. *Frontiers in Genome Editing* **4**, (2022).
32. Yu, B. J. *et al.* Minimization of the *Escherichia coli* genome using a Tn 5 -targeted Cre/ loxP excision system. *Nature Biotechnology* **20**, 1018–1023 (2002).
33. Goryshin, I. Y., Naumann, T. A., Apodaca, J. & Reznikoff, W. S. Chromosomal Deletion Formation System Based on Tn5 Double Transposition: Use For Making Minimal Genomes and Essential Gene Analysis. *Genome Res* **13**, 644–653 (2003).
34. Hashimoto, M. *et al.* Cell size and nucleoid organization of engineered *Escherichia coli* cells with a reduced genome. *Molecular Microbiology* **55**, 137–149 (2005).
35. Gyorfy, Z. *et al.* Engineered ribosomal RNA operon copy-number variants of *E. coli* reveal the evolutionary trade-offs shaping rRNA operon number. *Nucleic Acids Res* **43**, 1783–1794 (2015).
36. Mizoguchi, H., Sawano, Y., Kato, J. & Mori, H. Superpositioning of Deletions Promotes Growth of *Escherichia coli* with a Reduced Genome. *DNA Res* **15**, 277–284 (2008).
37. Hirokawa, Y. *et al.* Genetic manipulations restored the growth fitness of reduced-genome *Escherichia coli*. *Journal of Bioscience and Bioengineering* **116**, 52–58 (2013).
38. Morimoto, T. *et al.* Enhanced Recombinant Protein Productivity by Genome Reduction in *Bacillus subtilis*. *DNA Res* **15**, 73–81 (2008).
39. Reuß, D. R. *et al.* Large-scale reduction of the *Bacillus subtilis* genome: consequences for the transcriptional network, resource allocation, and metabolism. *Genome Res* **27**, 289–299 (2017).

40. Aguilar Suárez, R., Stülke, J. & van Dijl, J. M. Less Is More: Toward a Genome-Reduced *Bacillus* Cell Factory for “Difficult Proteins”. *ACS Synth. Biol.* **8**, 99–108 (2019).
41. Baumgart, M. *et al.* *Corynebacterium glutamicum* Chassis C1*: Building and Testing a Novel Platform Host for Synthetic Biology and Industrial Biotechnology. *ACS Synth. Biol.* **7**, 132–144 (2018).
42. Komatsu, M., Uchiyama, T., Ōmura, S., Cane, D. E. & Ikeda, H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc Natl Acad Sci U S A* **107**, 2646–2651 (2010).
43. Bu, Q.-T. *et al.* Rational construction of genome-reduced and high-efficient industrial *Streptomyces* chassis based on multiple comparative genomic approaches. *Microb Cell Fact* **18**, 16 (2019).
44. Leprince, A., de Lorenzo, V., Völler, P., van Passel, M. W. J. & Martins dos Santos, V. A. P. Random and cyclical deletion of large DNA segments in the genome of *Pseudomonas putida*. *Environ Microbiol* **14**, 1444–1453 (2012).
45. Martínez-García, E., Nikel, P. I., Aparicio, T. & de Lorenzo, V. *Pseudomonas* 2.0: genetic upgrading of *P. putida* KT2440 as an enhanced host for heterologous gene expression. *Microbial Cell Factories* **13**, 159 (2014).
46. Liang, P. *et al.* Deletion of genomic islands in the *Pseudomonas putida* KT2440 genome can create an optimal chassis for synthetic biology applications. *Microb Cell Fact* **19**, 70 (2020).
47. Fan, X. *et al.* Genome reduction enhances production of polyhydroxyalkanoate and alginate oligosaccharide in *Pseudomonas mendocina*. *Int J Biol Macromol* **163**, 2023–2031 (2020).
48. Park, M. K. *et al.* Enhancing recombinant protein production with an *Escherichia coli* host strain lacking insertion sequences. *Appl Microbiol Biotechnol* **98**, 6701–6713 (2014).
49. Pósfai, G. *et al.* Emergent properties of reduced-genome *Escherichia coli*. *Science* **312**, 1044–1046 (2006).
50. Karcagi, I. *et al.* Indispensability of Horizontally Transferred Genes and Its Impact on Bacterial Genome Streamlining. *Mol Biol Evol* **33**, 1257–1269 (2016).

51. Umenhoffer, K. *et al.* Reduced evolvability of Escherichia coli MDS42, an IS-less cellular chassis for molecular and synthetic biology applications. *Microbial Cell Factories* **9**, 38 (2010).
52. Csörgő, B., Fehér, T., Tímár, E., Blattner, F. R. & Pósfai, G. Low-mutation-rate, reduced-genome Escherichia coli: an improved host for faithful maintenance of engineered genetic constructs. *Microb Cell Fact* **11**, 11 (2012).
53. Rugbjerg, P. & Sommer, M. O. A. Overcoming genetic heterogeneity in industrial fermentations. *Nat Biotechnol* **37**, 869–876 (2019).
54. Chakiath, C. S. & Esposito, D. Improved recombinational stability of lentiviral expression vectors using reduced-genome Escherichia coli. *BioTechniques* **43**, 466–470 (2007).
55. Fehér, T., Karcagi, I., Blattner, F. R. & Pósfai, G. Bacteriophage recombineering in the lytic state using the lambda red recombinases. *Microb Biotechnol* **5**, 466–476 (2012).
56. Gohrbandt, S. *et al.* H9 avian influenza reassortant with engineered polybasic cleavage site displays a highly pathogenic phenotype in chicken. *Journal of General Virology*, **92**, 1843–1853 (2011).
57. Schoggins, J. W. *et al.* Dengue reporter viruses reveal viral dynamics in interferon receptor-deficient mice and sensitivity to interferon effectors in vitro. *Proc Natl Acad Sci U S A* **109**, 14610–14615 (2012).
58. Pöhlmann, C. *et al.* Periplasmic delivery of biologically active human interleukin-10 in Escherichia coli via a sec-dependent signal peptide. *J Mol Microbiol Biotechnol* **22**, 1–9 (2012).
59. Lieder, S., Nickel, P. I., de Lorenzo, V. & Takors, R. Genome reduction boosts heterologous gene expression in Pseudomonas putida. *Microbial Cell Factories* **14**, 23 (2015).
60. Ran, H., Wu, J., Wu, D. & Duan, X. Enhanced Production of Recombinant Thermobifida fusca Isoamylase in Escherichia coli MDS42. *Appl Biochem Biotechnol* **180**, 464–476 (2016).
61. Lee, J. H. *et al.* Metabolic engineering of a reduced-genome strain of Escherichia coli for L-threonine production. *Microbial Cell Factories* **8**, 2 (2009).

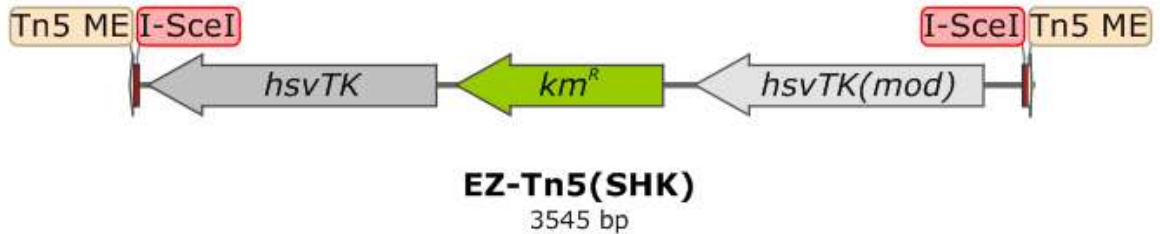
62. Fredens, J. *et al.* Total synthesis of *Escherichia coli* with a recoded genome. *Nature* **569**, 514–518 (2019).
63. Gibson, D. G. *et al.* Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. *Science* **329**, 52–56 (2010).
64. Hutchison, C. A. *et al.* Design and synthesis of a minimal bacterial genome. *Science* **351**, (2016).
65. Kurokawa, M., Seno, S., Matsuda, H. & Ying, B.-W. Correlation between genome reduction and bacterial growth. *DNA Res* **23**, 517–525 (2016).
66. Blattner, F. R. *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1462 (1997).
67. Platt, R., Drescher, C., Park, S.-K. & Phillips, G. J. Genetic System for Reversible Integration of DNA Constructs and *lacZ* Gene Fusions into the *Escherichia coli* Chromosome. *Plasmid* **43**, 12–23 (2000).
68. Tominaga, M. *et al.* Liquid-Based Iterative Recombineering Method Tolerant to Counter-Selection Escapes. *PLoS One* **10**, (2015).
69. Pósfai, G., Koob, M. D., Kirkpatrick, H. A. & Blattner, F. R. Versatile insertion plasmids for targeted genome manipulations in bacteria: isolation, deletion, and rescue of the pathogenicity island LEE of the *Escherichia coli* O157:H7 genome. *J. Bacteriol.* **179**, 4426–4428 (1997).
70. Kolisnychenko, V. *et al.* Engineering a Reduced *Escherichia coli* Genome. *Genome Res.* **12**, 640–647 (2002).
71. Hall, B. G. Activation of the *bgl* operon by adaptive mutation. *Molecular Biology and Evolution* **15**, 1–5 (1998).
72. Goryshin, I. Y., Jendrisak, J., Hoffman, L. M., Meis, R. & Reznikoff, W. S. Insertional transposon mutagenesis by electroporation of released Tn5 transposition complexes. *Nat Biotechnol* **18**, 97–100 (2000).
73. Pósfai, G., Kolisnychenko, V., Berczki, Z. & Blattner, F. R. Markerless gene replacement in *Escherichia coli* stimulated by a double-strand break in the chromosome. *Nucleic Acids Res* **27**, 4409–4415 (1999).
74. Szappanos, B. *et al.* Adaptive evolution of complex innovations through stepwise metabolic niche expansion. *Nat Commun* **7**, (2016).

75. Warringer, J., Ericson, E., Fernandez, L., Nerman, O. & Blomberg, A. High-resolution yeast phenomics resolves different physiological features in the saline response. *Proc Natl Acad Sci U S A* **100**, 15724–15729 (2003).
76. Nyerges, Á. *et al.* Conditional DNA repair mutants enable highly precise genome engineering. *Nucleic Acids Res.* **42**, e62 (2014).
77. Lenski, R. E., Rose, M. R., Simpson, S. C. & Tadler, S. C. Long-Term Experimental Evolution in *Escherichia coli*. I. Adaptation and Divergence During 2,000 Generations. *The American Naturalist* **138**, 1315–1341 (1991).
78. Maddamsetti, R. *et al.* Core Genes Evolve Rapidly in the Long-Term Evolution Experiment with *Escherichia coli*. *Genome Biol Evol* **9**, 1072–1083 (2017).
79. Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**, 2006.0008 (2006).
80. Goryshin, I. Y., Miller, J. A., Kil, Y. V., Lanzov, V. A. & Reznikoff, W. S. Tn5/IS50 target recognition. *Proc Natl Acad Sci U S A* **95**, 10716–10721 (1998).
81. Pitcher, R. S., Brissett, N. C. & Doherty, A. J. Nonhomologous end-joining in bacteria: a microbial perspective. *Annu Rev Microbiol* **61**, 259–282 (2007).
82. Lovett, S. T., Hurley, R. L., Suttera, V. A., Aubuchon, R. H. & Lebedeva, M. A. Crossing Over Between Regions of Limited Homology in *Escherichia coli*: RecA-Dependent and RecA-Independent Pathways. 10.
83. Chayot, R., Montagne, B., Mazel, D. & Ricchetti, M. An end-joining repair mechanism in *Escherichia coli*. *Proc Natl Acad Sci U S A* **107**, 2141–2146 (2010).
84. Tashiro, Y., Fukutomi, H., Terakubo, K., Saito, K. & Umeno, D. A nucleoside kinase as a dual selector for genetic switches and circuits. *Nucleic Acids Res* **39**, e12 (2011).
85. Schaller, H. The Intergenic Region and the Origins for Filamentous Phage DNA Replication. *Cold Spring Harbor Symposia on Quantitative Biology* **43**, 401–408 (1979).
86. Ma, S., Su, T., Liu, J., Lu, X. & Qi, Q. Reduction of the Bacterial Genome by Transposon-Mediated Random Deletion. *ACS Synth. Biol.* **11**, 668–677 (2022).
87. Foster, P. L., Hanson, A. J., Lee, H., Popodi, E. M. & Tang, H. On the Mutational Topology of the Bacterial Genome. *G3 (Bethesda)* **3**, 399–407 (2013).

88. Niccum, B. A., Lee, H., Mohammedismail, W., Tang, H. & Foster, P. L. The Symmetrical Wave Pattern of Base-Pair Substitution Rates across the Escherichia coli Chromosome Has Multiple Causes. **10**, 18 (2019).
89. Chen, Z. & Zhao, H. A highly sensitive selection method for directed evolution of homing endonucleases. *Nucleic Acids Res* **33**, e154 (2005).
90. Li, X., Thomason, L. C., Sawitzke, J. A., Costantino, N. & Court, D. L. Positive and negative selection using the tetA-sacB cassette: recombineering and P1 transduction in Escherichia coli. *Nucleic Acids Res* **41**, e204 (2013).
91. Wang, X. *et al.* Cryptic prophages help bacteria cope with adverse environments. *Nat Commun* **1**, 1–9 (2010).
92. Plague, G. R. *et al.* Transposable Elements Mediate Adaptive Debilitation of Flagella in Experimental Escherichia coli Populations. *J Mol Evol* **84**, 279–284 (2017).
93. Choe, D. *et al.* Adaptive laboratory evolution of a genome-reduced Escherichia coli. *Nat Commun* **10**, 1–14 (2019).
94. Ma, S. *et al.* Random genome reduction coupled with polyhydroxybutyrate biosynthesis to facilitate its accumulation in Escherichia coli. *Frontiers in Bioengineering and Biotechnology* **10**, (2022).

Függelék

a



b

CTGTCTCTTATACACATCTTAGGGATAACAGGGTAATCGATGCTGTTCGGAAACACGGTGCCTGACTGCAA
 GCTTTCAGTTAGCCTCCCCATCTCCCGGGCAAACGTGCGCGCCAGGTTCGCATATCGTTCGGTATGGAGCC
 GGGGTGGTACGTGGGTCTGGACCATCCCGGAGGTAAGTTGCAGCAGGGCGTCCCGGCAGCCGGCGGGC
 GATTGGTCGTAATCCAGGATAAAGACGTGCATGGAACGGAGGCGTTTGGCCAAGACGTCCAAGGCCAGG
 CAAACACGTTATACAGGTCGCCGTTGGGGGCCAGCAACTCGGGGGCCGAAACAGGGTAAATAACGTGTC
 CCCGATATGGGGTTCGTGGGCCCGCGTTGCTCTGGGGCTCGGCACCCTGGGGCGGCACGGCCGTCCCCGAA
 AGCTGTCCCCAGTCTCCCGCCACGACCCGCCGCACTGCAGATACCGCACCGTATTGGCAAGTAGCCCGT
 AAACGCGGCGAATCGCAGCCAGCATAGCCAGGTCCAGCCGCTCGCCGGGGCGCTGGCGTTTGGCCAGGCG
 GTCGATGTGTCTGTCTCCGGAAGGGCCCCAAGCACGATGTTGGTGCCGGGCAAGGTCGGCGGGATGAGG
 GCCACGAACGCCAGCACGGCCTGGGGGGTCATGCTGCCATAAAGTACCGCGCGGCCGGGTAGCACAGGA
 GGGCGGCGATGGGATGGCGGTGCAAGATGAGGGTGAGGGCCGGGGCGGGGCATGTGAGCTCCCAGCCTC
 CCCCCGATATGAGGAGCCAGAACGGCGTCGGTCACGGCATAAAGCATGCCATTGTTATCTGGGCGCTT
 GTCATTACCACCGCCGCGTCCCCGGCCGATATCTCACCTGGTTCGAGGCGGTGTTGTGTGGTGTAGATGT
 TCGCGATTGTCTCGGAAGCCCCCAGCACCCGCCAGTAAGTCATCGGCTCGGGTACGTAGACGATATCGTTC
 GCGCGAACCCAGGGCCACCAGCAGTTGCGTGGTGGTGGTTTTCCCATCCCGTGGGGACCGTCTATATAA
 ACCCGCAGTAGCGTGGGCATTTTCTGCTCCGGGCGGACTTCCGTGGCTTCTTGCTGCCGGCGAGGGCGCA
 ACGCCGTACGTCGGTTGCTATGGCCGCGAGAACGCGCAGCCTGGTTCGAACGCAGACGCGTGTGATGGCC
 GGGGTACGAAGCCATGGCTAGTATTTCTCCTCTTTCTCTAGTAGTGCTCAGTATCTCTATCACTGATAGG
 GATGTCAATCTCTAAGCTAATTAATTAGAAAAATTCATCCAGCATCAGATGAAATTCAGTGTGTTTCATAT
 CCGGATTATCAATGCCATATTTCTGAAACAGACGTTTTTGCAGGCTCGGGCTAAATTCGCCCAGGCAGTT
 CCACAGAATGGCCAGATCCTGATAACGATCCGCAATGCCACACGGCCACATCAATGCAGCCAATCAGT
 TTGCCTTCATCGAAAATCAGGTTATCCAGGCTAAAATCGCCGTGGGTACCACGCTATCCGGGCTAAAACG
 GCAGCAGTTTATGCATTTCTTTCCACACCTGTTCCACCGCCAGCCGTTACGTTTCATCATCAAAAATCGCT
 CGCATCCACCAGGCCGTTGTTTCATACGGCTCTGCGCCTGGGCCAGACGAAACACACGATCGCTGTTAAAC
 GGGCAGTTGCACACCGGAATGCTATGCAGACGACGCAGAAACACGGCCAGCGCATCCACAATGTTTTTCGC
 CGCTATCCGGATATTTCTTCCAGCACCTGAAACGCGTTTTTGCCCGGAATCGCGGTGGTCAGCAGCCACGC
 ATCATCCGGGTGCGAATAAAAATGTTTAATGGTCGGCAGCGGCATAAATTCGGTCAGCCAGTTCAGACGC
 ACCATTTTCATCGGTACATCGTTCGCCACGCTGCCTTTGCCATGTTTCAGAAACAGTTCGGCGCATCCG
 GTTTGCCATACAGACGATAAATGGTCGCGCCGCTCTGACCCACGTTATCACGCGCCCATTTATAGCCATA
 CAGATCCGCATCCATGTTGCTGTTTCAGACGCGGACGGCTACAGCTCGTTTCACGCTGAATATGGCTCATA
 ACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTTATTGTTTCATGATGATATATTTTTTATCTTGTG
 CAATGTAACATCAGAGATTTTGGACACAAGCGATGATAAGCTGTCCTCGAGTTAATTCGCTTACCCCAT
 TTCACGTGCGAAGGTACGGGCAAGATCACAGATGGTTCGGGATACTCCAGGCGTCGTCACATGCGTTTGC

ACCATGCCGCTAGTCAGTTGAAGCAGCGCATCACGACAGCCTGCTGGGCTTTGATCATAATCCAGAATAA
 AAACATGCATGCTGCGCAGACGCTTAGCCAGAACATCCAGCGCCCACGCGAAAACATTGTACAGATCCCC
 ATTTGGTGCCAGCAGCTCCGGCGCACGGAACAGGGTGAACAGGGTATCACCAATGTGCGGACGCGGACCT
 GCGTTAGACTGAGGTTCCGCACCCTGCGGAGGAACTGCCGTACCGCTCAGTTGACCCCAATCTTCACGCC
 AACTGCCACCGCATTCAGGTAACGACAGTGTTCGCCAGCAGGCCATATACACGACGAATCGCTGCCAG
 CATCGCCAGATCCAGGCGTTCACCAGGACGTTGACGTTTTGCCAGACGGTTCGATATGACGATCTTCTGGC
 AGCGCCCCCAGCACGATATTGGTCCCAGGCAGAGTAGGCGGAATCAGTGCAACGAATGCCAGGACGGCCT
 GTGGGGTCATAGAACCATCAGATAACGTGCAGCAGGATAGCACAGCAGAGCCGCAATCGGGTGACGATC
 AAAAATCAGGGTCAGGGCAGGCGGGCGGCATGGCTACTGCCGGCTTCACCACCAATATGCGGGCGCCAGA
 ACCGCATCGGTTACTGCGTAAGGCATACCCATCGTGATTTGCGCACTGGTCATCACTACGGCTGCATCAC
 CGGCGCTGATTTCCCCTTGGTCCAGACGATGCTGGGTCGTATAGATATTAGCAATAGTTTCACTCGCGCC
 CAGCACGCGCAATACGTCATCGGTTTCAGGCACATAAACAATATCATCACGGCTGCCAGCGCAACCAGT
 AATTGGGTGCTGGTGGTTTTACCCATACCATGCGGCCCATCAATGTATACGCGTAACAGGGTCGGCATCT
 TTTGCTCCGGACGAACCTCAGTGGCCTCCTGCTGACGACGCGGACGCAGTGCGGTACGACGATTAGAATG
 ACCACGGCTGCGCGCTGCCTGATCGAAAGCAGATGCATGCTGGTGACCCGGATAGCTCGCCATGGTGTA
 TCCTCCTAAAAGTTAAACAAAATTCTAGATATTATAATTGTTATCCGCTCACAAAGCAAATAAATTTTTC
 ATGATTTATCGATGCATGAAGCTCGTGTATGGGGCTGACTTCAGGTGGTCTGCGCGTAATCTCTTGTTC
 CGAAGGGATTACCCTGTTATCCCTAAGATGTGTATAAGAGACAG

Függelék 1. ábra. A módosított EZ-Tn5 transzpozon vázlatos rajza (a) és nukleotid szekvenciája (b). Tn5 ME: Tn5 transzpozáz felismerő szekvencia, ún. mozaik vég (ME = mosaic end); I-SceI: I-SceI meganukleáz felismerő hely; *km^R*: kanamicin rezisztencia gén; *hsvTK*: *Herpes simplex* vírus timidin-kináz gén; *hsvTK(mod)*: a *hsvTK* kodon-módosított duplikátuma.

Függelék 1. táblázat. A munka során felhasznált PCR primerek listája.

Név	Szekvencia (5'-3')	Felhasználás
pSG-ME-Sce_F	ATTACCCTGTTATCCCTAAGATGTGTATAAGAGA-CAGCTGGTAACTGCCGTCAGAGTCG	a pSG76-A plazmid linearizációja hiperaktív Tn5 fordított ismétlődések és I-SceI helyek hozzáadásával
pSG-ME-Sce_R	ATTACCCTGTTATCCCTAAGATGTGTATAAGAGA-CAGCTGATCCGGTCTCCCTATAGTG	
HKH_F	CCCTTCGCTGAACAAGAG	a <i>hsvTK-km^R-hsvTK_{mod}</i> szelekciós kazetta amplifikációja a pHKH plazmidről
HKH_R	CGATGCTGTCGGAAACAC	
T7	TAATACGACTCACTATAGGG	pST76-A-specifikus primerek genomi rekombinációk ellenőrzéséhez
pSTAR	GGATCTGAGGTTCTTATGGCT	
RD_1-BR	CTGGGGCGTCTTATGAGCCTGCTGTCACCCTTG	genomi homológia boxok amplifikációja a D1 deléció készítéséhez
RD_1-BF	CAGCAGGCTCATAAGACGCCCCAGCGTCGCCAT	
RD_1-C-BamHI	CCGGATCCTATCAGCTGTCCCTCCTGTT	
RD_1-A-EcoRI	GCGAATCCCGGTGTATCGGAATTTG	
RD_1-D	ACGCTTAATCACCTCATCATCTG	genomi primerek a D1 deléció létrejöttének ellenőrzéséhez
RD_1-E	ACGTGAAACATCCGGGCTGAAA	
RD_10-A-EcoRI	CCGAATTCACCACGATCAGAATGGAACC	genomi homológia boxok amplifikációja a D3 deléció készítéséhez
RD_10-C-BamHI	GCGGATCCATAAGCGATTCCCAGTGAC	
RD_10-D	GATGGTGAGGTTGATATTGT	genomi primerek a D3 deléció létrejöttének ellenőrzéséhez
RD_10-E	GCTGTAGCCAAATCGAACTG	
RD_6-A EcoRI	GCGAATTCACCTGACCGGCAAACCTTATC	genomi homológia boxok amplifikációja a D4 deléció készítéséhez
RD_6-C	CGCTATGTCCGCAATTTAC	

RD_6-D2	GTATGGCGGTAGCATTGTCC	genomi primerek a D4 deléció létrejöttének ellenőrzéséhez
RD_6-E	TCATCCTGTGGCGTACAAG	
RD_9-A3-BamHI	CCGGATCCCAGCTCGACCAAGAACAAG	genomi homológia boxok amplifikációja a D5 deléció készítéséhez
RD_9-C3-EcoRI	CCGAATCCGCAGTTGCACGTCATTCTC	
RD_9-C4-EcoRI	CCGAATCCGCACCGGAATCAGCATAAG	genomi primer a D5 deléció ellenőrzéséhez
RD_3-BR	ACCTGAACAAATTACAAAAACGTCTGCGTC	genomi homológia boxok amplifikációja a D6 deléció készítéséhez
RD_3-BF	GTTTTTGTAAATTTGTTTCAGTTGTATTGTTTC	
RD_3-A-EcoRI	CCGAATTCTCGGAGGCATACCTTTCATC	
RD_3-C-BamHI	CCGGATCCTAATACCTTCACTGCCCTAC	
RD_3-D2	TGCCGGGTGCAAGAGTAAAC	genomi primerek a D6 deléció létrejöttének ellenőrzéséhez
RD_3-E	GATGATATTGTCGCTTTAGG	
RD_2-BF	TTGTTGGCAAGGTAATGACTCCAATTATTGATAG	genomi homológia boxok amplifikációja a D7 deléció készítéséhez
RD_2-BR	AGTCATTACCTTGCCAACAATCTTTGATGGCTGAA	
RD_2-C-BamHI	GCGGATCCCATTATGGGCGTTGGCCTC	
RD_2-A-EcoRI	GGGAATTCGCTGCTGCAGTTTGAAG	
RD_2-D	TGCAGCTCCGGTTATCATCC	genomi primerek a D7 deléció létrejöttének ellenőrzéséhez
RD_2-E2	GGAAGCGAGAGTAATTAAC	
RD_4-A-EcoRI	GCGAATCCGGTAATCATCTGCGTTGTC	genomi homológia boxok amplifikációja a D9 deléció készítéséhez
RD_2-C-BamHI	GCGGATCCCATTATGGGCGTTGGCCTC	
RD_4-D	CCACCAGCAGAATCAATACC	genomi primerek a D9 deléció létrejöttének ellenőrzéséhez
RD_2-E2	GGAAGCGAGAGTAATTAAC	
RD_7-A-EcoRI	GCGAATCCCGGAATGGTGGCGACTATG	genomi homológia boxok amplifikációja a D10 deléció készítéséhez
RD_7-C-BamHI	CGGGATCCTCGAGCCGTTTTACCCCAAG	
RD_7-D	GAGTCCCAATAAGCGCCTCT	genomi primerek a D10 deléció létrejöttének ellenőrzéséhez
RD_7-E	TATGCCTGTTCCGAGATG	
RD_8-A-EcoRI	CGGAATTCTCGCATGGATGTGATGAAG	genomi homológia boxok amplifikációja a D11 deléció készítéséhez
RD_8-C-BamHI	GCGGATCCCGGATGAATGGCATAGTTGG	
RD_8-D	CTTCGTTACTGGCAAAAGG	genomi primerek a D11 deléció létrejöttének ellenőrzéséhez
RD_8-E	GTGCGGTAAGTCACCGTATC	
RD_11-A3	CCGCTGCGGTGTAGTAGTC	genomi homológia boxok amplifikációja a D12 deléció készítéséhez
RD_11-C3	GCGGATCCTGGCGAAGCGGAAAGATCTC	
RD_11-D	TGTACCGCCATCGCTTCAC	genomi primerek a D12 deléció létrejöttének ellenőrzéséhez
RD_11-E	CAGCCAACAGTTTGATAACG	
yehB-fw	GCGGATCCCATCAATCTTTACCGAAGG	genomi homológia boxok amplifikációja a mutáns <i>yehB</i> allél létrehozásához
yehB-rev	GCGAATCCGATGCCACTAAGTCGCATAGTAAA	

Függelék 2. táblázat. A szekvenálásra kiválasztott törzsek növekedési paraméterei. (a) erősen mutagén protokollal kapott MG1655-alapú törzsek, (b) gyengén mutagén protokollal kapott MG1655-alapú törzsek, (c) erősen mutagén protokollal kapott MDS42-alapú törzsek.

A növekedési ráta és OD-növekmény értékeket a szülői törzshöz hasonlítottuk. A feltüntetett átlagok és szórások 4 ismételt mérésen alapulnak.

a

Ciklusok száma	Kiválasztás alapja	Törzs	Relatív növekedési ráta		Relatív OD-növekmény	
			Átlag	Szórás	Átlag	Szórás
1 kontroll ciklus	jó növekedés	contr-56	1,080	0,025	1,081	0,019
		contr-59	1,079	0,033	1,090	0,019
1 ciklus	jó növekedés	c1-16	0,999	0,029	0,971	0,042
		c1-17	1,019	0,042	0,979	0,009
		c1-53	1,077	0,040	1,061	0,014
		c1-30	1,021	0,035	1,121	0,028
		c1-21	1,024	0,063	1,127	0,020
		c1-23	1,017	0,047	1,086	0,013
		c1-24	1,072	0,054	1,121	0,002
		c1-48	1,026	0,053	1,094	0,024
		c1-72	1,064	0,033	1,088	0,033
		c1-29	1,031	0,065	1,048	0,017
	c1-32	1,017	0,067	1,071	0,029	
	random	c1-5	0,671	0,014	0,795	0,076
		c1-9	1,009	0,052	0,579	0,047
		c1-10	0,717	0,052	0,253	0,017
		c1-11	1,039	0,064	0,693	0,087
		c1-12	0,724	0,096	0,218	0,039
		c1-15	0,992	0,031	0,994	0,030
		c1-6	0,623	0,076	0,356	0,014
		c1-14	0,840	0,014	0,565	0,012
		c1-7	0,989	0,037	1,065	0,004
c1-8		0,901	0,017	0,939	0,016	
c1-13	1,039	0,034	1,057	0,004		
2 ciklus	jó növekedés	c2-31	1,058	0,055	0,965	0,036
		c2-17	0,988	0,006	1,109	0,009
		c2-36	1,035	0,028	1,085	0,012
		c2-23	1,112	0,041	1,152	0,073
		c2-11	1,025	0,023	1,121	0,050
		c2-8	1,083	0,023	0,968	0,038
		c2-12	1,063	0,020	1,102	0,022
		c2-20	1,030	0,044	1,092	0,060
		c2-26	1,086	0,025	1,120	0,020
c2-32	1,107	0,049	0,947	0,052		
5 ciklus	jó növekedés	c5-158	0,841	0,058	1,042	0,028
		c5-170	0,962	0,044	1,134	0,035

b

Ciklusok száma	Kiválasztás alapja	Törzs	Relatív növekedési ráta		Relatív OD-növekmény	
			Átlag	Szórás	Átlag	Szórás
1 ciklus	jó növekedés	c1_lowmut-80	0,997	0,049	1,028	0,042
		c1_lowmut-87	1,055	0,041	0,763	0,045
	random	c1_lowmut-51	0,986	0,036	1,030	0,028
		c1_lowmut-52	0,968	0,035	1,024	0,015
		c1_lowmut-53	0,989	0,027	1,031	0,027
		c1_lowmut-54	0,998	0,026	1,040	0,026
		c1_lowmut-56	0,966	0,039	1,052	0,028
		c1_lowmut-57	1,041	0,022	0,762	0,025
		c1_lowmut-58	0,955	0,031	0,989	0,032
		c1_lowmut-59	0,981	0,017	1,026	0,035
		c1_lowmut-55	0,989	0,025	1,044	0,017
		c1_lowmut-60	0,974	0,024	1,017	0,033

c

Ciklusok száma	Kiválasztás alapja	Törzs	Relatív növekedési ráta		Relatív OD-növekmény	
			Átlag	Szórás	Átlag	Szórás
1 ciklus	random	c1_MDS42_1	0,995	0,001	1,051	0,067
		c1_MDS42_2	1,005	0,015	1,052	0,091
		c1_MDS42_3	1,008	0,033	1,074	0,053
		c1_MDS42_4	0,992	0,019	1,017	0,005
		c1_MDS42_5	1,003	0,027	1,006	0,045
		c1_MDS42_6	0,990	0,010	1,049	0,070
		c1_MDS42_8	0,995	0,017	1,017	0,010
		c1_MDS42_10	0,993	0,020	1,053	0,040
		c1_MDS42_18	0,973	0,017	1,004	0,041
		c1_MDS42_7	0,963	0,017	0,985	0,023
		c1_MDS42_14	0,952	0,016	1,028	0,024
		c1_MDS42_16	0,982	0,012	1,071	0,016