A kompenzáló evolúció rendszerszintű vizsgálata élesztőben

Boross Gábor Zoltán

Ph.D. értekezés

Témavezető: Dr. Papp Balázs

Biológia Doktori Iskola
MTA Szegedi Biológiai Kutatóközpont, Biokémiai Intézet
Szegedi Tudományegyetem, Természettudományi és Informatikai Kar
Szeged, 2016

Tartalomjegyzék

1. Bevezetés 3
1.1. Az adaptáció elméleti modelljei . . . 3
1.1.1. Fisher geometriai modellje . . . 4
1.1.2. DNS szekvencia alapú modellek . . . 9
1.1.3. A geometriai modell és a DNS szekvencia alapú modellek összehasonlítása . . . 13
1.2. A kompenzáló evolúció . . . 14
1.3. Kísérletes munkák a kompenzáció vizsgálatára . . . 20
1.4. Géndeléció kompenzációjának vizsgálata sörélesztőben . . . 27

2. Célkitűzés 29

3. Anyag és Módszer 31
3.1. Felhasznált élesztő törzsek . . . 31
3.2. Tápoldatok és táptalajok . . . 31
3.3. Evolúció a laboratóriumban . . . 31
3.4. Nagy skálájú fitneszmérések . . . 33
3.5. Kompenzációt mutató gének azonosítása . . . 34
3.6. Környezeti profilok meghatározása . . . 34
3.7. Genomszekvenálás . . . 35
3.8. Pozitív szelekció azonosítása . . . 36
3.9. Transzkriptomikai vizsgálat . . . 37
3.10. Bioinformatikai analízis során használt adatsorok . . . 37

4. Eredmények 39
4.1. Az élesztő géneknek csak kis hányada alkalmas a kompenzáló evolúció laboratóriumi vizsgálatára . . . 39
4.2. A kompenzáló evolúció génvesztés után gyakori, és gyorsan bekövetkezik 39
4.3. Nem minden gén kompenzálható egyformán . . . 42
4.4. A gének pleiotrópiája nem befolyásolja a kompenzációt . . . 44
4.5. A kompenzáló evolúció általában nem a paralóg génekre hat . . . 44
4.6. A génkiütés fitneszköltsége korrelál a kompenzáció sebességével . . . 44
4.7. A kompenzáló evolúció elősegíti a genomok diverzifikációját . . . 47
4.8. A kompenzáló evolúció nem állítja vissza a gének expresszióját a vad típus szintjére . . . 49
4.9. A kompenzáló evolúció diverzifikációs hatása változatos környezetekben felderíthető . . . 52

5. Diszkusszió 57

6. Köszönetnyilvánítás 63

7. Hivatkozások 64

8. Összefoglaló 76

9. Summary 79

10. Függelék 82

10.1. A genomszekvenálásra és transzkriptomikai vizsgálatra kiválasztott evolúciós vonalak 82

10.2. A gének kompenzálhatóságban mutatott különbözősége nem pusztán a géndeléció fitneszhatásától függ 83

1. Bevezetés

A természetes szelekció által megvalósuló adaptáció során egy élőlény a környezetéhez alkalmazkodik, ezáltal folyamatosan növelve rátermettségét. Ugyanakkor káros hatású, a rátermettséget csökkentő mutációk is elterjedhetnek természetes populációkban. Ezek a káros mutációk kompenzáló evolúciót eredményezhetnek. Kompenzáló evolúciónak nevezzük azt a jelenséget, amikor egy mutáció káros hatását a genom egyéb területein egy másik mutáció részben, vagy egészben kijavítja, ezért elterjed a populációban. Tehát a kompenzáló evolúció tulajdonképpen az adaptív evolúció egy formája, melyben az adaptáció a genetikai háttér megváltozására történik.

1.1. Az adaptáció elméleti modelljei

Adaptációnak nevezzük, amikor egy populáció összetétele olyan fenotípus irányába mozog, amely jobban megfelel a környezetének. Adaptáltnak tekintünk egy adott jelleget, ha az az egyednek az adott környezetben, illetve azon környezetek összességében, amelyet élete során megtapasztal, maximális rátermettséget (fitneszt) biztosít [Fisher, 1930]. Charles Darwin is részben az élőlények adaptált jellegeinek megfigyelésén keresztül fogalmazta meg a természetes szelekció segítségével működő evolúció elvét [Darwin, 1859]. Az adaptációk természetes szelekció segítségével terjednek el a populációkban (kivéve pre-adaptációk [Gray et al., 2010], lásd később), ugyanakkor a természetes szelekció nem okoz mindig adaptációt. Gondolhatunk ilyenkor az önző genetikai elemekre [Hurst, 1992]. Például egy transzpozon elterjedése a genomban a természetes szelekció eredménye, hiszen a gyorsabban duplikálódó transzpozonok fognak fennmaradni egy élőlény genomjában, ugyanúgy, ahogy egy gyorsabban szaporodó egyed utódai elterjedhetnek egy állati populációban. Azonban egy transzpozon elterjedését (hacsak nem derül fény az organizmusra előnyös voltára) nem nevezhetjük adaptációnak, hiszen nem segíti elő az egyed túlélését és szaporodását környezetében. Mi több, "önző" gének közötti kompetíció a populáció ivararányát is megváltoztathatja [Dyson and Hurst, 2004], amely egy fenotípus szintjén jelentkező különbség, melynek adaptív (azaz az egyed rátermettségét környezetében növelő) voltát semmi nem indokolja.

Mivel a kompenzáció is egy adaptív folyamat, a kompenzáló evolúciót vizsgáló elméleti és kísérleti kutatások irodalmának áttekintése előtt, néhány klasszikus elméleti adaptációs modellt mutatok be. Ezek az elméletek az adaptációs folyamatokat, evolúcióbiológiai és genetikai ismereteink felhasználásával, egy matematikai modellben igyekeznek leírni. A modellek egyszerűsítenek, és sok szempontból nem realisztikusak, ezért predikcióikat mindenképpen fenntartással kell kezelnünk, ugyanakkor kutatásunknak adnak egy olyan koncepcionális vázat, amely mentén kvantitatívan tudjuk vizsgálni a kompenzáló evolúciót, illetve ami segítségünkre lesz a korábbi kutatási eredmények és saját méréseink

értelmezésekor.

A fenotípus szintjén megjelenő adaptációk adják az evolúció működésének legismertebb példáit, melyeket megfigyelve kitűnik, hogy egyes élőlények milyen változatos módokon alkalmazkodtak környezetükhöz. Ugyanakkor nem csak a zsiráf hosszú nyaka, vagy a darazsakat utánzó zengőlegyek színezete adaptáció eredménye, hanem sok más, kevésbé szembetűnő jelleg. Felmerül a kérdés, hogy vajon milyen mechanizmus állhat az adaptáció mögött. Bár Darwinnak a 19. században még nem voltak helyes ismeretei az öröklődés genetikai mechanizmusáról, mégis tőle származik a gradualizmus elve, miszerint az evolúció nagyon kis lépésekben történik. Darwin szerint a természetes szelekció csak egymás utáni kis variációkon keresztül tud előrelépni, sosem ugorhat előre, kis és lassú lépéseken keresztül történik az adaptáció [Darwin, 1859]. Ez a nézet vezetett a kvantitatív jellegeket statisztikai eszközökkel vizsgáló biometria kialakulásához. A biometria tudományág képviselői a folytonos jellemzők (úgymint testmagasság) örökletességének mértékét, a szelekció általi megváltozásukat (szelekcióra adott válasz) vizsgálják. Leghíresebb képviselőjük Karl Pearson több új statisztikai módszert is kidolgozott munkássága során.

A tudományterületen problémát jelentett a biometria eredményeit a mendeli öröklődés szabályaival egyeztetni. Hogyan értelmezhetjük a folytonos változók apró különbségeinek öröklődését a mendeli genetika szabályaival? Az ellentmondás feloldására Ronald Fisher dolgozta ki az "infinitesimal" modell néven elhíresült elméletet [Fisher, 1918], mely egy elméleti, matematikai modell. Ebben feltételezi, hogy a folytonos, kvantitatív változókat végtelenül sok gén alléljai határozzák meg, ahol minden gén hatása végtelenül kicsi, a gének között interakció nincs, az allélok között a dominancia különböző mértékű lehet. A modell természetesen sok feltevésében (végtelen sok gén, végtelenül kicsi hatással a fenotípusra, interakciók hiánya a gének között) nem tükrözi a valóságot. Mint minden modell, egyszerűsít, idealizálja a modellezni kívánt világot, cserébe segít megmagyarázni, megérteni egyes megfigyelt jelenségeket. Fisher "infinitesimal" modellje által megérthető, hogyan lehetséges a mendeli genetikai szabályok mentén folytonos fenotípusos változók öröklődését elképzelni. Ugyanakkor, sok, az adaptációval kapcsolatban felmerülő kérdés és megfigyelt jelenség kívül esik az "infinitesimal" modell látókörén. Az adaptáció elméleti megértésében a következő előrelépést ugyancsak Ronald Fisher szolgáltatta.

1.1.1. Fisher geometriai modellje

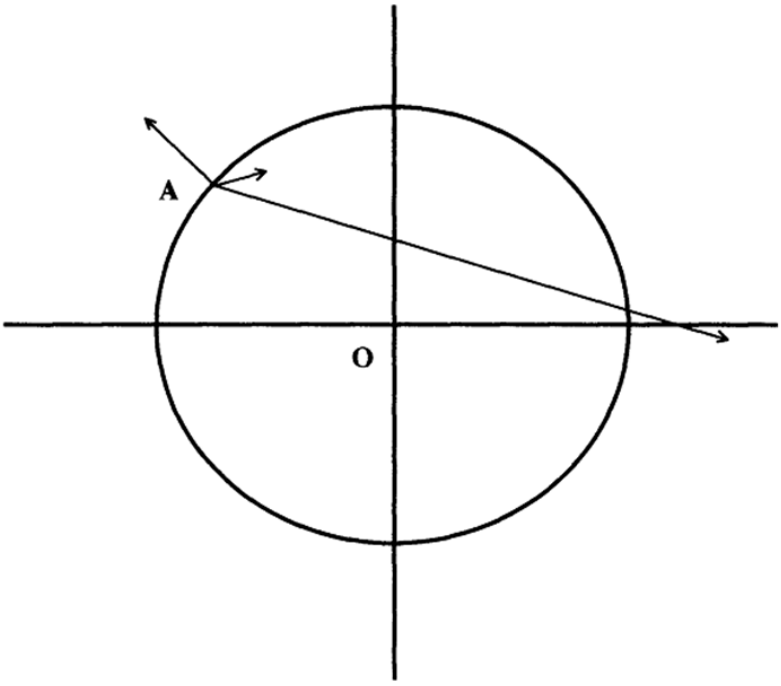
Fisher 1930-ban jelentette meg könyvét, a "The genetical theory of natural selection"-t [Fisher, 1930], melyben az adaptációkkal részletesen foglalkozik. Fisher akkor tekint egy egyedet adaptáltnak egy adott szituációhoz, vagy szituációk összességéhez, amely az egyed környezetét adja, ha el tudunk képzelni olyan kicsit különböző szituációkat, vagy környezeteket, amelyekhez kevésbé lenne az egyed adaptált. Vagy, hasonlóképpen, ha el

tudunk képzelni kissé különböző organizmusokat, amely az adott környezethez kevésbé adaptált. Az adaptációt, mely során egy organizmus egy környezethez több különböző tényező mentén hasonul, egy geometriai modellel illusztrálja. Ebben a geometriai modellben a hasonulás mértékét a távolság mutatja. Minél közelebb van A pont O origóhoz egy koordináta-rendszerben, annál magasabb fokú adaptációról beszélhetünk. A dimenziók (koordinátatengelyek) száma az adott, egymástól független fenotípusos jellegek száma, amelyekre az adaptáció zajlik. A pont reprezentálja az organizmus által mutatott fenotípusos jellegek állapotát, míg az origó mutatja a tökéletesen adaptált állapotot. Adaptációkor tehát A pont az origót (O) közelíti. Ha ezt a jelenséget egy háromdimenziós térben akarjuk ábrázolni, az azt jelenti, hogy 3 jelleg alapján történik az adaptáció. Ez természetesen sok esetben egy egyszerűsítés, de három dimenzióban is bemutatható a modell alapvető működése és a belőle levonható következtetések.

Képzeljük el, hogy A pont d/2 távolságra van az origótól. Ha meg akarunk találni a koordináta-rendszeren belül minden olyan pontot, ami A-nál jobban adaptált állapotot jelez, az minden olyan pont lesz, amely közelebb van O-hoz, mint A. Ez pedig egy olyan gömbön belül eső pontok halmaza, mely gömbnek középpontja O, és a gömb áthalad A-n, azaz sugara d/2, átmérője d (lásd 1. ábra). A pont mutációkkal mozog a térben. Egy mutációt egy adott vektorral történő elmozdulásként értelmezhetünk a modellben. Fontos, hogy ezek a mutációk véletlenszerűen történnek, nincs kiemelt irányuk. Nem mozdul gyakrabban, vagy könnyebben A pont az adaptációt jelentő origó felé, mint attól elfele. Az adaptációt jelentő mutációk, melyeket a természetes szelekció kiválogat, azok a mutációk lesznek, melyek közelebb viszik A-t O-hoz, ezáltal magasabb rátermettséget okozva az organizmusnak. Előbb, vagy utóbb megjelenhet egy mutáció, amely A-t O-hoz közelebb viszi, ez a mutáns a természetes szelekció által elterjed a populációban (két dimenzióban szemléltetve a 1. ábrán). Ennek a folyamatnak az ismétlődése az adaptáció Fisher modelljében.

Fisher feltette a kérdést, vajon a mutációknak mekkora része okoz fitneszemelkedést? Modellje azt mutatja, hogy egy adott mutáció esélye arra, hogy A-t O-hoz közelebb mozdítsa, attól függ, hogy mekkora a mutáció által okozott fenotípusos hatás mértéke, azaz a koordináta-rendszerben mekkora az elmozdulás vektorának hossza. Ezt a hosszt jelöljük r-rel[1]. A kérdés tehát az, hogy milyen összefüggés van a mutáció hatáserőssége (r) és annak valószínűsége (P) között, hogy A-t a mutáció O-hoz közelebb, azaz a fent leírt gömbön belülre viszi. Könnyű elképzelni, hogy ha r nagyobb, mint d, azaz a mutációs hatás nagyobb, mint a gömb átmérője, akkor annak az esélye, hogy a mutáció emelje az organizmus fitneszét nulla. Hiszen A a gömb felszínén helyezkedik el, innen bármely

[1] r nem a gömb sugarát jelenti, hanem a mutáció általi fenotípusos változás nagyságát, A elmozdulásának hosszát a koordináta rendszerünkben. Azért nevezzük mégis r-nek ezt a változót, hogy Fisher eredeti nevezéktanát megtartsuk.



1. ábra. **Fisher geometriai modelljének két dimenziós szemléltetése.**
Az organizmus A pontban helyezkedik el, O jelzi az optimumot két fenotípusos karakterre nézve, melyek felvett értékét a két koordinátatengely szemlélteti. A fenotípus véletlen irányú és nagyságú lépésekkel (mutációkkal) tud mozogni a síkon. Azon változások lesznek előnyösek, melyek A-t O-hoz közelebb viszik, azaz az ábrázolt körön belülre mutatnak. Ábra forrása: [Orr, 1998]

irányba mozdítjuk el A-t a gömb átmérőjénél nagyobb távolságra, nem helyezkedhet el A a gömbön belül (lásd ezt az esetet két dimenzióban szemléltetve a 1. ábrán). A másik véglet, ha r egy nagyon kicsi érték. Ilyenkor A r sugarú környezetében a gömb közelíthető egy síkkal, amely tartalmazza A-t. A képzeletbeli sík egyik oldala a gömbön belüli pontokat tartalmazza, másik oldala a gömbön kívülieket (ez a közelítés csak A pont nagyon szűk környezetében igaz, ha r egy nagyon kicsi érték). Annak az esélye, hogy a sík egy pontjából véletlenszerűen elmozdulva a sík egyik, vagy másik oldalára érkezünk, 0,5. Tehát ahogy r közelíti a 0-t, úgy közelíti P a 0,5 -öt. Fisher megadja e két szélsőséges érték között ($0 \leq r \leq d$) P értékét:

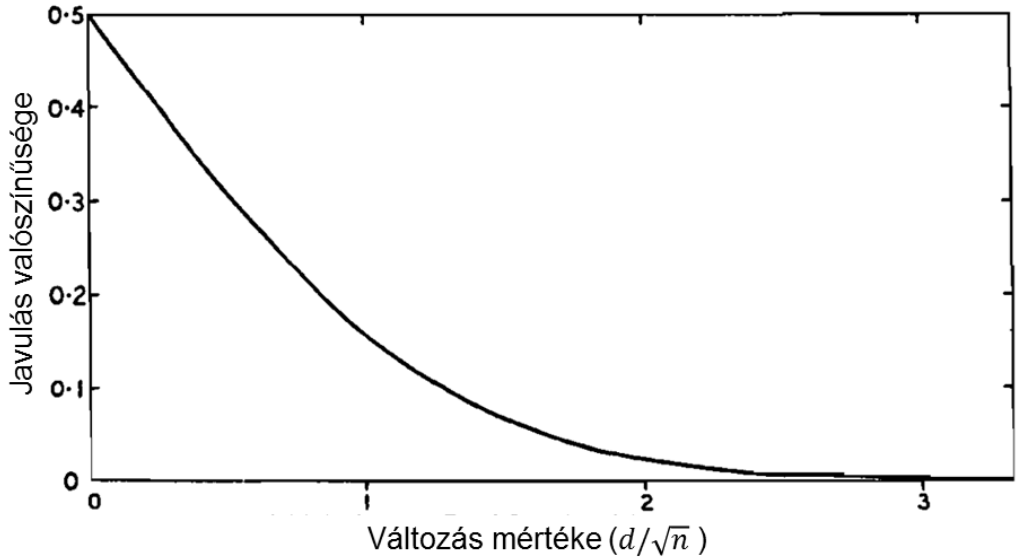
$$P = \frac{1}{2}(1 - \frac{r}{d}) \tag{1}$$

Természetesen a három dimenzió egy adaptációs eseményhez kevés. Általában nem

csak három karakter mentén történik az adaptáció. Több dimenzió mentén megvizsgálva ugyanezt az összefüggést, Fisher azt találta, hogy az alapvető tulajdonsága a fenti összefüggésnek megmarad magasabb dimenziók esetén is. Nagyon kicsi r esetén az előnyös mutációra az esély 50%, majd r növekedésével a valószínűség gyorsan csökken. Azt találta ugyanakkor, hogy minél magasabb a fenotípusos karakterek (dimenziók) száma, ami mentén az adaptáció történik, annál gyorsabban csökken az előnyös változás valószínűsége (2. ábra). Emellett több fenotípusos karakter esetén azért is lassabb lesz az adaptáció, mert egy koordinátatengely mentén az optimum felé történő lépés kisebb hatású lesz [Orr, 2000]. Ezt legkönnyebben úgy tudjuk demonstrálni, ha elképzelünk egy csupán egy karaktert tartalmazó modellt, ahol A pont az origótól x távolságra van. Itt az adott karakter mentén történő tökéletes adaptáció (x hosszúságú lépés) az origóba viszi A-t. Ha most elképzelünk egy kétdimenziós modellt, ahol két karakter mentén történhet adaptáció és mindkét dimenzió szerint x távolságra van A az origótól, akkor beláthatjuk, hogy az egyik karakter mentén x nagyságú lépés A pontot $\sqrt{2}$ x távolságból x távolságra viszi. Ebben az esetben tehát egy x hosszúságú adaptációs lépés az A pontot csak ~0.4 x hosszal vitte közelebb az origóhoz. A további dimenziónövekedések is hasonló elv szerint fogják csökkenteni az egy karakter mentén történő adaptációs lépések hatékonyságát.

Fisher geometriai adaptációs modellje eszerint azt jósolja, hogy az adaptáció kis változások révén megy végbe, hiszen a genetikai változatosság, az utódokban keletkező mutációk nem irányított események, és a nagy hatású mutációk esélye arra, hogy előnyösek legyenek nagyon kicsi. Ez a modell tulajdonképpen megerősítette a Darwintól származó gradualizmus elvét, miszerint az evolúció kis lépésekben történik.

Azonban ebből a gondolatmenetből hiányzik a populáció szintű értelmezés, ami a populációgenetika alapja. A modell egy alapvető revíziója a japán evolúcióbiológus, Motoo Kimura nevéhez fűződik [Kimura, 1983]. Kimura, a molekuláris evolúció neutrális elméletének megalkotója, munkássága során kiemelte a genetikai sodródás fontosságát a molekulák evolúciójában. Kimutatta, hogy a mai rokon fajokban megfigyelt molekuláris különbségek nagy része nem lehet a természetes szelekció eredménye, hanem ezek neutrális mutációk elterjedésére vezethetők vissza, melyeknek megjelenésükkor nem volt hatásuk az organizmus rátermettségére és a genetikai sodródás által fixálódtak a populációban. Ez a molekuláris evolúció neutrális elméletének egyik alappillére, mely azt mondja ki, hogy a molekuláris (nem fenotípus) szinten megfigyelt rokon fajok közötti különbségek döntő többsége neutrális evolúció eredménye, és a szubsztitúciós ráta csak a mutációs rátától függ. Később ezt Tomoko Ohta kiegészítette, a nem egészen neutrális, de csak kis mértékben káros mutációk fontosságát hangsúlyozva, miszerint populációmérettől függően, de ezek is fixálódhatnak a fajokban a genetikai sodródás által [Ohta, 1973, Ohta, 1992]. A neutrális elmélet jelenleg is vita tárgyát képezi. A vita azonban valójában csak arról szól, hogy milyen arányban vannak jelen adaptív szubsztitúciók a neutrálisok mellett.

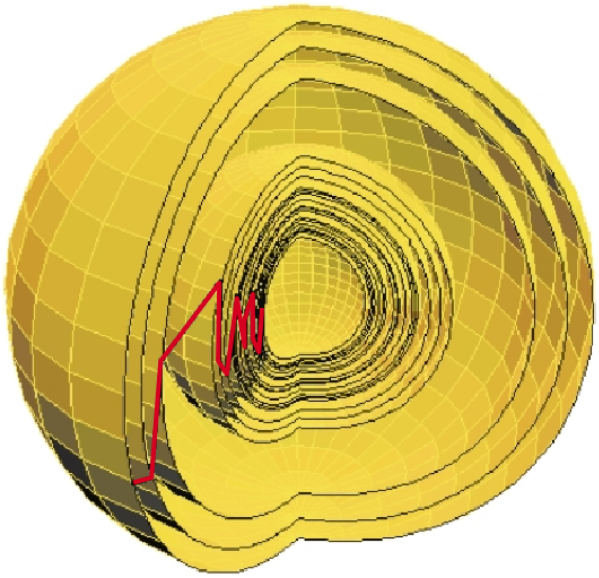


2. ábra. **Az előnyös változás valószínűségének a változás nagyságától és a fenotípusos változók számától való függése.**
Az x tengelyen bemutatott változó, a mutáció által okozott fenotípusos változás nagysága, amely normálva van a dimenziók (fenotípusos változók) számára. Egysége d/$\sqrt{n}$, ahol d A távolsága az origótól, n pedig a dimenziók száma. Ábra forrása: [Fisher, 1930]

Kimura és Ohta munkássága előtérbe helyezte a genetikai sodródás fontosságát az evolúció vizsgálatában. Ezek után nem meglepő, hogy Kimura volt az, aki a populációgenetika egyik atyja, Ronald Fisher geometriai adaptációs modelljét megvizsgálta a genetikai sodródás szemszögéből. Míg Fisher modellje azt mondja, hogy a kis hatású mutációknak nagyobb az esélyük, hogy adaptívak legyenek, azaz növeljék az organizmus rátermettségét, addig a genetikai sodródás miatt éppen a kis hatású mutációk lesznek azok, amelyek, bár adaptívak, de nagyobb eséllyel fognak eltűnni egy véges méretű populációból a genetikai sodródás miatt. Tehát a kis hatású mutációk nagyobb eséllyel növelik a fitneszt, de kisebb eséllyel fixálódnak a populációban, míg a nagy hatású mutációk csak kis eséllyel viszik az organizmust közelebb az optimumhoz, kis eséllyel növelik a fitneszét, de ez esetben viszont kisebb az esélye annak is, hogy a genetikai sodródás áldozatává váljanak. Ezek alapján Kimura azt a következtetést vonta le, hogy a közepes hatással bíró mutációk okozzák az adaptáció jelentős részét [Kimura, 1983].

Valójában még ez a megközelítés sem egészen pontos, mivel csak az adaptáció első lépésével, az első szubsztitúcióval foglalkozik. Ugyanakkor az adaptáció során több előnyös

mutáció is fixálódhat, ami által a populáció egyre közelebb kerül az optimumhoz. Elméleti kutatások, melyek szimuláltak ilyen adaptációs utakat Fisher geometriai modelljében, arra a következtetésre jutottak, hogy az egymás után fixálódó mutációk hatásai közel exponenciális görbe mentén csökkennek [Orr, 1998, Orr, 1999]. Ez azt jelenti, hogy általában az első fixálódó mutáció viszonylag nagyobb hatású. A második fixálódó változás mértéke az elsőnek csak egy bizonyos hányada, a harmadik a másodiknak egy hasonló hányada és így tovább. Ahogy közeledünk az optimumhoz, exponenciálisan csökkenő mértékű adaptációs változások következnek be ("finomhangolás"). Egy ilyen adaptációs utat mutat be három dimenzióban a 3. ábra.

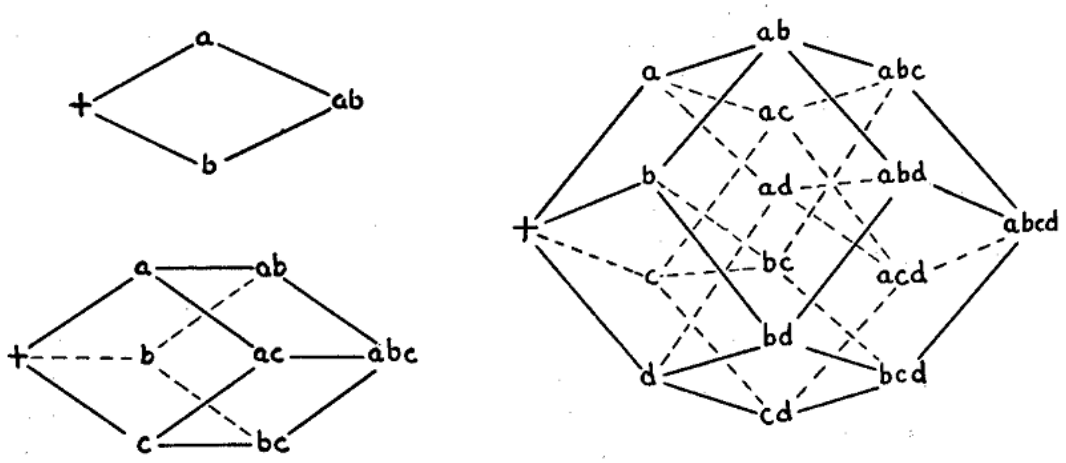


3. ábra. **Egy adaptációs útvonal Fisher geometriai modelljében.**
Fisher geometriai modelljét három dimenzióban szemléltetve az ábra egy adaptációs útvonalat mutat. Az optimum a gömb közepe, az organizmus az ábrázolt legnagyobb gömb felszínéről indul és adaptációs lépésekkel (piros vonalak) halad a gömb közepe felé. Ábra forrása: [Barton and Keightley, 2002]

1.1.2. DNS szekvencia alapú modellek

Fisher a mendeli öröklődést használja geometriai modelljében, de az öröklődés molekuláris alapját nem. A genetikai információ átadásáért felelős örökítőanyag, a DNS, egy lineáris molekula, amely a genetikai kód által meghatározott módon fehérjeszekvenciákat kódol. Ezek ugyancsak lineáris szekvenciái aminosavaknak. Ez a rendszer diszkrét elemekből

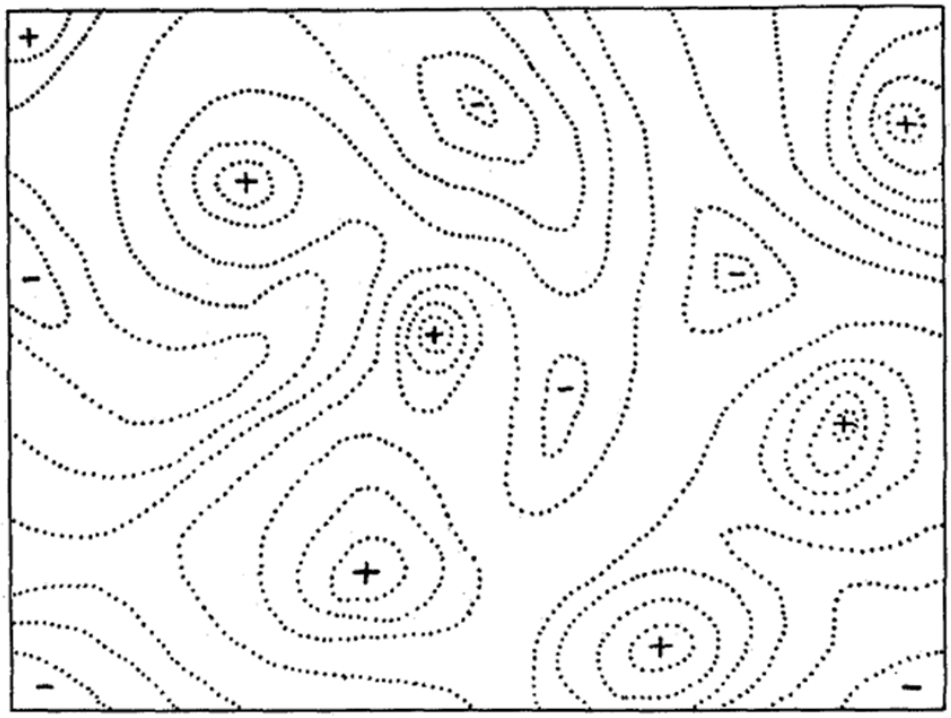
áll, bár nagyon sok, de mégis véges variációja létezhet a különböző DNS és fehérje szekvenciáknak. A hatvanas évek elején kezdték kidolgozni a DNS és fehérje alapú modelleket, ezekhez a koncepcionális alapot azonban sokkal korábban alkotta meg Sewall Wright. Wright 1932-ben [Wright, 1932] még nem DNS, vagy fehérje szekvenciáról írt modelljében, hanem gén kombinációkról. Ezen génkombinációk egy hálózatot alkotnak, amelyben az egymástól egy mutációs távolságra lévő "genotípusok", amelyek egy egyszerű mutációval egymásból elérhetők, össze vannak kötve. Minden génből két változatot feltételezve a hálózatot "legkönnyebben" egy sokdimenziós térben lehet elképzelni, ahol a genotípusok egy annyi dimenziós kockával reprezentálhatóak, ahány génünk van (4. ábra)



4. ábra. **Génkombinációk hálózata Wright modelljében.**
A genotípusok kombinációja reprezentálható egy hálózattal, ahol minden génkombináció össze van kötve, amelyet egymástól egy mutáció választ el. Ez génenként (a,b,c és d) két változat esetén egy annyi dimenziós kockával ábrázolható legkönnyebben, ahány génünk van. (Az ábra jobb oldalán egy négydimenziós kockát látunk 2 dimenzióban ábrázolva. Wright eredeti publikációjában az 5 dimenziós kockát is ábrázolta a síkban.) Ábra módosítva [Wright, 1932] alapján.

Wright ezek után azt írja, hogy minden génkombinációhoz rendeljünk egy fitneszértéket. Ezt ábrázolhatjuk még egy dimenzióban. Ha meg akarjuk könnyíteni a dolgunkat, elképzelhetjük a génkombinációkat 2 dimenzióban. Ezt tegyük meg egyszerűen úgy, hogy az egymástól kis mutációs távolságra lévő génkombinációk közel, a sok mutációs lépésre lévő génmutációk pedig egymástól távol legyenek a síkon. Ezek után minden genotípus fitnesze lesz a harmadik dimenzió. Ezáltal egy úgynevezett fitnesztájképet (nevezik adaptív tájképnek is) kapunk, ahol a magas fitneszű genotípusok csúcsokként jelennek meg (5. ábra). Ez a tájkép koncepció Wright-tól ered és meghatározó lett a későbbi adaptációs kutatásokban. Az adaptációt egy ilyen tájképen megtett útként értelmezhetjük. Adaptáció során az organizmus környezetéhez alkalmazkodik, rátermettsége emelkedik. A

tájképen ez egy csúcs megmászását jelenti, hasonlatosan ahhoz, ahogy Fisher modelljében az organizmus megközelítette a gömb középpontját.



5. ábra. **Wright fitnesztájképe.**
A genotípusokat két dimenzióra vetítve és mindegyikhez egy fitneszértéket rendelve, amelyet a harmadik dimenzióban ábrázolunk, egy fitnesztájképet kapunk, amelyek adaptív csúcsokból és maladaptív völgyekből állnak. Ábra forrása: [Wright, 1932]

A fitnesztájkép szolgáltatta az alapkoncepciót a későbbi DNS (és fehérje) szekvencia alapú adaptációs modellekhez. Ezekben a tájkép alapját nem génkombinációk, hanem konkrét szekvenciák adják. Bár a lehetséges szekvenciáknak száma véges, egy N hosszúságú szekvencia esetén 4^N darab szekvencia lehetséges, ami egy óriási szám. John Maynard Smith azonban úgy csökkentette le a lehetséges mutációs lépések számát a modellben [Maynard Smith, 1970], hogy ez a szám 3N legyen. Ezt pedig úgy tette meg, hogy feltételezte, hogy a mutációs ráta elég alacsony és a szelekció elég erős ahhoz, hogy egyszerre maximum egy mutáció legyen jelen a populációs minden egyes egyedében. A megjelenő mutációk vagy negatív szelekcióval kiszelektálódnak, vagy pozitív szelekcióval fixálódnak a populációban. Ezek az úgynevezett "strong selection-weak mutation (SSWM)" feltételek. Ekkor tehát azt feltételezhetjük, hogy egy adott populáció egy ponton van a tájképen és onnan csak az egy egységnyi távolságra lévő pontokat "éri el", azaz mutációk sorával az egy nukleotid mutációs távolságra lévő szekvenciákat tudja feltérképezni. Ilyen szekvenciából pedig 3N darab van, hiszen egy bázis mutálódhat egy másik bázissá (3) a

szekvencia teljes hosszán (N). Maynard Smith ezzel jelentősen lecsökkentette a variációk számát. Hogy az SSWM feltételek milyen mértékben teljesülnek a természetben és együtt elterjedő többszörös mutációknak mekkora szerepe van az adaptációban, ez kérdéses. Mindenesetre Maynard Smith modelljében, a populáció a környezetében lévő pontok közül azok egyikére fog "átlépni" mutáció és természetes szelekción keresztüli fixálódás által, amelyek fitnesze nagyobb, mint a jelenlegi vad típusé. Természetesen a genetikai sodródás miatt még adaptív mutációk sem mindig fixálódnak, de a modell ezt figyelmen kívül hagyhatja. A genetikai sodródás lassíthatja az adaptáció menetét, de a mutációk ismétlődhetnek. Előbb, vagy utóbb fixálódik egy előnyösebb változat, ezáltal megtörténik egy lépése az úgynevezett "adaptív sétának".

A modellt később Stuart Kauffman és Simon Levin továbbfejlesztette. Kihangsúlyozták, hogy a tájkép egyáltalán nem biztos, hogy olyan egyenletes lesz, ahogy azt korábban gondoltuk [Kauffman and Levin, 1987]. Gének (mutációk) közötti episztázis, kapcsoltság és a rekombináció is eredményezheti azt, hogy a tájkép "rögösebb" legyen[2]. Kidolgoztak egy olyan modellcsaládot, amelyben ez a rögösség egy paraméter mentén állítható, ezáltal robusztusabb következtetések levonására alkalmasak. Arra jutottak, hogy összetettebb organizmusok esetén nehezebb természetes szelekció által a mutációs nyomást ellensúlyozni. Az összetettség növekedésével az elérhető optimum kevésbé rátermett lesz és ezen optimum közelében is nehezebben tartható a populáció. Ez hasonló Fisher modelljében ahhoz az eredményhez, miszerint egy összetettebb organizmus (több fenotípusos karakter, azaz több dimenzió) esetén lassabb lesz az adaptáció, hiszen egy koordinátatengely mentén az optimum felé történő lépés a karakterek számának gyökével skálázódik [Orr, 2000] (ugyanakkor ellenérvként lásd: [Wang et al., 2010]).

A DNS szekvencia alapú modellekben komoly problémát jelentett, hogy milyen fitneszértékeket rendeljünk az egyes genotípusokhoz. Milyen lehet a fitneszeloszlása egy adott vad típusú organizmus szekvencia térben lévő környezetének, ahova egy mutációval eljuthat. John Gillespie a valószínűségszámítás területéről az extrémérték-elméletet alkalmazva bemutatta, hogy szinte teljesen mindegy milyen eloszlást mutatnak a fitneszértékek [Gillespie, 1984]. Ha a kiindulási pontunk ezen az eloszláson egy magas érték, akkor az adaptáció során fixálódott mutációk fitneszhatásai az összes vad típussal szomszédos genotípus fitneszérték eloszlásától független lesz. Az a feltevés, hogy a jelenlegi vad típus magas fitneszű legyen, azaz mutációs szomszédainak többsége nála rosszabb fenotípusú legyen, egyáltalán nem valószínűtlen. Logikus azt feltételeznünk, hogy a jelenlegi vad típus véletlen szekvenciáknál magasabb fitneszt mutat. Az extrémérték-elmélet szerint majdnem mindegy milyen eloszlásból vesszük mintáinkat, ha azokat az eloszlás extrém értékeiből

[2]Ez a meggondolás, hogy a mutációk fenotípusos hatásai egymástól függhetnek, hiányzott Fisher geometriai modelljéből is, amely nem is foglalkozott a genotípus fenotípus leképezéssel. Fisher modellje csak egy fenotípus térben rendel egy adott ponthoz fitneszértéket az optimumtól való távolság által.

vesszük, akkor az így kapott extrém értékek eloszlása nagyon hasonló lesz. Gillespie felismerte, hogy egy adaptációs séta esetén a lehetséges mutációk többsége rontani fogja a vad típus fitneszét, és a mutációknak csak egy nagyon kis része fogja emelni. Ezek fognak fixálódni a populációban és kialakítani az új vad típust. Ezek a mutációk ugyanakkor a korábbi vad típus összes mutációs szomszédai közül extrémek fitnesz tekintetében. Ahogy az adaptáció során egyre nő a fitneszérték, egyre nehezebben találni adaptív mutációt, ezért egyre lassabbak lesznek az adaptív lépések. A fixálódott adaptív mutációk gyakorisága egyre csökken, mindemellett az általuk elért fitnesznövekmény is exponenciálisan csökken [Gillespie, 1984, Orr, 2002, Orr, 2005b]. Ez a konklúzió ismét hasonlít Fisher geometriai modelljéhez abban a tekintetben, hogy egyre ritkább és egyre kisebb fitnesznövekmény figyelhető meg az adaptáció során. Ugyanakkor Fisher modellje folytonos volt, ott az adaptáció egy soha meg nem álló folyamat, mindig lehet az origóhoz közelebb jutni. A DNS alapú modellek viszont természetüknél fogva diszkrétek, itt el lehet érni a (lokális) optimumot, egy olyan pozíciót a genotípus térben, amelynek minden szomszédja rosszabb fitneszt mutat, mint önmaga. Ezért megvizsgálható számítógépes szimulációval, hogy egy adaptációs séta alatt átlagosan hány mutáció történik. Számítógépes szimulációk arra utalnak, hogy az első adaptációs lépés az egész adaptáció alatti fitneszemelkedés igen nagy részéért, legalább a feléért felelős [Orr, 2002], és, hogy viszonylag kevés mutáció fixálódik a populációban egy átlagos adaptáció során [Gillespie, 2002]. A DNS alapú modelleket vizsgálva a kutatók azt is tapasztalták, hogy az adaptáció során, neutrális evolúcióhoz viszonyítva, jelentősen megnő a parallel mutációk száma [Orr, 2005a], azaz ismételt adaptációs utakban gyakrabban megtaláljuk ugyanazon szubsztitúciókat. Ez érthető is, hiszen az adaptáció, még ha több csúcsra is vezethet fel, a kiindulási állapot függvényében egyfajta irányultságot ad az evolúció menetének, a neutrális evolúcióhoz képest.

1.1.3. A geometriai modell és a DNS szekvencia alapú modellek összehasonlítása

Láthattuk, hogy Fisher geometriai modellje és a DNS alapú modellek egyaránt az adaptáció jelenségét magyarázzák, de teljesen különböző módon. Fisher modelljében a fenotípusok terében a populáció a környezetnek való jobb megfelelés irányába mozog. A DNS szekvencia alapú modellekben a populáció a DNS szekvenciák terében mozog, a környezethez való adaptáció a különböző szekvenciákhoz rendelt fitneszértékek által fejeződik ki. Fisher modellje folytonos, míg a DNS alapú modell az örökítőanyag sajátságai miatt diszkrét, véges. Ugyanakkor az adaptációról a modellek által levonható több fő konklúzió is egybeesik [Orr, 2005b]. Az adaptáció alatt fixálódott független mutációk fitneszhatásai mindkét modellben exponenciálisan csökkenek, tehát a fitneszemelkedése az egymás utáni adaptív mutációk által egyre kisebb mértékű lesz. Mindkét modell alapján azt várjuk

emellett, hogy ahogy közelítünk az optimumhoz, egyre ritkábban történjenek adaptív változások. A hasonlóságok mögött valószínűleg mindkét modell "matrjoska baba" jellege rejtőzik. Miszerint bármely adaptációs lépés után tulajdonképpen ugyanaz a folyamat ismétlődik meg, csak egy kisebb skálán. Ezek a hasonlóságok óvatos optimizmusra adhatnak okot, miszerint a két igen különböző modell alapján levont azonos, vagy nagyon hasonló következtetések robusztusak lehetnek a modellek fajtáira.

Ugyanakkor mindkét modell olyan feltevéseken alapul, amelyek a természetben nem feltétlenül teljesülnek, ezért predikcióikat ezek ismeretében kell kezelnünk. Mindkét modellben egy fix optimumhoz történik az adaptáció. Ezt elképzelhetjük úgy, hogy az adott vad típusú organizmus adaptálódott már a környezethez, de történik egy hirtelen környezeti változás, ami lehet egy inorganikus hirtelen hőmérséklet-változás, vagy akár egy új faj (kompetítor, vagy predátor, esetleg prédafaj) megjelenése a környezetben. A vad típus fenotípusa ekkor már nem optimális, és megkezdődik egy adaptáció, ami az új optimum eléréséig tart. Fix helyzetű optimummal találkozunk sok mikrobiális laborevolúciós kísérletnél, de gyakran egy folyamatosan változó optimumhoz történik az alkalmazkodás [Bell and Gonzalez, 2009, Burger and Lynch, 1995, Gomulkiewicz and Holt, 1995, Lynch et al., 1991, Orr and Unckless, 2008, Pease et al., 1989, Zhang and Hill, 2005]. Emellett mindkét modell abból indul ki, hogy van egy izogenikus vad típusú populáció, amely mutációk révén adaptálódik, melyek vagy rögtön kiszelektálódnak a populációból, vagy fixálódnak. A kezdeti genetikai variancia, a populáción belüli polimorfizmus nem kap szerepet a modellekben. Az ismertetett modellek egészen bizonyosan sok tekintetben nem realisztikusak, ezáltal alkalmazhatóságuknak, predikcióiknak korlátai vannak. Ugyanakkor általuk lehetőség nyílik a természetes szelekció által történő adaptáció kvantitatív vizsgálatára, mind elméleti síkon, mind kísérletek által.

1.2. A kompenzáló evolúció

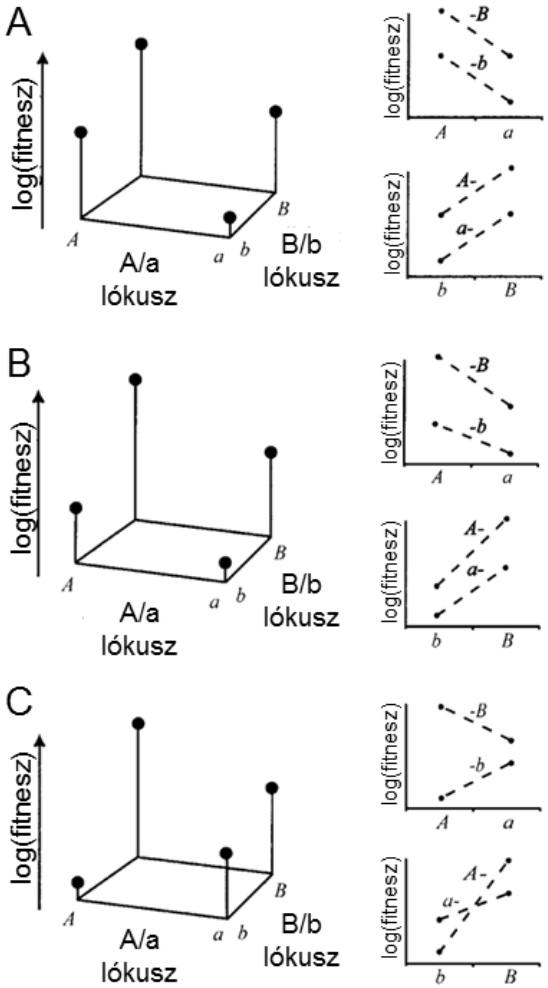
Kompenzáló evolúciónak nevezzük azt a jelenséget, amikor egy káros hatású mutáció a genom egyéb területein adaptív evolúciót indukál: megjelenhet egy az eredetitől független mutáció, amely az eredeti mutáció káros fitneszhatását részben, vagy egészben kijavítja, és pozitív szelekcióval elterjed a populációban. Tehát a kompenzáló evolúció létrejöttéhez szükség van egy káros mutációra, amely vagy a populációban összes egyedében, vagy legalábbis egy részükben jelen van. Természetesen a káros mutációk egy része megjelenése után a negatív szelekció áldozatául esik és nem fixálódik a populációban. Ideális esetben, konstans környezetben, egy végtelen méretű (vagy legalábbis "elég nagy") populációban, amelyben a mutációk ráadásul egyenként, egymástól időben izoláltan jelennek meg, valóban nem történhetne meg, hogy egy káros hatású mutáció felhalmozódjon a genomban. Valójában azonban egy káros fitneszhatású mutáció többféleképpen is elterjedhet

a populációban, melynek különböző módjait alább kifejtem, de példaként elég csak kis populációkban a genetikai sodródásra gondolnunk [Wright, 1970, Kimura, 1968, Kimura, 1983, Ohta, 1973].

Ha egy káros mutáció fixálódott a genomban, nem szelektálódott ki, akkor reverzióval visszamutálódhat, de ennek esélye igen kicsi, hiszen egyetlen meghatározott mutációnak kell ehhez megtörténnie. Egy alternatív lehetőség lehet a genom egyéb helyein véletlenszerűen megjelenő kompenzáló hatású mutációk elterjedése. Az ilyen kompenzáló mutációk specifikusak a kezdeti káros genetikai változásra. Ez a specificitás azt jelenti, hogy interakció, azaz episztázis van a mutációk között [Kimura, 1985, Weinreich and Chao, 2005]. Episztázisnak nevezzük, amikor egy mutáció fenotípusos hatása függ a genetikai háttértől [Dixon et al., 2009]. Két mutáció között akkor van episztázis, amikor együttes hatásuk különbözik attól, amit egyedi hatásaik alapján várnánk. Ezek lehetnek akár kis, kvantitatív eltérések is, ugyanakkor a kompenzáció legnyilvánvalóbb eseteiben a kompenzáló mutáció csak az eredeti káros mutációs háttéren előnyös, vad háttéren épphogy hátrányos. Ezt a speciális, nagyon erős episztázist nevezzük előjel episztázisnak (angolul "sign epistasis") [Weinreich et al., 2005], amelyben az egyik mutáció hatása a másik mutáció hatására előjelet vált (6. ábra).

A fitnesztájképen a vad típus egy adott csúcs, egy lokális (nem feltétlenül globális) optimum. Egy ilyen csúcs megmászása adaptív evolúcióval történik (5. ábra), ugyanakkor egy adott csúcsról más, akár "magasabb" csúcsok elérése bonyolultabb kérdés. Ehhez először egy mélyebben fekvő völgyön (tehát alacsonyabb rátermettséggel járó genotípuson) kell keresztülhaladnunk. Tulajdonképpen a kompenzáló evolúció során is ez történik, a csúcsot jelentő vad típusról a káros mutáció fixálódásával lejjebb kerülünk, ahonnan adaptációval elkezdődik egy újabb "hegymászás" a fitnesztájképen. Ez akár más csúcsokra is vezethet. Bár a természetes szelekció a csúcsról lefelé történő elmozdulás ellen dolgozik a kezdeti káros mutáció elleni szelekción keresztül, ha hatása nem túl erős, akkor mégis elképzelhető új csúcs meghódítása káros mutáció genetikai sodródás általi elterjedése révén. Sewall Wright sokat foglalkozott ezzel a kérdéssel és ő alkotta meg a "shifting balance" elméletet [Wright, 1932, Wright, 1970]. Azzal foglalkozott, hogy vajon egy adott lokális optimumról hogyan lehetséges más csúcsok elérése, és ezt a folyamatot milyen tényezők befolyásolják, segítik elő, vagy hátráltatják. A folyamat szerinte három részből áll [Wright, 1970]. Először genetikai sodródás által fixálódik egy káros mutáció egy kis populációban. (Ő ezt a folyamatot még mintavételi sodródásnak hívta és sodródás alatt sokféle folyamatot értett. Azóta a genetikai sodródás kifejezés az általa mintavételi sodródásként specifikált folyamatra lett használatos.) Ebben segíthet a populáció részleges izoláció által alpopulációkra történő darabolódása, mert ez a beltenyészettséget növeli, ezáltal a sodródást elősegíti. Második lépésben adaptáció által egy másik csúcsot hódít meg az adott alpopuláció, harmadik lépésben pedig, ha az új csúcs az eredetinél magasabb

fitneszt mutat, akkor az adott alpopuláció kiszorítja a faj többi alpopulációját, hiszen izolációjuk csak részleges, így közöttük génáramlás történik. Ezt a folyamatot a mutációs ráta, a természetes szelekció hatásfoka és a lokális beltenyészettség mind befolyásolja, azok egyensúlyára van szükség, ahogy írja is cikkében [Wright, 1932] : *"There must be gene mutation, but an excessive rate gives an array of freaks, not evolution ; there must be selection, but too severe a process destroys the field of variability, and thus the basis for further advance ; prevalence of local inbreeding within a species has extremely important evolutionary consequences, but too close inbreeding leads merely to extinction. ... life depends on the maintenance of a certain balance among its factors."*



6. ábra. **Előjel episztázis két mutáció között.**
Az ábrák bal oldalán a vízszintes síkon A és B gén (haploid) genotípusa van ábrázolva. Tegyük fel, hogy AB genotípus a vad típus, a és b pedig mutáns genotípusok. A függőleges tengelyen az adott genotípushoz tartozó fitnesz van ábrázolva (fitnesztájkép). Az ábrák jobb oldalán a három dimenziós tájkép van leképezve a síkra kétféleképp, hogy könnyebben érzékelhessük az A -> a (jobb felső alábrák), illetve a B -> b (jobb alsó alábrák) mutációk hatásait különböző genetikai háttéren. Az **A,** panelen a két mutáció között nincs episztázis. Az ab kettős mutáns olyan fitneszértéket mutat, amire az aB és Ab egyes mutánsok fitnesze alapján számítunk. Az episztázis hiánya a panel jobb alábráin úgy jelenik meg, hogy a mutációk hatását jelző szaggatott vonalak egymással párhuzamosak, tehát a genetikai háttér nincs hatással a mutáció hatáserősségére. A **B,** panelen egy episztázist láthatunk, ami nem előjel episztázis, mert a mutációk hatásának nem az előjelére, csak a mértékére van hatással a genetikai háttér. Ezt mutatja a jobb alábrákon, hogy a szaggatott vonalak nem párhuzamosak egymással. Az A -> a mutáció káros hatását csökkenti a b mutáns háttér, ugyanígy a B -> b mutáció hatását is csökkenti az a mutáns háttér. A **C,** alábra az előjel episztázist mutatja be. Vegyük észre a jobb felső alábrán, hogy az A -> a mutációnak nem csak a hatáserőssége, de a hatás előjele is változik a genetikai háttér függvényében. Míg b mutáns háttéren az A -> a mutáció előnyös, addig B vad háttéren ugyanaz a mutáció káros. Ábra módosítva [Weinreich et al., 2005] alapján.

Mennyire fontos fitneszvölgyek áthidalása az evolúció során? Ez a kérdés képezi alapját sok elméleti kutatásnak, amelyek populációgenetikai analízissel, RNS molekulák modellezésével, vagy digitális organizmusok evolúciójának szimulációjával próbálják megválaszolni, vajon a káros mutációk fixálódása és az azt követő kompenzáló evolúció milyen mértékben járul hozzá az adaptációs folyamathoz [Weissman et al., 2009, Gokhale et al., 2009, Lenski et al., 2003, Cowperthwaite et al., 2006]. Covert és munkatársai [Covert et al., 2013] is digitális organizmusokon vizsgálták a káros mutációknak az adaptációt segítő hatását. Ezek a digitális organizmusok olyan számítógépes programok, amelyek replikációra és mutációra képesek, evolválódnak és egy véges térben versengenek egymással az erőforrásokért. Szimulációik során azt találták, hogy a káros mutációk felgyorsítják az adaptációs rátát, de eredményeik függtek a mutációs ráta mértékétől. Nem meglepő módon, a rekombináció jelenléte csökkenti a káros mutációknak az adaptáció sebességét növelő hatását. Nyitott kérdés még a káros hatású mutációkat követő kompenzáló evolúció szerepe az adaptációban.

Ahogy láttuk, káros mutációk fixálódhatnak a populációban genetikai sodródás által. Ennek valószínűsége függ a mutáció fitneszhatásától és a populáció effektív méretétől [Kimura, 1983, Ohta, 1973]. Ugyanakkor nem feltétlenül kell a káros mutációnak fixálódnia a populációban ahhoz, hogy kompenzáció történhessen. Elég, ha kellő ideig fennmarad egy alpopuláció, amely tartalmazza a mutációt, hogy a polimorf populációban megjelenhessen a káros mutációt tartalmazó egyedek egyikében egy kompenzáló mutáció. Weinreich és munkatársai [Weinreich and Chao, 2005] elméleti modelljükkel megvizsgálták, hogy milyen feltételeknek kell teljesülnie ahhoz, hogy a káros mutáció fixálódása nélkül is tapasztalhassunk kompenzáló evolúciót. Azt találták, hogy míg kis populációknál a genetikai sodródás és a káros mutáció fixálódása áll elsősorban az új adaptív csúcsok elérésének hátterében, addig nagyobb populációknál már a fent leírt, fixációt nem igénylő folyamat fog dominálni. Modelljükben e két rezsimet egy kritikus populációméret választja el. Szimulálták a rekombináció hatását is, ami az utóbbi, fixációt nem igénylő folyamat hatásfokát nagyban rontja. Ugyanakkor eredményeik arra utalnak, hogy kapcsoltság esetén így is gyakran megtörtént egy új csúcs meghódítása a káros mutáció fixációja nélkül természetes populációkban [Weinreich and Chao, 2005].

Ha összehasonlítjuk egy adott faj különböző populációit, megfigyelhetjük a köztük mutatkozó különbségeket. Ennek a polimorfizmusnak egy része bizonyára neutrális vagy előnyös, de egy másik részük negatív fitneszhatású lehet. Ha elfogadjuk Wright koncepcióját, ezek a még nem fixálódott káros polimorfizmusok is alkothatják az alapját a kompenzáló evolúciónak. De vajon a fajokban megfigyelt polimorfizmusok milyen arányban mutatnak káros hatásokat? Az ezt vizsgáló kutatások meglepően sok káros variánst találtak. Becslések szerint *Saccharomyces cerevisiae*-ben 36% [Doniger et al., 2008], *Drosophila melanogaster*-ben 27% [Fay et al., 2002], emberben 20-38% [Fay et al., 2001, Sunyaev et al.,

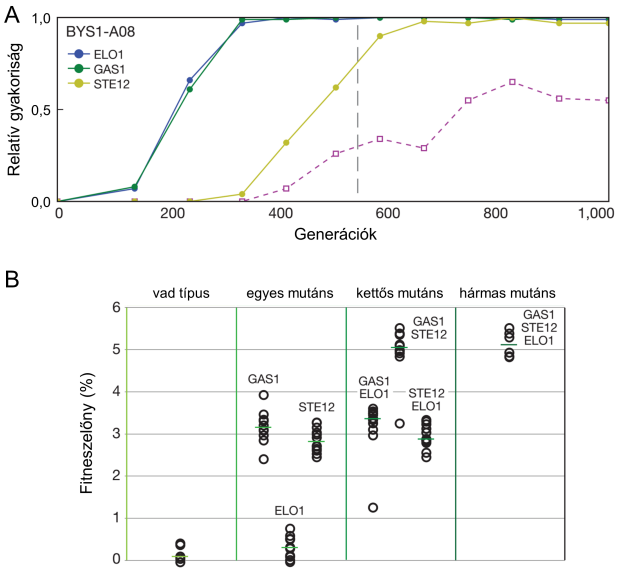
2001, Ng and Henikoff, 2002, Chasman and Adams, 2001] a káros polimorfizmusok aránya a kódoló régióban lévő nem szinonim pontmutációk között. McArthur és munkatársai [MacArthur et al., 2012] 185 emberi genom polimorfizmusainak vizsgálata során arra jutottak, hogy átlagosan minden ember 20 olyan homozigóta polimorfizmust hordoz, amely egy gént teljesen inaktivál.

Ezek a káros mutációk nem csak a genetikai sodródás által a természetes szelekciót elkerülve terjedhettek el, illetve maradhattak fenn a populációban. Ha egy káros mutáció egy pozitív szegregáló mutációhoz kapcsolva jelenik meg a populációban, az indirekt irányító szelekció következtében azzal együtt elterjedhet. Ez a genetikai hitchhiking jelensége (másik nevén genetic draft). Negyven aszexuálisan szaporodó (nem rekombinálódó) élesztő populáció részletes megszekvenálásával egy kutatás kimutatta, hogy nagyon gyakoriak az ilyen esetek, amikor egy káros, vagy neutrális mutáció szerencsés genetikai háttéren keletkezik és ezáltal terjed el a populációban, akár fixálódásig, akár csak magasabb arányig, mint amit fitneszhatása alapján várhatnánk [Lang et al., 2013]. Egy kiragadott példán szemléltei a szekvenálási eredmények bemutatásával ezt a folyamatot a 7. ábra.

Egy másik mód, amellyel egy káros mutáció elterjedhet, ha nem minden tekintetben káros. Antagonisztikus pleiotrópiának azt a jelenséget nevezzük, amikor egy adott mutációnak több független hatása van a fenotípusra. Ezekből az egyik pozitív hatású, míg egy másik negatív. Ha a pozitív hatás nagyobb, mint a negatív, a mutáció elterjedhet a populációban, majd káros hatására történhet kompenzáció. Illetve az is elképzelhető, hogy a mutáció egy környezetben előnyös, így a természetes szelekció által elterjed a populációban, majd a környezet megváltozásával káros hatásúvá válik [Elena and Lenski, 2003, Blank et al., 2014, Qian et al., 2012]. Ennek szemléletes és részletesen kutatott példája baktériumokban az antibiotikum rezisztencia kialakulása, mikor antibiotikum jelenlétében a rezisztencia-mutáció elterjed a populációban, majd az antibiotikum kezelés után az addig előnyt biztosító mutáció a rátermettséget csökkentheti a gyógyszer hiányában [Andersson and Levin, 1999, Andersson and Hughes, 2010].

Összességében elmondhatjuk, hogy a természetben nagy számban találunk káros hatású szegregáló mutációkat, amelyek a kompenzáló evolúció alapját képezhetik. Azután, hogy egy káros mutáció megjelenik és bizonyos mértékig elterjed a populációban, a negatív hatás kompenzáló evolúció által kijavításra kerülhet.

Nem történik valódi kompenzáló evolúció ugyanakkor abban az esetben, amikor a kompenzáló mutáció megelőzi a nélküle káros hatású mutációt. Ez az úgynevezett preszupresszió például akkor történhet meg [Gray et al., 2010], ha a későbbi kompenzációt okozó változás, valamilyen más okból előnyös, és fixálódásával lehetőséget teremt a nélküle káros mutációnak a neutrális vagy majdnem neutrális elterjedésre. Ez az elterjedés azért valósulhat meg, mert a kompenzáló mutáció fixálódása miatt az addig káros mutációnak már nincs, vagy csak nagyon kicsi a káros hatása, ezért egy kisméretű populációban

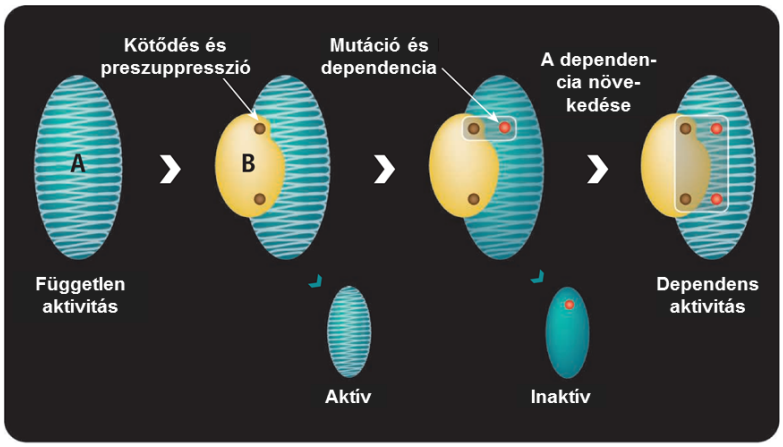


7. ábra. **Genetikai hitchhiking kísérletes feltárása.**
A, Különböző időpontokban a teljes populáció megszekvenálásával Lang és munkatársai követték egyes mutációk szegregációját egy élesztő populációban. **B**, A különböző mutánsok fitneszét az 545. generációban (szürke, szaggatott vonal) megbecsülve kiderül, hogy míg GAS1 és STE2 mutációk adaptívak, addig a neutrális ELO1 mutáció hozzájuk kapcsolva terjedt el a populációban. Ábra forrása: [Lang et al., 2013]

genetikai sodródás által elterjedhet. Gray és munkatársai állítása szerint [Gray et al., 2010] ez a folyamat gyakran előfordulhat molekuláris komplexek evolúciója során. Egy példájuk szerint képzeljük el, hogy A és B fehérje interakciója valamilyen szempontból előnyös, az A fehérjének azonban létezik egy ettől független funkciója is. A és B interakciója ezen funkció szempontjából lényegtelen, ugyanakkor az interakció megengedi A-ban olyan pontmutációk fixálódását, amelyek kizárólag B jelenlétében nem okozzák A funkcióvesztését. Ezáltal az addig A funkciója szempontjából semleges A-B interakció az A-ban fellépő pontmutációk kompenzátora lesz, az interakciótól függő lesz A funkciója (8. ábra). Gray gondolatmenete szerint a molekuláris szinten tapasztalt komplexitásért részben ez a neutrális folyamat is felelőssé tehető.

1.3. Kísérletes munkák a kompenzáció vizsgálatára

A kompenzáló evolúciónak, mint láthattuk, kiterjedt elméleti háttere van, de a kutatások nem rekedtek meg pusztán a modellek szintjén. Már a hetvenes években is vizsgálták kísérletesen a kompenzáló evolúciót [Campbell et al., 1973, Hall and Hartl, 1974], később pedig a molekuláris biológiai módszerek fejlődésével a kompenzáció gyakorlati, laborban történő kutatása még inkább lehetővé vált. Ezen kutatások egy része a kompenzáló

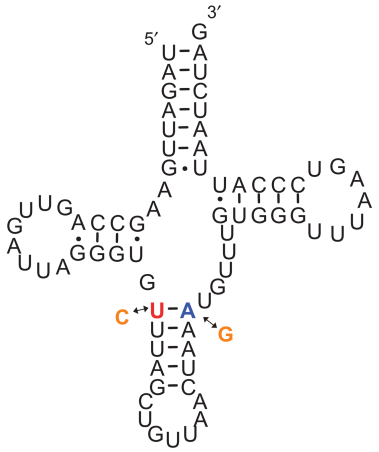


8. ábra. **A pre-szuppresszió folyamata.**
Egy molekuláris változás (A és B fehérje interakciója) szupresszálhatja A fehérje mutációjának negatív hatását, ezáltal facilitálhatja A fehérjében (közel) neutrális mutációk fixációját. Ábra forrása:[Gray et al., 2010]

evolúció lenyomatát próbálja felfedezni a mai élőlények összehasonlításával, mások pedig laborevolúciós kísérletet végeznek, melyeknél az evolúciós folyamatok egy rövid időskálán a tudósok szeme láttára zajlanak. Előbbi kutatási eredményből jóval kevesebb van, hiszen a mai élőlények közötti különbségek között kompenzációs kapcsolatot találni nagyon nehéz. Ehhez egyrészt szükség lenne két mutáció közötti episztázis kimutatására, másrészt a mutációk időrendiségét sem ismerjük általában. Ráadásul két mai faj közötti szubsztitúciók száma is magas, ami tovább nehezíti az effajta vizsgálódást. Éppen ezért fajok összehasonlításával kapott eredményeink csak néhány speciális esetben vannak.

Kompenzáció nyomát találták a génexpresszió regulációjában cisz és transz elemek között [Landry et al., 2005, Wittkopp et al., 2004, Goncalves et al., 2012], a nukleoszóma organizációjának [Kenigsberg et al., 2010], fehérje-fehérje interakciós partnerek [Pazos and Valencia, 2008], transzkripciós faktor kötőhelyek turn-overének [Doniger and Fay, 2007], illetve fehérjeszekvenciák [Kondrashov et al., 2002], vagy RNS molekulák másodlagos szerkezetének vizsgálatakor [Ivankov et al., 2014, Kern and Kondrashov, 2004, Meer et al., 2010, Kirby et al., 1995, Chen et al., 1999, Stephan and Kirby, 1993]. Az egyik ilyen kutatásban [Kern and Kondrashov, 2004] humán mitokondriális transzfer RNS-ekben találtak 52 olyan káros mutációt, amelyek más emlős fajokban szubsztitúcióként jelen vannak a természetes populációban. Feltételezték, hogy az utóbbi esetekben intramolekuláris kompenzáló szubsztitúciók által történhetett az emberben káros mutáció fixációja más fajokban. Az 52 patogén mutáció közül 32 egy Watson-Crick bázispárt tett tönkre. A kompenzációnak több lehetséges módját is kimutatták. Egy másik bázispár kötéserősségének növelése, egy új bázispár kötés kialakítása az RNS szerkezetben, vagy magasabb

rendű, bonyolultabb módosítások a transzfer RNS másodlagos szerkezetében is szerepet játszhattak abban, hogy növekedjen a molekula stabilitása. A 9. ábra mutat be egy példát a bázispár kijavítására csimpánz (*Pan troglodytes*) transzfer RNS-ben.



9. ábra. **Intramolekuláris kompenzáció emlős transzfer RNS-ben.**
A humán transzfer RNS-ben hátrányos C -> U polimorfizmus (pirossal jelölve) a csimpánz (*P. troglodytes*) transzfer RNS-ben szubsztitúcióként jelen van. Káros hatását ebben a fajban egy második szubsztitúció (G -> A, kékkel jelölve) kompenzálja. Ábra módosítva [Kern and Kondrashov, 2004] alapján.

Az eddig ismertetett munkák kurrens fajok összehasonlításával vizsgálják a kompenzáló evolúciót. Ahhoz, hogy a kompenzációt detektálni, vizsgálni lehessen, szükség van egy, a genotípusból a fenotípust, a fenotípusból fitneszt leképező, prediktív modell használatára. Számítógépes módszerekkel kell megjósolni, hogy például adott RNS szekvencia milyen másodlagos szerkezetet, adott fehérjeszekvenciák milyen fehérje-fehérje interakciót, illetve, transzkripciós faktor esetén, milyen DNS kötést eredményeznek. Ezeknek a predikcióknak a segítségével egy adott mutációhoz tulajdonképpen egy fitneszváltozást rendelünk, amely adatok segítségével találhatunk olyan mutációs partnereket, melyek egymással előjel episztázist mutatnak. Ennek egy alternatívája lehet a megfigyelt természetes variáció különböző kombinációinak mesterséges létrehozása és azok rátermettségének kísérletes tesztelése [Natarajan et al., 2013], de erre csak nagyon speciális esetekben és a mutációk limitált körében van lehetőség.

Ezzel szemben, laborevolúciós kísérletekben, bár nem a természetben valóban lezajlott evolúció végtermékét vizsgáljuk, nincs szükségünk sem a genotípus - fenotípus leképezésre valamely számítógépes modellel, sem a mutációk sorrendjének feltételezésére. Ezekben a kísérletekben a laborban mesterségesen hozunk létre káros mutációt, vagy mutációkat a populációban, majd ezután laborevolúció folyamán történik a kompenzáció, az adaptáció a megváltozott genetikai háttérhez. A mutációk sorrendisége ismert, hiszen mi hozzuk létre a kompenzáció kezdetét jelentő káros mutációt. A kiindulási vad típus, a

káros mutációt hordozó ős és a kompenzáció végén az evolvált törzs fitneszét is le tudjuk mérni kísérletesen. Amit az ilyen kísérletekben sem tudunk biztosan meghatározni, az az episztázis megléte. Ehhez azt kéne tudnunk, hogy a kompenzációt okozó mutáció előnyös hatása vad háttéren csökken, vagy megszűnik. Azaz, hogy a kompenzáció specifikus a genetikai háttérre, arra és nem a környezethez (tápoldathoz) történő adaptáció. Ennek eldöntésére specifikus kísérleteket lehet tervezni, de sok vizsgálatban az alapján valószínűsítik a kompenzáció meglétét, hogy vad típusú, káros mutációt el nem szenvedett kontroll törzsekkel is végzik a laborevolúciós kísérletet. Ha ezeknél az evolvált kontroll törzseknél nagyobb fitneszemelkedést mutatnak a mutáns populációk, akkor valószínűsíthetjük, hogy a káros mutációra specifikus kompenzáció történik ezekben [Moore et al., 2000]. Másik előnye a laborevolúciós kísérleteknek, hogy nem csak intragenikus kompenzációkat fedezhetünk fel velük, ellentétben a természetes szubsztitúciók vizsgálatával, ahol legmeggyőzőbb eredményeink RNS és fehérje szekvenciákon belüli kompenzációkról vannak. Ezzel szemben, laborevolúciós kísérletekben, bár nem a természetben valóban lezajlott evolúció végtermékét vizsgáljuk, nincs szükségünk sem a genotípus - fenotípus leképezésre valamely számítógépes modellel, sem a mutációk sorrendjének feltételezésére. Ezekben a kísérletekben a laborban mesterségesen hozunk létre káros mutációt, vagy mutációkat a populációban, majd ezután laborevolúció folyamán történik a kompenzáció, az adaptáció a megváltozott genetikai háttérhez. A mutációk sorrendisége ismert, hiszen mi hozzuk létre a kompenzáció kezdetét jelentő káros mutációt. A kiindulási vad típus, a káros mutációt hordozó ős és a kompenzáció végén az evolvált törzs fitneszét is le tudjuk mérni kísérletesen. Amit az ilyen kísérletekben sem tudunk biztosan meghatározni, az az episztázis megléte. Ehhez azt kéne tudnunk, hogy a kompenzációt okozó mutáció előnyös hatása vad háttéren csökken, vagy megszűnik. Azaz, hogy a kompenzáció specifikus a genetikai háttérre, arra és nem a környezethez (tápoldathoz) történő adaptáció. Ennek eldöntésére specifikus kísérleteket lehet tervezni, de sok vizsgálatban az alapján valószínűsítik a kompenzáció meglétét, hogy vad típusú, káros mutációt el nem szenvedett kontroll törzsekkel is végzik a laborevolúciós kísérletet. Ha ezeknél az evolvált kontroll törzseknél nagyobb fitneszemelkedést mutatnak a mutáns populációk, akkor valószínűsíthetjük, hogy a káros mutációra specifikus kompenzáció történik ezekben [Moore et al., 2000]. Másik előnye a laborevolúciós kísérleteknek, hogy nem csak intragenikus kompenzációkat fedezhetünk fel velük, ellentétben a természetes szubsztitúciók vizsgálatával, ahol legmeggyőzőbb eredményeink RNS és fehérje szekvenciákon belüli kompenzációkról vannak.

A laborevolúciós kompenzáció kezdetén egy vagy több káros mutációt kell létrehoznunk. Mivel ennek a mutáns törzsnek a fitneszét le tudjuk mérni, tesztelhetjük, hogy a kezdeti törzs fitneszértéke hogyan befolyásolja a kompenzáló evolúciót. Fisher geometriai modellje alapján azt várjuk, hogy a kisebb rátermettségű mutáns törzsek könnyebben kompenzálhatóak, mivel távolabb kerül a törzs a vad típust jelentő optimumtól, ezáltal

megnő az adaptív mutációk aránya. Érdekes kérdés az is, hogy a kezdeti mutáció által okozott funkcióvesztés vajon kijavításra kerül-e a kompenzáló evolúció által (funkció visszaállítása), vagy az adott funkcióra való igény csökken, azaz a funkcióvesztés negatív hatása mérséklődik (helyettesítés)?

A kompenzációt indukáló káros mutációt többféleképpen előidézhetjük. Vagy mesterségesen, irányított módon okozunk mutációt, vagy lemásoljuk a természetben meglévő folyamatokat. Az egyik módja, hogy egy rátermettséget csökkentő mutáció elterjedjen a populációban, a genetikai sodródás. Ennek egy extrém változatát kísérletesen az úgynevezett mutáció akkumulációs kísérletekben szokták használni. Ezekben a kísérletekben a populáció ismételten egyetlen (vagy nagyon kevés) egyedből (aszexuálisan szaporodó fajok esetén) álló "palacknyakon" vezetik keresztül. Azaz véletlenszerűen egyetlen egyedet vesznek a populációból és azt szaporítják tovább. Ezáltal a szelekció hatását szinte teljesen ki tudják iktatni, kivéve persze a letalitást okozó mutációkat, így káros mutációk halmozódnak fel a genomban [Muller, 1964]. Vírusban [Burch and Chao, 1999] és *Caenorhabditis elegans*-ban is [Estes and Lynch, 2003, Denver et al., 2010, Estes et al., 2011] vizsgálták mutáció akkumulációs kísérletekben a kompenzáló evolúciót. Utóbbi kísérletsorozatban vad típusú törzsből 74 független vonalban végeztek mutáció akkumulációt 240 generáción keresztül, majd az így keletkezett leromlott fitneszű törzseket adaptív laborevolúciónak tették ki. Azt találták az evolvált vonalak elemzésekor, hogy a nagyobb fitneszcsökkenést mutató ősök nagyobb fitneszemelkedést mutatnak. A fitneszemelkedés kezdetben gyors volt, majd lelassult az adaptáció során [Estes and Lynch, 2003], amely megfelel az elméleti modellek által jósoltaknak. A törzsek megszekvenálásakor nem találtak reverziót, csak kompenzáló mutációkat [Denver et al., 2010].

Ahogy az előző alfejezetben említésre került, a kezdeti káros mutáció fixálódásának egy másik lehetséges útja az antagonisztikus pleiotrópia jelensége. Erre egy természetes példa, amely a klinikumban is kiemelt jelentőségű, az antibiotikum rezisztencia kialakulása. Egy antibiotikum megjelenése a környezetben adaptációt indukál, rezisztencia terjedhet el a baktériumpopulációban. A rezisztenciát kialakító mutáció az antibiotikum környezetben való csökkenésével ugyanakkor káros hatású lehet. Tehát pozitív szelekcióval fixálódik a populációban egy mutáció (antibiotikum rezisztencia), amely a környezet megváltozásával (antibiotikum mentes környezet) negatív fitneszhatású lesz. A kompenzáció ilyenkor a rezisztencia-mutáció káros hatását csökkenti, természetesen csak akkor, ha nem történik reverzió [Andersson and Hughes, 2010]. Az antibiotikum rezisztencia kompenzációját vizsgáló kutatásokból [Andersson and Levin, 1999] az alábbi konklúziók vonhatók le: Gyakran találni kompenzáló mutációkat, vagy ugyanabban a génben, amelyben a rezisztenciát kialakító mutáció történt [Björkman et al., 1998, Johanson et al., 1996], vagy másikban [Bjorkman et al., 1999, Schrag et al., 1997, Schrag and Perrot, 1996, Sherman et al., 1996]. A kompenzáció antibiotikum rezisztencia esetén leggyakrabban a funkció visszaállításán

keresztül ment végbe [Andersson and Hughes, 2010], nem a funkció helyettesítése által. Emellett nagyon ritkán figyelték meg a rezisztencia-mutáció reverzióját. Ennek egyik oka, hogy reverzió mutációból egyetlen egy történhet, a potenciális kompenzáló mutációk száma ennél jóval magasabb. Ráadásul, ha fixálódott már egy kompenzáló mutáció, akkor a reverzió már nem feltétlenül fog fitneszelőnyt biztosítani, antibiotikum jelenléte nélkül sem, a kompenzáló és a rezisztencia-mutációk esetleges előjel episztázisa miatt.

A kompenzáló evolúció kísérletes vizsgálatakor a fenti kutatások során evolúciós folyamat révén (genetikai sodródás, természetes szelekció) fixálódtak a populációban a káros mutációk. Ennek egy alternatívája, ha a kísérlet elején mesterségesen hozzuk létre a káros mutációt, majd laborevolúció során történhet a kompenzáció. Az alábbi kutatások ez utóbbi stratégiát alkalmazták. Moore és munkatársai [Moore et al., 2000] *Escherichia coli*-ban vizsgálták random inzerciók kompenzációját, Poon és munkatársai [Poon and Chao, 2005, Poon and Chao, 2006] pedig a ϕX174 bakteriofágok adaptációját vizsgálta nem szinonim kódoló pontmutációk hatására. A kezdeti mutáció reverzióját ezekben a kísérletekben is csak nagyon ritkán tapasztalták: Baktériumban egyáltalán nem találtak reverziót, a fágban is csak egyetlen egyszer, szemben 9 kompenzáló mutációval. Mindkét kutatás során azt tapasztalták, hogy a kezdeti káros mutáció fitneszhatása és a kompenzációjának valószínűsége és mértéke között erős pozitív összefüggés van.

A kezdetben létrehozott káros mutáció, pontmutáció vagy inzerciók helyett, egyes modellszervezetekben, például élesztőben, elérhető egy gén teljes kiütésével. A génkiütött mutánsokban, ellentétben a mesterségesen létrehozott pontmutációkkal, vagy a mutáció akkumulációs vonalakkal, egy molekuláris sejtfunkció hiányát hozzuk létre. A kikapcsolt géntől függően ez a funkció akár a sejt fennmaradása, növekedése szempontjából alapvető fontosságú is lehet. Bár esszenciális gének elvesztése esetén nem tudunk laborevolúciót végrehajtani (hiszen a mutáns sejtek nem osztódnak), de bármely olyan mutáns, amely képes az osztódásra, bármilyen lassan is, alkalmas arra, hogy a kompenzáló evolúciót teszteljük rajta. A génspecifikusan bevitt pontmutációkhoz képest tehát a funkció teljes vesztése történik itt, a mutáció akkumulációs vonalakhoz képest pedig nagy előny, hogy pontosan ismerjük milyen funkció veszik el a mutáns törzsekben. Nem egyértelmű, hogy ezekben a géndeléciós törzsekben kompenzálható lesz-e a fehérje és ezzel együtt a funkció elvesztése. Egy pontmutáción átesett fehérje kissé megváltozott aminosav szekvenciája okozhat aktivitáscsökkenést, egy ilyen mutáció kompenzációját könnyű elképzelni. A szekvenciában bekövetkezett kompenzáló mutációk helyreállíthatják az aktivitást, vagy akár a fehérje expressziójának megnövelésével is csökkenthető a kezdeti mutáció káros hatása. Ezzel szemben egy gén teljes deléciójakor kérdéses, hogy történik-e egyáltalán kompenzáció, és ha igen, akkor milyen mechanizmusokon keresztül. Az eredeti funkció visszaállítása talán nehezebb, mint egyéb káros mutációk esetén, de géndeléció esetén is elképzelhető, amelyekre példákat mutatok a következő fejezetben. Természetesen az

eredeti funkcióra való igény csökkentésével (helyettesítés) is megtörténhet a kompenzáció. A kétféle kompenzáció megoszlását még nem vizsgálták.

Doktori munkám is a génvesztések kompenzációjával foglalkozik, melyek korábban csak kis skálájú kutatások témái voltak. Egy kutatássorozatban [Harcombe et al., 2009, Rokyta et al., 2002] DNS ligázt ütöttek ki egy bakteriofágban. Reverzió nem történt, kompenzáló mutációkat találtak, melyek egy része DNS metabolizmusban részt vevő génben történt, de nem mind volt funkcionálisan kapcsolható a DNS ligázhoz. A funkció teljes helyreállítását találta ugyanakkor két kutatás is [Campbell et al., 1973, Hall and Hartl, 1974], amelyekben a laktáz gént ütötték ki *Escherichia coli*-ban. Mindkét esetben azt találták, hogy más gének úgy mutálódtak, hogy képessé váltak a laktóz lebontására. Heineman és munkatársai [Heineman et al., 2005] bakteriofágban a lízisért felelős lizozim gén elvesztésére végeztek adaptációt és ugyancsak egy másik gén vette át a funkciót. Élesztőben egy citokinézis motorfehérje génjét, a *MYO1* gént ütötték ki és 400 generáción keresztül laboratóriumban evolváltattak 45 független populációt, amelyből 29 mutatott kompenzáló evolúciót [Rancati et al., 2008]. Vizsgálataik során azt találták, hogy nem pontmutációk, hanem aneuploidián keresztül megvalósuló, sok génre kiterjedő expressziós változások okozták a mutáns törzsek kompenzációját. Az eredmények amellett, hogy egyetlen gén elvesztését vizsgálták csak, azért sem tekinthetőek általánosnak, mivel a mitózisban részt vevő, ahhoz szükséges citokinézis motorfehérje elvesztése aneuploidiát és/vagy poliploidiát indukál.

Összefoglalva tehát a kompenzáló evolúciót kísérletesen vizsgáló eddigi kutatásokat, a mai fajok összevetésével széles körben találjuk nyomát valószínűsíthető kompenzáló evolúciónak, génexpresszió reguláció, nukleoszóma organizáció, fehérje és RNS szerkezet szintjén. Laborevolúciós kísérletes vizsgálatokból a következő fő konklúziók vonhatók le a kompenzáló evolúcióról:

1. A revertáló mutációk nagyon ritkák, ami érthető is, hiszen reverzió egyetlen mutációval hozható létre, míg a potenciálisan kompenzáló mutációk száma nagy.

2. Minél nagyobb mértékű a káros mutáció által okozott fitneszcsökkenés, annál nagyobb valószínűséggel és nagyobb hatású kompenzáló mutációk fixálódnak.

3. A rátermettség javulásában telítődés mutatkozik, az adaptív változások idővel egyre kisebb fitneszemelkedést okoznak.

4. Az eddigi laborkísérletekben a kompenzáló evolúció során fixálódott mutációk sokszor a káros mutációhoz funkcionálisan kapcsolhatóak, de gyakran olyan génekben is történik kompenzáló mutáció, amely funkcionálisan nem függ össze az eredeti káros mutációval.

A 2. és 3. megfigyelés egybevág Fisher geometriai adaptációs modelljével. Ha a populáció távolabb van az optimumtól (alacsony fitnesz), akkor a mutációk nagyobb hányada lesz

pozitív hatású, tehát gyorsabb, valószínűbb, erősebb kompenzációt várunk. Az idő előrehaladtával újabb és újabb kompenzáló mutációk elterjedésével a populáció közelebb kerül az optimumhoz, ezáltal egyre kisebb lépésekben növekedhet a fitnesz.

Az eddigi kutatások példákat mutatnak ugyan a kompenzáló evolúcióra, de nem rendszerszinten vizsgálódnak. Többségük nem teljes génvesztést, hanem kisebb hatású mutációkat vizsgált, azoknál is a molekulán belüli kompenzációt kutatták elsősorban. Kérdéses, hogy milyen mechanizmus alapján történhet egy teljes gén elvesztésének kompenzációja. Vajon helyreállítódik az eredeti funkció? Környezetspecifikus lesz a kompenzáció által biztosított fitneszjavulás, vagy általánosan érvényes, a környezetek széles körében megfigyelhető? Az elvesztett gén paralógjainak (ha vannak) kiemelt szerep jut-e a kompenzáló evolúció során? Vannak a kompenzációnak további evolúciós következményei? Az egyik ilyen következmény lehet, hogy a fixálódott káros mutációk és az általuk indukált kompenzáló evolúció a vad típus körüli genotípus térben való "sétákat" tesz lehetővé, ami által a populáció jobban feltérképezheti a genotípus teret, potenciálisan akár magasabbat fitneszcsúcsokat is meghódíthat (lásd. 1.1.2. fejezet és 10. ábra). Ezeket a kérdéseket doktori munkám előtt tudtommal még egy kutatás sem próbálta megválaszolni egy organizmusban sem rendszerszintű vizsgálatokkal. Az elméleti modellek, az adaptáció és a kompenzáció elméleti háttere és a felgyülemlett kísérletes eredmények is előkészítették a terepet doktori munkámhoz, a génvesztésre mutatott kompenzáló evolúció élesztőben történő, rendszerszintű vizsgálatához.

1.4. Géndeléció kompenzációjának vizsgálata sörélesztőben

Munkánk kezdetekor célunk a génvesztésre mutatott kompenzáló evolúció rendszerszintű laborevolúciós vizsgálata volt élesztőben. Az élesztő (*Saccharomyces cerevisiae*) ideális eukarióta modellorganizmus vizsgálatunkhoz. Nemcsak géndeléciós könyvtárak érhetőek el kísérletünk megkezdéséhez, a genetikai módosítások módszertana is rendkívül kidolgozott, emellett pedig rengeteg genomikai és funkcionális genomikai adat áll rendelkezésre. Így például ismert a legtöbb gén funkciója, kiütésének hatása az organizmusra, a gének expressziója és genetikai interakcióik, a fehérjék fizikai kölcsönhatásai és még sok egyéb, a kompenzáló evolúcióban potenciálisan szerepet játszó génváltozót tártak már fel élesztőben közel teljes genom lefedettségben. Haploid fázisban aszexuálisan szaporítható, rövid generációs idővel, ezáltal belátható időn belül várhatunk kompenzációt a laborevolúció során.

A kompenzációt kiváltó kezdeti káros mutációnak egy gén teljes elvesztését választottuk. Egy ilyen mutációnak jelentős fitneszhatása lehet. A gén funkciójáról és ezáltal a mutációnak az organizmusra kifejtett funkcionális hatásáról sok esetben van információnk, ellentétben például egy mutáció akkumulációs kísérlet véletlen mutációival, irányítottan

bevitt nem-szinonim kódoló mutációkkal, vagy regulációs mutációkkal, amelyek hatását kevésbe ismerjük. A sörélesztő géndeléciós könyvtárral már számtalan kutatás foglalkozott és a deléciós törzsek sok tulajdonsága ismert. Tudtommal rendszerszinten előttünk semmilyen kísérleti munka nem vizsgálta még a kompenzáló evolúciót.

2. Célkitűzés

Kísérletünk tervezésekor a célunk az volt, hogy a sörélesztőben minden olyan ismert gén elvesztésének kompenzációját megvizsgáljuk a laborban, amely gének deléciója fitneszhatást mutat ugyan, de a géndeléciós törzs még életképes, a törzs a kísérlet során életben tartható, osztódásra, ezáltal evolúcióra képes. A kisebb léptékű vizsgálatok eredményei alapján az alábbi célokat tűztük ki:

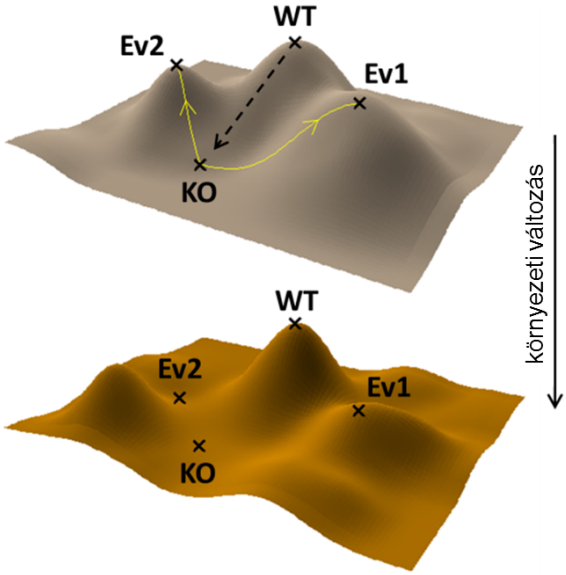
1. Először is arra voltunk kíváncsiak, milyen a kompenzációs potenciálja az élesztő genomnak. Meg akartuk határozni, hogy a gének mekkora hányadának elvesztését képes kompenzálni az organizmus.

2. Meg akartuk vizsgálni, hogy ez a kompenzáció mennyire teljes, azaz a kezdeti fitneszesést milyen mértékben kerül kijavításra.

3. Meg akartuk határozni, hogy mely gének elvesztését lehet kompenzálni. Azonosítani akartunk olyan genomikai változókat, amelyek elárulják, hogy egy adott gén elvesztése kompenzálható-e.

4. Célunk volt meghatározni a kompenzáció környezetfüggését. Ennek felderítésére a kompenzációt mutató vonalak és ősi, kiindulási törzseik fitneszét különböző környezetben lemértük.

5. Egy ehhez kapcsolódó kérdés, hogy a kompenzáció a kiütött gén funkciójának visszaállításán vagy helyettesítésén keresztül történik-e. Mint a bevezetőben láthattuk, géndeléciók esetén is találtak már példát mindkét esetre. Azt feltételezhetjük, hogy a funkció visszaállítása esetén a fitnesznövekedés, a kompenzáció általános jellegű lesz, nem függ például a külső környezettől. Ezzel szemben, ha funkcióhelyettesítés történt, akkor a kompenzáció megléte függhet a külső környezettől (10. ábra), hiszen a környezet megváltozásával megváltozhatnak a funkcionális szükségletek.



10. ábra. **Kompenzáló adaptáció bemutatása adaptív tájképen.**
Az adaptív evolúció a szürkével jelölt környezetben a génkiütött (KO) törzset különböző adaptív csúcsokra vezetheti. Hogy az evolvált vonalak (Ev1/Ev2) a vad típushoz (WT) képest mennyire divergáltak, ezt a környezet megváltozásával felderíthető. A vízszintes tengelyek az absztrakt genotípus teret mutatják, míg a függőleges tengely a fitneszt, amit meg tudunk becsülni a laborban.

3. Anyag és Módszer

A doktori munkám kísérleti és számítógépes módszereinek ismertetése előtt szeretném kiemelni, hogy a doktorim alapját képező kutatómunka Dr. Szamecz Bélával közös erőfeszítésünk volt. A labormunkát Dr. Szamecz Béla, az adatok kiértékelését, bioinformatikai analíziseket én, a kísérletek megtervezését és optimalizálását közösen végeztük.

3.1. Felhasznált élesztő törzsek

Munkánk során használt minden törzs a BY4741-es *Saccharomyces cerevisiae* törzsből származik. Egyszeres géndeléciós törzseket egy publikált gyűjteményből [Giaever et al., 2002] használtunk, hogy azonosíthassunk minden olyan gént, amelynek elvesztése nem letális, ugyanakkor szignifikáns fitneszcsökkenést okoz. Ezeket a lassan növő törzseket két lépésben azonosítottuk. Egy korábbi munka 671 gént talált, melyek homozigóta mutánsként diploid élesztő lassú növekedését okozták gazdag és szintetikus táptalajon is [Deutschbauer et al., 2005]. Ezekre a génekre nézve deléciós törzseket választottunk ki a MATa párosodási típusú haploid élesztő deléciós könyvtárból [Giaever et al., 2002] és mértük meg a fitneszét. 187 törzset találtunk, melyek szignifikáns és a vad típushoz képest legalább 10%-os fitneszcsökkenést mutattak (Mann-Whitney próba, fdr korrekció utáni p érték $< 0{,}05$). Ez a 10%-os érték jóval a mérésünk hibahatára fölött van. Ezen a 187 géndeléciós törzsen, mint ősi genotípuson kezdtük el a laborevolúciót.

3.2. Tápoldatok és táptalajok

A gazdag tápoldat (YPD) 1% élesztő kivonatot, 2% peptont, 2% glükózt tartalmazott. Ebben a tápoldatban zajlott a laborevolúció és, ha másképp nincs jelölve, a fitneszmérések is. A szilárd táptalaj 2% agar felhasználásával készült, amit korábbi mérésekkor optimálisnak találtunk ahhoz, hogy megismételhetőek legyenek a telepméret értékeink a kísérletek folyamán. A környezeti profilok meghatározásakor, a gazdag tápoldattól eltérő környezetekben történő mérésekkor a táptalajok a 1. táblázatban bemutatott anyagokat tartalmazták.

3.3. Evolúció a laboratóriumban

Mind a 187 lassú növekedést mutató géndeléciós törzzsel laborevolúciós kísérletet végeztünk. Minden ősi törzsből négy független vonallal indítottuk el a kísérletet. Emellett a vad típusú törzsünket is evolváltattuk 22 független evolváló vonalként (evolvált kontroll). Vad típusként a YOR202W géndeléciós törzset használtuk, mivel ennek fitnesze megkülönböztethetetlen a BY4741-es valódi vad típusú törzsétől [Qian et al., 2012], ugyanakkor

Környezet	Rövidítés	Szénforrás	Növekedés inhibitor
Glükóz	GLU	2% glükóz	-
Glicerin	GLY	3% glicerin	-
Tejsav	LAC	2% tejsav	-
Kálium-acetát	KAC	2% kálium-acetát	-
Etanol	ETH	3% etanol	-
Etanol + glicerin	ETHGLY	2% etanol + 2% glicerin	-
Raffinóz	RAF	2% raffinóz	1 μg/ml antimicin A
Koffein	CAF	2% glükóz	2 mg/ml koffein
Cikloheximid	CYC	2% glükóz	0,18 μg/ml cikloheximid
Ecetsav	ACE	2% glükóz	0,4% ecetsav
Glükóz + etanol	GLUETH	2% glükóz	6% etanol
Higromicin	HYG	2% glükóz	50 μg/ml higromicin
Nátrium-klorid	NACL	2% glükóz	1,2 M nátrium-klorid
Benomil	BEN	2% glükóz	15 μg/ml benomil

1. táblázat. **A gazdag táptalajtól eltérő mérésekben alkalmazott táptalajok összetétele.**

tartalmazza a kanamycin kazettát (KanMX4) a nem funkcionális his3Δ1 allél helyén. Ez azért fontos, mert így tudunk kontrollálni a korábban leírt mutagén hatására a KanMX4 kazettának [Teng et al., 2013]. Minden törzset 96-os lemezeken (96 db tápoldatot tartalmazó vályat egy műanyag lemezen) tartottunk. Minden lemezen 4 vályatot üresen hagytunk a keresztfertőzések vizsgálata céljából. A sejteket gazdag tápoldatban növesztettük, hogy a tápanyaglimitáltság általi szelekciós nyomást minimalizáljuk. Vizsgálatunkhoz fontos volt, hogy a génvesztés okozta fitneszcsökkenésen kívül minden más szelekciós erőt minimalizáljunk, hogy kompenzáló evolúciót figyelhessünk meg a laborevolúció végén, ne egyéb adaptációkat. A kanamycin kazetta jelenlétére nem szelektáltunk, nem adtunk kanamycint a tápoldathoz, hiszen a kanamycin egyrészt lecsökkenti növekedési rátát és ezért szelekciós nyomást jelentene rezisztencia kialakulására, másrészt mutagenezist indukálhat az élesztő stresszválaszán keresztül. Az optimális növekedési feltételek eléréséhez a lemezeket 350 fordulat/perc fordulatszámon rázattuk 30 °C-on. Egy kézi replikátor segítségével körülbelül 10^5 sejtet (~0,5 μl térfogatot) oltottunk át 100 μl friss tápoldatba 96-os lemezeken minden második nap. Ez a kísérleti felállás két átoltás között átlagosan 7,6 generációt jelentett. Az evolúciós kísérlet összesen 104 napig folyt, ami körülbelül összesen 400 generációt jelent. A 0., 26., 52., 78. és 104. napon a mintákat 15%-os glicerolban lefagyasztottuk és -80 °C-on tartottuk a növekedési ráta méréséig. Keresztfertőzések jelenlétére rendszeresen szűrtünk PCR vizsgálattal és az üres vályatok befertőződésének megfigyelésével.

3.4. Nagy skálájú fitneszmérések

A populációk növekedését optikai denzitásuk (OD) rendszeres mérésével követtük nyomon. Folyékony tápoldatban növesztettük a törzseket 384 vályatot tartalmazó lemezeken, ugyanolyan gazdag (YPD) tápoldatban, amiben az evolúció is történt. Relatív (kontroll vad típushoz viszonyított) növekedési rátát használtunk a fitnesz mérésére (lásd később). Ez a módszer lehetővé teszi nagy áteresztőképességgel lemérni egyszerre sok törzs növekedési rátáját (48 óra alatt ~7500 populáció növekedési rátáját tudjuk lemérni). Az evolúciós kísérletből vett mintákat tartalmazó, korábban lefagyasztott lemezeket először 96-os lemezekre oltottuk ki, ahol 48 órát nőttek a sejtek az evolúcióval megegyező körülmények között. Innen az uniformitás eléréséhez egy Microlab STARlet (Hamilton Robotics) robottal átoltottuk a sejteket 384 vályút tartalmazó lemezekre. Minden vályú 60 μl gazdag tápoldatot tartalmazott. Egy 384-es lemezre négy 96-os lemez tartalmát oltottuk rá. Három lemez ugyanannak a 96-os laborevolúcióban részt vevő lemeznek a különböző időpontú (0, 26, 52, 78, 104 nap) változata volt, a negyedik lemez mindig egy csak vad típusú kontroll törzseket tartalmazó lemez volt, melyet szisztematikus hatások kontrollálásához használtunk. A 384-es lemezeket 30 °C-on tartva automata inkubátorban tartottuk, 1000-1200 fordulat/perc fordulatszámú rázással és egy automatizált rendszer segítségével egy lemezolvasóban 20 perces időközönként mértük minden egyes lemezen minden egyes vályú optikai denzitását 600 nanométeres hullámhosszon. Minden evolúciós vonalat minden időpontban 6 technikai ismétlésben mértünk. Adataink közül kiszűrtük az abnormális növekedési görbékhez (pl. növekedés hiánya, bakteriális fertőződés gyanúja) tartozó értékeket, majd azokat a törzseket vizsgáltuk csak, ahol legalább 4 adatpont megmaradt a 6 technikai ismétlés közül.

A törzsek fitneszének mérésére a területen szokásos növekedési rátát alkalmaztuk [Warringer and Blomberg, 2003]. Dr. Kovács Károly által írt algoritmust használva a felvett növekedési görbék meredekségét minden mérési pontban kiszámítottuk, majd a 7 legnagyobb meredekség értékből (az első két legnagyobb értéket kiszűrve az esetleges kiugró értékek miatt) átlagot számoltunk. Így kaptuk meg egy adott technikai ismétlésre a növekedési rátát. A lemezeken belüli illetve lemezek közötti szisztematikus hibák kiszűrésére a lemezeken minden negyedik pozícióban egy vad típusú referencia törzset mértünk le. Minden mérési értékünket normáltuk az adott vályú szomszédságában lévő 4-9 db referencia vad típus növekedési rátájának mediánjára. Ezzel egyrészt egy relatív fitneszértéket kaptunk a vad típus fitneszére normálva, másrészről pedig kontrolláltuk mérésünket az esetleges szisztematikus (lemezen belüli és lemezek közötti) pozícióhatásokra. Egy adott törzs adott időpontban mutatott fitneszét az adott időpontban számolt vad típusra normált növekedési ráták mediánjaként becsültük meg. A 0 napos, ősi fitnesz számolásakor mind a négy evolúciós vonal 0 napos összes technikai replikátumának

növekedési rátájából együtt számoltunk mediánt, hiszen ekkor még izogenikus törzsek voltak.

3.5. Kompenzációt mutató gének azonosítása

Egy evolvált vonal akkor mutatott kompenzációt definíciónk szerint, ha két feltételt teljesített. Egyrészt, az őshöz képest szignifikáns fitneszemelkedést mutatott (egyoldalú Mann-Whitney próba fdr korrigált p érték $< 0{,}05$). Másrészt, hogy biztosak legyünk benne, hogy kompenzáció és nem a tenyésztési körülményekhez történő "tápanyagadaptáció" történt, az emelkedés mértéke nagyobb kellett, hogy legyen, mint azt a vad típusú evolvált kontroll törzstől várnánk. Ehhez normál elosztást illesztettünk a 22 vad típusú evolvált kontroll vonal fitnesznövekményére, és csak azokat az evolvált géndeléciós vonalakat tekintettük kompenzáltnak, amelyek fitnesznövekménye kisebb, mint 5% valószínűséggel származhatott volna az evolvált kontroll eloszlásból.

A kompenzáló evolúció mértékének meghatározásához bevezettük a relatív kompenzációs index (RKI) fogalmát,

$$RKI = \frac{\Delta_{végső} - \Delta_{kezdeti}}{kontroll_{végső} - \Delta_{kezdeti}} \tag{2}$$

ahol Δ a deléciós törzset, *kontroll* a vad típusú evolvált kontrollt, *végső* a 104 napos míg *kezdeti* a 0 napos állapotokat jelöli. Ez alapján a 100%-os relatív kompenzáció (RKI = 1) azt jelenti, hogy deléciós törzs az evolúció végére ugyanakkor fitneszt ért el, mint az evolvált kontroll törzs.

3.6. Környezeti profilok meghatározása

Hogy a kompenzáló evolúció során felhalmozott mutációk pleiotróp hatásait felderítsük (10. ábra), 237 kompenzációt mutató evolvált vonalat, a hozzá tartozó ősi genotípusokat, valamint mind a 22 evolvált kontroll vonalat fitneszmérésnek vetettük alá 14, a laborevolúció tápoldatától eltérő környezetben. A környezetek többségét egy korábbi kutatás alapján választottuk [Dudley et al., 2005], ahol a géndeléciós könyvtár fitneszét mérték meg különböző környezetekben (evolúció nélkül). A 14 környezet különböző szénforrásokat és növekedést gátló anyagokat tartalmazott (lásd 1. táblázat). Mivel nagyszámú mérést kellett végeznünk, illetve azért, hogy [Dudley et al., 2005] eredményeivel össze tudjuk vetni a saját ősi genotípusaink fitnesz adatait, ezeket a méréseket szilárd agar táptalajon végeztük. Hogy egységes, ismételhető átoltást és növekedésmérést érjünk el, egy robotikus replikáló rendszert állítottunk be telepméret alapján való fitneszmérésre. Téglalap alakú, egyenként 768 élesztő telepet tartalmazó agar lemezeken végeztük a mérést. Ezek mérete

megegyezik azokkal a 96 vályút tartalmazó lemezekkel, amelyekben a laborevolúció zajlott. Így egy agar lemezre összesen nyolc 96-os lemez oltható rá, melyet az alábbi módon végeztünk: három 96-os, evolúciós vonalakat tartalmazó lemezt oltottunk egy 768-as agar lemezre, mindegyiket a kísérlet kezdeti (0 napos) és végső (104 napos) időpontjából is. A hat 96-os lemez mellé oltottunk még két, csak referencia vad típust tartalmazó 96-os lemezt, amely a fitneszértékek normálására és az esetleges lemezen belüli és lemezek közötti szisztematikus hibákra való korrekcióra használtunk. A lemezeket 48 óráig tartottuk 30°C-on, majd azonos környezetbe való átoltás után és újabb 48 óra inkubáció (30°C) elteltével fényképet készítettünk róluk. A digitális képeket feldolgoztuk, telepméretet számoltunk a pixelek alapján Fekete Gergely által kidolgozott algoritmussal, a szisztematikus hibákat elimináltuk a vad típusú referencia mérések alapján. A lemezeken belüli pozícióhatásokra és lemezek közötti különbségekre illesztettünk egy modellt, ami felismeri, ha két lemez referencia vad törzseinek mérési értékei szisztematikusan különböznek, vagy ha egy lemezen belül egy adott irányba folytonosan nőnek vagy csökkenek az értékek. Előbbit okozhatja két mérési nap közötti környezeti különbségek, utóbbi oka lehet például egy esetlegesen az inkubátorban lévő hőmérséklet-gradiens. Az illesztett modell alapján korrigáltunk az adott hatásra. Minden törzs fitneszének adott időpontban, adott környezetben a technikai ismétlések (4 db) korrigált, vad típusú referenciára normált telepméreteinek mediánját vettük. Az ős és vad típus közötti, illetve az evolvált és vad típus közötti szignifikáns fitneszkülönbségeket egyoldalú Mann-Whitney próbával teszteltük (fdr korrigált p érték $<$ 0,05).

3.7. Genomszekvenálás

A kompenzáló evolúció mechanizmusának felderítése céljából genomszekvenálást végeztünk összesen 41 kiválasztott törzsön (lásd Függelék 10.1., 2. táblázat). Azokat az evolvált vonalakat választottuk ki elsősorban, amelyek őse jelentős fitneszesést mutatott a gén deléciójakor, majd a kompenzáló evolúció által nagymértékű javulás történt a törzs növekedési rátájában, és ez a javulás lehetőleg több lépésben ment végbe (a törzsek fitneszét három köztes időpontban is lemértük). Ez utóbbi feltétel azt jelezheti előre, hogy több kompenzáló mutáció is történt az adott evolválódó vonalban. Ahhoz, hogy a parallel evolúciót vizsgálni tudjuk, azonos kezdeti géndeléciós törzshöz több (2-4 db) egymástól függetlenül evolválódó vonalat is kiválasztottunk szekvenálásra. Összesen 41 evolúciós vonal és hozzájuk tartozó 14 db géndeléciós ős genomját szekvenáltunk meg. Izogenikus, tiszta tenyészetet készítettünk belőlük és a kompenzáció alatti fitneszemelkedést új, független méréssel validáltuk.

A szekvenálást Illumina HighScan SC készülékkel végezte az UD-GenoMed Medical Genomic Technologies Kft., átlagosan 806-os lefedettséget elérve a teljes genomon (a

törzsek szekvenálás előtti preparációjának részleteit lásd [Szamecz et al., 2014]). A megszekvenált DNS szakaszok a *Saccharomyces cerevisiae* EF4 genomhoz illesztettük a BWA szoftvercsomaggal [Li and Durbin, 2010] az ismétlődő szekvenciákat a RepeatMasking programmal távolítottuk el [Kohany et al., 2006]. A referenciától különböző variánsokat a GATK szoftvercsomaggal vizsgáltuk [McKenna et al., 2010]. Azon pontmutációkat figyelmen kívül hagytuk, amelyeknek "phred-scaled quality score" értéke 200 alatti volt, vagy a mutáns/referencia aránya 0,3 alatti volt. Egész kromoszómák, vagy kromoszómaszegmensek duplikációját a megnövekedett lefedettséggel azonosítottuk. Azokat az eltéréseket fogadtuk el duplikációnak, amelyek hossza legalább 25 kilobázis volt, és amelyek esetében mind a Control-FREEC [Boeva et al., 2012], mind a CNV-seq [Xie and Tammi, 2009] szoftverek szignifikáns eltérést jeleztek.

Célunk a *de novo* mutációk azonosítása volt. Ezek azok a mutációk, amik az ősben még nem voltak jelen, csak később, a kompenzáló evolúció során keletkeztek. Azokat a mutációkat, amelyek a génkiütés után jelentek meg a deléciós törzsben, de még a kísérlet megkezdése előtt, "másodlagos ősi mutációknak" nevezzük. Ezek a mutációk tehát a gén elvesztését elszenvedett ősben is jelen voltak már, nem csak egy adott evolvált vonalban, ugyanakkor semelyik másik ősi géndeléciós törzsben nem találtunk ilyen mutációt. Ezek a mutációk még lehetnek a génvesztésre specifikus kompenzáló mutációk (bár nem kísérletünk alatt keletkeztek). Ezzel szemben azok a mutációk, amelyek több ősi deléciós törzsben is megtalálhatóak, még a gén elvesztése előtt kellett, hogy létrejöjjenek. Ezek semmiképpen nem lehetnek kompenzáló mutációk, ezért figyelmen kívül hagytuk ezeket.

3.8. Pozitív szelekció azonosítása

A pozitív szelekció általános nyoma a kódoló régióban történő mutációkon belül a nem-szinonim változások feldúsulása. Bár a szinonim mutációk sem feltétlenül teljesen neutrálisak, a nem-szinonim mutációk jóval nagyobb hatásúak és az adaptív evolúció esetén, ahogy kompenzáló evolúciónál is, ezek gyakoriságának emelkedését várjuk. A teljes genomszekvenálás megmutatta, hogy a kódoló régióban található pontmutációk 86%-a nem-szinonim mutáció volt. A szelekció hatását a szinonim és nem szinonim mutációk aránya alapján, a neutrális esetre vonatkoztatva teszteltük. Ehhez Barrick és munkatársai módszerét használtuk [Barrick et al., 2009] használtuk. Röviden, a kódoló régióban talált pontmutációkra kiszámoltuk, hogy ha ezek a mutációk az összes kódoló régióban véletlen pozíciókban helyezkednének el, mekkora lenne az esélye annak, hogy legalább 86%-uk nem-szinonim mutációt okozna.

3.9. Transzkriptomikai vizsgálat

Transzkriptomikai vizsgálatra nyolc evolvált vonalat választottunk ki, melyek mindegyike jelentős fitneszesést mutatott génvesztéskor (min. 20%) és magas fitneszemelkedést mutatott a laborevolúció végére (min. 20%). Az evolvált vonalakhoz tartozó ősi törzsek és egy vad típusú kontroll expressziós profilját is felvettük. Az oligonukleotid alapú microarray méréseket holland kollaborátoraink, Patrick Kemmeren, Marian Groot Koerkamp és Frank Holstege végezték. A 10.1. függelék 2. táblázata tartalmazza a vizsgált törzseket. A kiválasztott törzsekből izogenikus, tiszta tenyészetet készítettünk és a kompenzáció alatti fitneszemelkedést új, független méréssel validáltuk. Azon gének expresszióját vettük megváltozottnak, melyek legalább 1,7-szeres változást mutattak 0,05 alatti p értékkel. Minden expressziós vizsgálatot és későbbi analízist elvégeztünk egy olyan adatbázison is, amelyből kiszűrtük a nagy skálájú duplikáción (CNV, teljes kromoszóma, vagy egy kromoszómaszegmens duplikációja) átesett géneket, illetve azokat a géneket, amelyek expressziójáról ismert, hogy függ az adott aneuploidiától [Torres et al., 2007], illetve a növekedési rátától [Airoldi et al., 2009, Brauer et al., 2008]. A transzkriptomikai adatok kiértékelését részben Dr. Kovács Károly végezte. A nyers adatok elérhetőek az alábbi url címen: http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-2352/. A törzsek preparációjának részleteit lásd: [Szamecz et al., 2014].

Hogy megvizsgáljuk a kompenzáló evolúció során a vad típusú génexpressziós állapot milyen mértékben állt vissza, minden egyes törzset expressziós profilja alapján egy hipertérben helyezhetünk el. Ez a tér annyi dimenzióból áll, ahány gén expresszióját mértük (több ezer). Ebben a térben vizsgáljuk meg minden egyes vonal esetén három törzs (vad típus, ős, evolvált) elhelyezkedését. Akárhány dimenziós is legyen a terünk, 3 pont benne mindig kijelöl egy síkot, ugyanúgy, ahogy 2 pont akár egy két dimenziós síkon, akár a 3 dimenziós térben mindig kijelöl egy egyenest. A 3 pont által kijelölt síkon egyszerűen ábrázolható a 3 pont távolsága. Ez a három távolság, a vad típus, az ős és az evolvált vonal euklideszi expressziós távolsága. A három pont ábrázolásával megtudhatjuk, hogy míg a génvesztés az ősi törzsünket eltávolította vad típusú expressziós állapotból, a kompenzáló evolúció vajon az evolvált vonalunkat a vad típusú állapot felé visszamozdítja-e, vagy esetleg távolítja.

3.10. Bioinformatikai analízis során használt adatsorok

Hogy megvizsgáljuk a kompenzáló evolúció mértékét befolyásolja-e a kiütött gén pleiotrópiája, három különböző élesztő génpleiotrópia adatsort használtunk. Környezeti pleiotrópiaként azon unikális környezetek számát tekintettük, amelyekben az adott gén elvesztése szignifikáns fitneszcsökkenést okozott [Dudley et al., 2005]. Egy gén hálózati pleiotrópiáját az adott génből átíródó fehérje BioGRID adatbázis [Stark, 2006] szerinti

fehérje-fehérje fizikális kapcsolatainak számaként kaptuk meg. A multifunkcionalitás mérőszám azt fejezi ki, hogy az adott gén hány olyan Gene Ontology kategóriához kapcsolható, amely kategóriát egy korábbi kutatásban élesztő genetikus funkcionálisan specifikusnak választottak [Myers et al., 2006].

Annak kivizsgálására, hogy a kompenzáló evolúció során történő mutációk a kiütött génnel funkcionális kapcsolatban álló génekben történtek-e, különböző mérőszámokat használtunk a funkcionális kapcsolat jellemzésére: stabil fehérje komplexekben lévő együttlét, genetikai interakciós profil hasonlóság, koexpresszió, paralógia. Fehérje komplexek definiálásához a Wodak csoport adatbázisát [Pu et al., 2009] használtuk, ami "tandem affinity purification" technikán és tömegspektrometrián alapszik (YHTP2008). A gének funkcionális kategóriákba való besorolását a MIPS Functional Catalogue Database-ből töltöttük le [Ruepp et al., 2004]. Genetikai interakciós profilhasonlóságot a Costanzo és munkatársai által lemért [Costanzo et al., 2010] nagy skálájú genetikai interakciós adatbázisból számoltunk. A profilt úgy kapjuk meg, hogy az adott génnek minden más génnel mutatott genetikai interakciós értékét egy vektorban megadjuk. A profilok hasonlóságát a két génhez tartozó profilok között számolt Pearson korrelációs koefficiens adja meg. Koexpressziós érték számolásához 247 normalizált microarray kísérlet eredményeit használtuk az M3D adatbázisból [Faith et al., 2008]. Ez alapján egy expressziós profilt állítottunk elő minden génhez. Ismételt mérések esetén a normalizált expressziós értékek átlagát számoltuk ki és használtuk fel a továbbiakban. Minden génpár esetén a koexpressziós értéket a két expressziós profil Pearson korrelációs koefficiense adta meg, hasonlóan a genetikai interakciós profilhasonlósági értékhez. Paralóg génpárokat BLASTP hasonlósági keresés alapján azonosítottunk. Minden nyitott leolvasási keretet (open reading frame, ORF) mindegyik másikkal szemben kerestük. Paralógnak tekintettünk két gént, ha a BLAST értékének várható értéke $E>10^{-8}$, az illesztés hossza legalább 100 aminosavhely hosszú volt, illetve ha a szekvencia hasonlóság legalább 30%-os volt és nem volt egyikük sem része transzpozonnak.

4. Eredmények

4.1. Az élesztő géneknek csak kis hányada alkalmas a kompenzáló evolúció laboratóriumi vizsgálatára

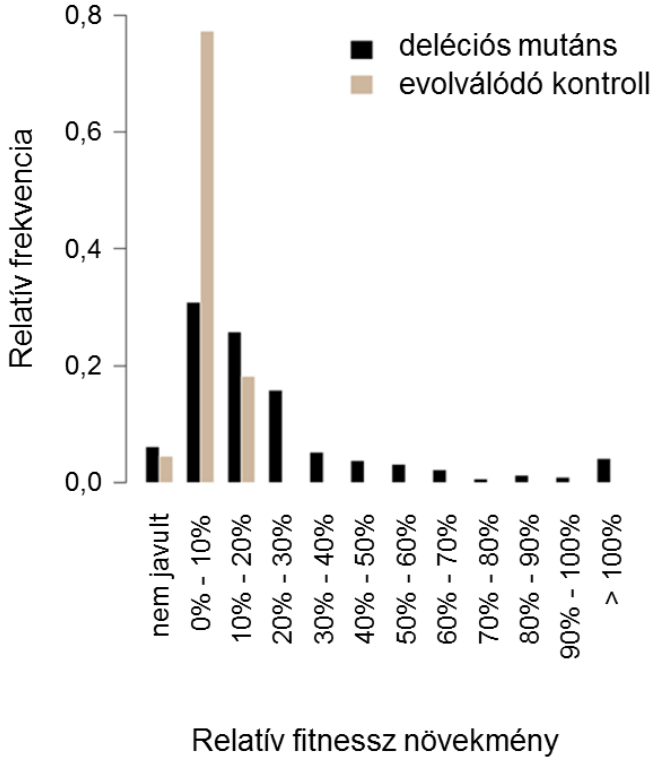
A kompenzáló evolúciót élesztőben vizsgáltuk haploid deléciós mutáns törzsek vizsgálatával. Ahhoz, hogy vizsgálatunkat el tudjuk kezdeni, legelőször is ki kellett választanunk a vizsgálandó gének csoportját. Az élesztő ~6500 génje közül ~1000 esszenciális gazdag tápoldatban [Giaever et al., 2002]. Esszenciális gének deléciójakor az organizmus nem életképes, nem mutat osztódást. Ezek a mutációk alkalmatlanok evolúciós kísérletekhez, hiszen nem vagyunk képesek egy osztódó populációt fenntartani a laborban, ezért evolúció sem mehet végbe. A maradék ~5500 nem esszenciális gén még mindig bőséges alapot szolgáltathat kísérletünkhöz. Ugyanakkor korábbi mérések kimutatták [Giaever et al., 2002, Costanzo et al., 2010, Breslow et al., 2008], hogy a nem esszenciális gének döntő többsége teljes inaktivációkor sem mutatott jelentős fitneszhatást gazdag tápoldatban. Egy ilyen génre deléciós törzs gazdag tápoldatban szinte a vad típusnak megfelelő fitneszt mutat. Ez azt jelenti, hogy egy ilyen törzsnek alig van mit kompenzálnia, hiszen a létrehozott mutációnak nincsen, vagy csak nagyon kicsi hatása van az organizmusra kísérletünkben. Összefoglalva tehát, kísérletünkben olyan kiinduló deléciós mutánsokra van szükség, melyek osztódásra és ezáltal evolúcióra képesek a laboratóriumban, ugyanakkor az adott gén deléciójának érzékelhető káros hatása van a törzs fitneszére, hogy a kompenzáló evolúciót kimutatható legyen a fitnesz javulásán keresztül.

Hogy megtaláljuk az összes ismert gént élesztőben, amely eleget tesz a fenti feltételeknek, egy többlépcsős kiválasztási stratégiát alkalmaztunk. Egy korábbi vizsgálatban lemérték homozigóta mutánsok fitneszét diploid élesztőben gazdag és minimális táptalajon, teljes genomskálán [Deutschbauer et al., 2005], és azonosítottak 671 nem esszenciális gént, melyek gazdag és minimális táptalajon is szignifikáns fitneszcsökkenést mutattak. Ezen törzsek fitneszét mi is lemértük. Ebből a kezdeti halmazból saját méréseink alapján végül 187 törzset azonosítottunk, amelyek legalább 10%-os szignifikáns fitneszcsökkenést mutattak. Ezzel a 187 törzzsel kezdtük meg a laboratóriumi evolúciót négy független vonalban.

4.2. A kompenzáló evolúció génvesztés után gyakori, és gyorsan bekövetkezik

Az evolvált vonalak és a hozzájuk tartozó ősi törzsek fitneszét összehasonlítva azt találtuk, hogy a géndeléciós törzsek többsége jelentős javulást mutatott (11. ábra). Összehasonlításképpen jelölve van az ábrán az evolvált vad típusú kontroll vonalak relatív fitneszemelkedése is. Míg a géndeléciós ősből induló evolvált vonalak átlagosan 23% fitneszemelkedést mutattak, addig ugyanez az érték a vad típusú kontroll vonalaknál csak 5% volt. Ez arra

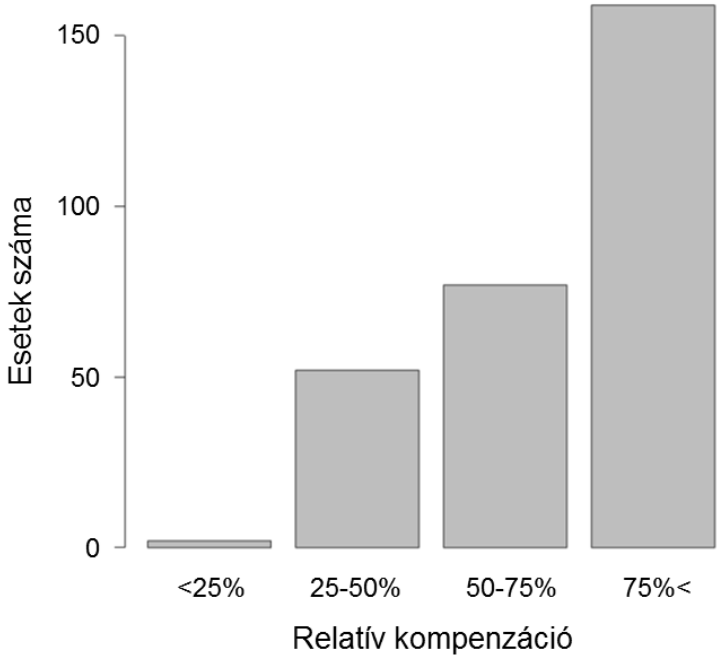
 utal, hogy a laboratóriumi evolúciós kísérletünk alatt a gén delécióra mutatott evolúciós válasz, a fitneszemelkedése nagyrészt kompenzáló evolúció eredménye és legfeljebb kisebb részben adaptáció a tápoldathoz. Ha alkalmazzuk a feltételt, miszerint a vad típusú kontrollnál nagyobb szignifikáns fitneszemelkedést kell mutatnia a kompenzáló vonalaknak, azt találjuk, hogy az általunk vizsgált géndeléciós genotípusok 68%-a (128 genotípus) mutatott kompenzációt legalább az egyik evolúciós vonalban.



11. ábra. **A kompenzáló evolúció jelentős fitneszemelkedést okozott.**
Az ábra a relatív fitneszemelkedés (evolvált fitnesz / ősi fitnesz - 1) gyakorisági eloszlását mutatja a gén deléciós ősből kiinduló evolvált vonalak (fekete oszlopok) és az evolvált vad típusú kontroll (szürke oszlopok) esetében.

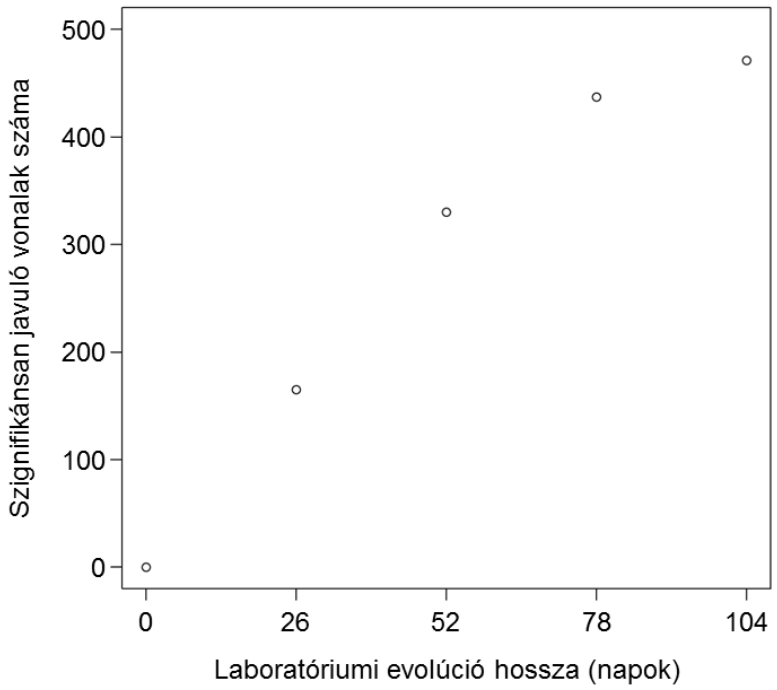
A kompenzáció mértékét más módon is vizsgálhatjuk. Megnézhetjük nem csak azt, hogy az őshöz képest hány százalékot nőtt az evolvált vonal fitnesze (relatív fitneszemelkedés), hanem azt is, hogy a kezdeti fitneszesés, amit az ős mutat a vad típushoz képest, a kompenzáló evolúció során milyen mértékben került kijavításra (relatív kompenzációs indexnek). A 12. ábra alapján a kompenzációt mutató evolvált vonalak döntő többsége a génvesztés okozta kezdeti fitneszcsökkenés legalább felét kompenzálta, de igen gyakori, hogy a vad típusú fitneszt is megközelítik. A relatív kompenzációs index számolásakor figyelembe vesszük az evolvált kontroll törzsek fitnesznövekedését, ezért az itt bemutatott magas kompenzációs értékek sem tulajdoníthatóak a tápoldathoz való adaptációnak.

Kimutattuk, hogy a kompenzáló evolúció a fitneszcsökkenést okozó, de nem esszenciális



12. ábra. **A génvesztés fitneszköltségének nagy része kompenzálható.**
Az ábra a relatív kompenzációs index eloszlását mutatja a kompenzációt mutató evolvált vonalak esetében.

gének nagy részében (68%) végbemegy, és a génvesztés okozta defektus nagy részben kompenzálható (12. ábra). A megfigyelt kompenzáció igen rövid idő alatt végbemegy, hiszen kísérletünk evolúciós időskálán szemlélve rövid volt, körülbelül 400 generáción (104 nap) keresztül zajlott. Felmerülhet a kérdés, hogy eredményeink mennyiben függnek az evolúciós idő hosszától. Laborevolúciós kísérletünk összesen 104 napon keresztül tartott. Mi történt volna, ha kétszer, vagy akár ezerszer ennyi idő áll rendelkezésre? A gének nagyobb részének elvesztését találtuk volna kompenzálhatónak és nagyobb mértékben? Tökéletes választ erre adatainkból nem tudunk adni, de méréseink megerősítenek minket abban, hogy elég hosszan végeztük a laborevolúciós kísérletet, az evolúció továbbvitele során sokkal kevesebb változást tapasztaltunk volna. Törzseink növekedési rátáját nem csak kísérletünk végén (104 nap után) mértük meg, de három köztes időpontban is, egyenletes időközökkel (26, 52 és 78 nap evolúció után). A kompenzáció mértékében telítődést tapasztaltunk (13. ábra). Ez arra utal, hogy a 104 napos kísérleti időtartam elegendő a kompenzálható géndeléciók döntő többségének felderítésére.

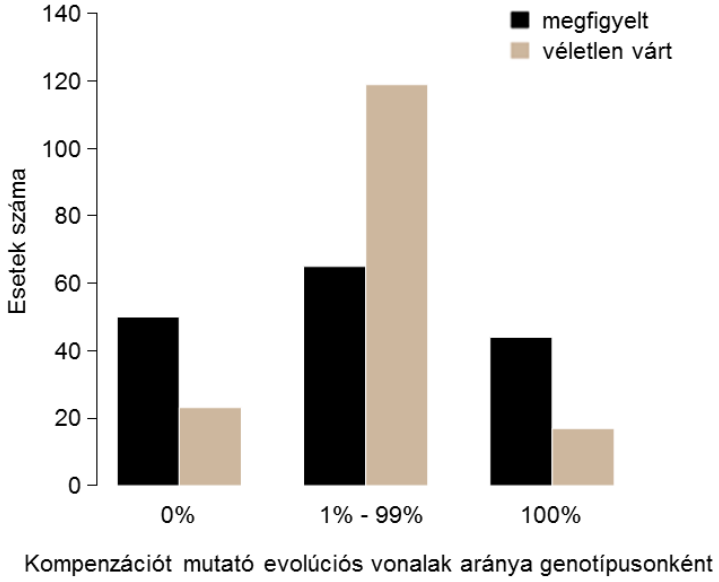


13. ábra. **A kompenzáció telítődést mutat kísérletünk során.**
A kompenzációt mutató vonalak száma kísérletünk során telítődést mutatott. A kísérlet kezdetén sok vonal mutatott kompenzációt, a kísérlet utolsó 26 napja alatt ugyanakkor az addig kompenzációt nem mutatott vonalak nagyon kis része volt képes kompenzáló evolúcióra.

4.3. Nem minden gén kompenzálható egyformán

Tehát megfigyeltük, hogy a gének egy nagy részének elvesztése kompenzálható nagy mértékben. Felmerül a kérdés, hogy vajon az élesztő gének mely tulajdonságai befolyásolják a kompenzálhatóságot? Mely gének elvesztését nem lehet evolúciósan kompenzálni és melyeket igen? Mielőtt a kérdést megvizsgálnánk és ismert géntulajdonságok és a kísérletünkben mutatott kompenzációs mérték közötti összefüggéseket keresnénk, feltehetjük a kérdést, hogy egyáltalán különbözik-e az általunk vizsgált 187 gén kompenzálhatóság tekintetében? Ha nem különböznének, akkor abban, hogy mely evolúciós vonalaknál tapasztaltunk kompenzációt és melyekben nem, csupán a véletlen döntött volna. Mivel minden egyes géndeléciós mutánsból négy független evolúciós vonalat indítottunk, ezért ezt a kérdést statisztikailag tesztelhetjük. Amennyiben véletlenszerűen kompenzálódott egyik vagy másik géndeléciós genotípus, úgy azt kell látnunk, hogy genotípusonként a kompenzációt mutató evolúciós vonalak száma (0-4) binomiális eloszlást mutat. Amennyiben a véletlen kompenzáció alapján vártnál gyakrabban találunk olyan genotípusokat, ahol az összes evolúciós vonal, illetve egyik vonal sem mutatott kompenzációt, ez arra enged

következtetni, hogy az eltérő gének inherensen különböznek azon tulajdonságukban, hogy elvesztésüket mennyire képes kompenzálni az organizmus. A kompenzálódást mutató vonalakat a genotípusok között véletlenszerűen megkeverve (randomizációs teszt) azt találtuk, hogy a véletlen várttól szignifikánsan nagyobb arányban vannak gének, amelyek teljes inaktivációját minden esetben, vagy egy esetben sem sikerült kompenzálni az evolúciós kísérletünk során (14. ábra, χ^2 próba, $P < 10^{-20}$). Tehát egyes gének elvesztését könnyebb, másokét nehezebb kompenzálni.



14. ábra. **Különböző gének elvesztése eltérő mértékben kompenzálható.** Minden géndeléciós ősből négy független vonalon indítottuk a laborevolúciót. A géneket csoportosítottuk aszerint, hogy a belőlük indított evolúciós vonalak egyike sem (0%)/ legalább egyike, de nem az összes (1% - 99%) / mindegyike (100%) mutatott-e kompenzációt. A kompenzálódást mutató vonalakat a genotípusok között véletlenszerűen megkeverve (randomizációs teszt) vizsgáltuk, hogy a különböző gének eltérnek-e kompenzálhatóságukban. 1000-szer randomizáltuk a kompenzációt mutató vonalakat a genotípusok között, ezáltal kaptunk egy eloszlást, amelyet akkor várnánk, ha a genotípusok nem különböznének kompenzálhatóságukban (szürke oszlopok). Az általunk megfigyelt valós eloszlás (fekete oszlopok) ettől szignifikánsan eltér.

4.4. A gének pleiotrópiája nem befolyásolja a kompenzációt

A következőkben azt kívántuk megvizsgálni, hogy mely tulajdonságai egy génnek azok, amelyek összefüggésben állnak a kompenzáció valószínűségével. Első hipotézisünk arra, hogy mely gének elvesztése kompenzálható gyakrabban, az adott gén pleiotrópiáján alapszik. A pleiotróp gének azok, melyek több funkció ellátásában vesznek részt. Egy ilyen génben bekövetkező mutációnak nem csak nagyobb lehet a fitneszhatása [Costanzo et al., 2010], de az inaktivációjával okozott kárt potenciálisan kompenzáló mutációk száma is nagyobb lehet, hiszen több funkcióban vehetnek részt, több fehérje állhat velük kapcsolatban, szélesebb körben fejtik ki hatásukat, mint nem pleiotróp társaik [Poon and Chao, 2005].

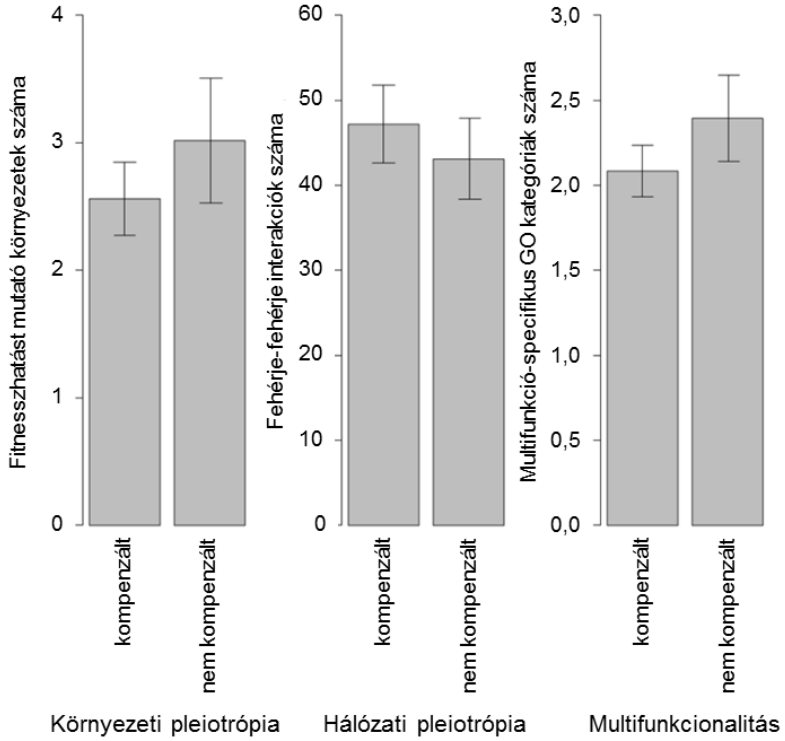
Hipotézisünk szerint a gének pleiotrópiája, úgymint 1) a génenkénti fehérje-fehérje fizikai interakciók száma, 2) a funkcionális kategóriák száma, amelyekbe az adott fehérjét besorolták 3) illetve a gén környezeti érzékenysége, azon környezetek száma, amelyben a gén deléciója fitneszcsökkenést okoz mind befolyásolhatják az adott gén elvesztésének kompenzálhatóságát. Azt találtuk, hogy nem térnek el szignifikánsan sem környezeti érzékenységben, sem fehérje-fehérje interakciós számban, sem a génhez rendelt funkcionális kategóriák számában a kompenzálódott és nem kompenzálható genotípusok (15. ábra).

4.5. A kompenzáló evolúció általában nem a paralóg génekre hat

A sörélesztő (*Saccharomyces cerevisiae*) evolúciós története során teljes genom duplikáción esett át [Wolfe and Shields, 1997], és természetesen kisebb, egy, vagy néhány gént érintő duplikációk is történtek őseiben. Feltételezhetnénk, hogy egy gén elvesztésekor a kompenzáló evolúció az adott gén duplikátumait érintheti, amely hasonlósága miatt átveheti a kiütött gén funkcióját. Az általunk vizsgált 187 génnek 27%-a (50 gén) rendelkezik paralóggal az élesztő genomban. A vizsgálatunkban kompenzációt mutató 128 genotípusnak pedig 25%-a (32 gén). Abban az esetben, ha a több kópiában meglévő gének elvesztését könnyebb lenne paralógjaikon keresztül kompenzálni, a paralóggal rendelkező gének mutánsainak gyakrabban kellene kompenzációt mutatniuk. Ha megvizsgáljuk, hogy a duplikált gének könnyebben kompenzálhatóak-e, azt kapjuk, hogy nincs különbség a két csoport között (χ^2 próba, p = 0,54). Mindemellett, a genomszekvenálási eredmények csak egyetlen genotípus esetében mutatták ki, hogy a kompenzáló evolúció az inaktivált gén duplikátumára hatott, részletekért lásd: [Szamecz et al., 2014].

4.6. A génkiütés fitneszköltsége korrelál a kompenzáció sebességével

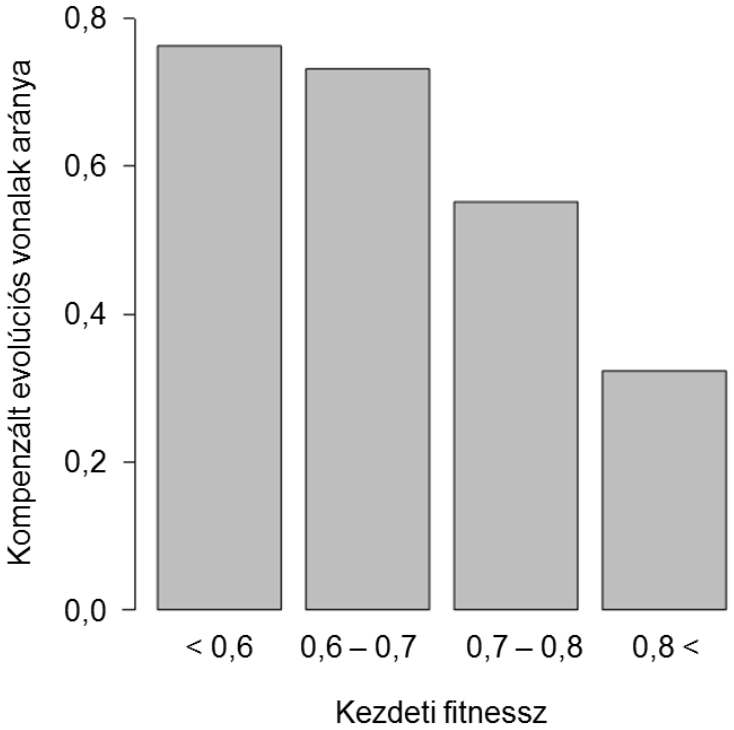
Az előbbiekben bemutattuk, hogy egy gén elvesztése után mutatott kompenzációt sem a gén pleiotrópiája, sem duplikátumainak jelenléte nem befolyásolja általánosan. Miért van tehát,



15. ábra. **A kompenzálhatóság nem függ az elvesztett gén pleiotrópiájától.** Az oszlopok az átlagot ± standard hibát mutatják. A különbség egyik esetben sem szignifikáns (Mann-Whitney próba, balról jobbra p = 0,71; 0,44; 0,36).

hogy egyes gének inaktivációja könnyebben, míg másoké nehezebben kompenzálható? Adatbázisunkban majdnem 200 genomikai és funkcionális genomikai géntulajdonság szerepel (pl. fehérje hossza, genetikai interakciós fokszám, fehérje komplexben való jelenlét, expresszió mérőszámok, különböző evolúciós ráták, pleiotrópia mérőszámok), amelyeknek megvizsgáltuk a kompenzációval mutatott összefüggését. A kompenzáció rátája azonban kizárólag egy változóval mutatott erős korrelációt, a génvesztés utáni fitnesszel. Minél alacsonyabb a génvesztés után mért fitnesz, azaz minél nagyobb a génkiütés hatása, annál inkább kompenzálhatónak találtuk a mutációt(16. ábra, χ^2 próba, $P < 10^{-13}$).

Törzseink rátermettségét csak egy adott hibával tudjuk megmérni. Emiatt a technikai korlát miatt előfordulhat, hogy kis fitneszemelkedést nem tudunk kimutatni. Ezáltal, a nagyon kis fitneszromlást mutató deléciós törzsek kompenzációját nem biztos, hogy képesek lennénk kimutatni. Ugyanakkor kizárhatjuk, hogy az általunk megfigyelt mintázat a mérés hibájának műterméke legyen, hiszen 1) csak a kezdetben legalább 10% fitneszromlást mutató törzseket vizsgáltuk. Ez a különbség a fitneszben jóval a detektálási küszöbünk felett van. 2) Az általunk talált összefüggés a detektálási küszöbünknél jóval nagyobb fitneszskálán áll fenn(lásd 16. ábra). Itt említeném meg, hogy a 4.3. fejezetben



16. ábra. **Az alacsony fitneszű géndeléciós genotípusok gyakrabban mutatnak kompenzációt.**
Minden evolúciós vonalat az ősben mutatott fitnesz szerint 4 csoportba osztottuk. Minden csoportban kiszámoltuk a kompenzált vonalak arányát. Az alacsony kezdeti fitneszt mutató vonalak kompenzálása gyakrabban következett be.

leírt mintázatot, miszerint a gének elvesztésük után mutatott kompenzációja egymástól különbözik, nem (csak) amiatt látjuk, hogy a gének különböznek egymástól deléciójukkor mutatott fitneszhatásukban (Függelék 10.2.).

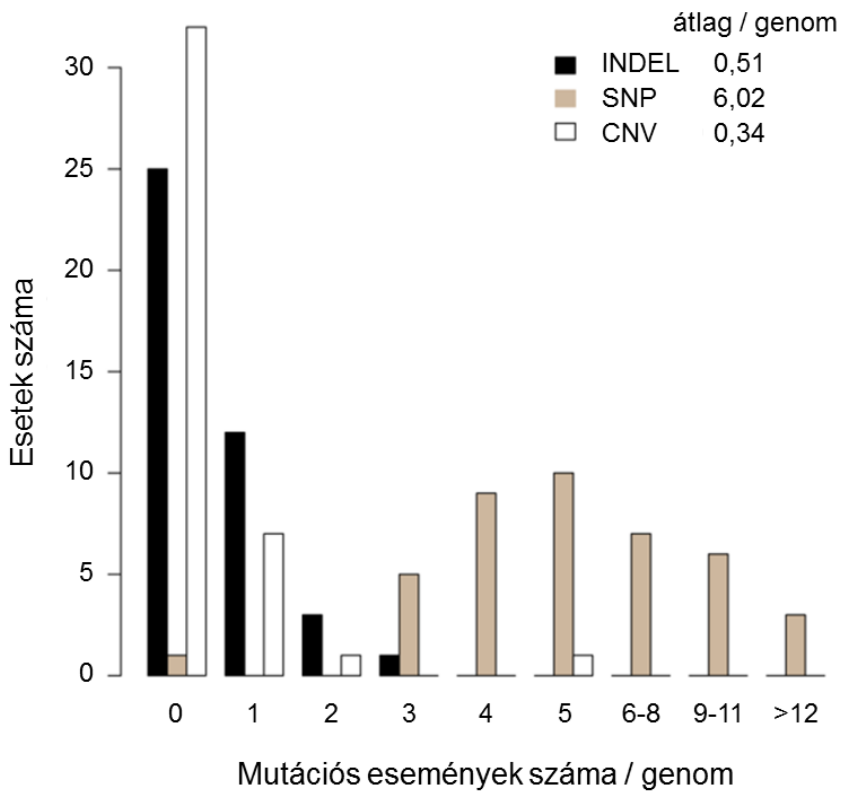
4.7. A kompenzáló evolúció elősegíti a genomok diverzifikációját

A kompenzáló evolúció mechanizmusának megvizsgálásához genomszekvenálást végeztünk egyes evolúciós vonalakon, amelyek mindegyike magas fitneszemelkedést mutatott kísérleti evolúciónk során. Egy adott géndeléciós genotípushoz több párhuzamos független vonalat is megszekvenáltunk, hogy észlelni tudjuk ha párhuzamos evolúció történt az azonos kiindulási genotípusú evolválódó vonalakban. Mindösszesen 41 evolúciós vonalon végeztünk szekvenálást, amelyek együtt 14 ősi genotípushoz tartoznak. A *de novo* mutációkra koncentráltunk, ezért is szekvenáltuk meg az ősi, kiindulási törzseket is, hogy el tudjuk különíteni a már korábban (a deléciós könyvtár készítésekor, vagy még korábban) felhalmozódott és ezért nem a delécióra specifikus mutációktól a potenciálisan kompenzációt létrehozó mutációktól.

Genomonként átlagban 6 pontmutációt (SNP), 0,5 inzerciót, vagy deléciót (INDEL) és 0,34 nagy skálájú duplikációt (CNV, teljes kromoszóma, vagy kromoszómaszegmens duplikációja) találtunk a vizsgált evolválódott vonalakban (17. ábra). Míg a vonalak majdnem 98%-ában találtunk legalább egy pontmutációt, addig kisebb inzerciót/deléciót csak 39%, nagy skálájú duplikációt pedig csak a vonalak 22%-a mutatott.

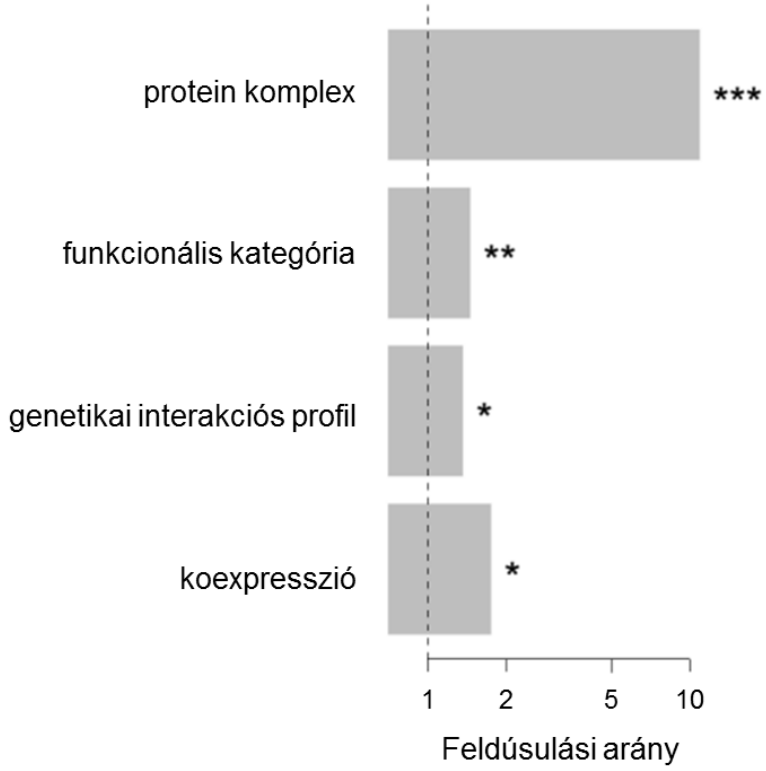
Ahogy a 17. ábra is mutatja, elsősorban pontmutációkat halmozódtak fel az evolválódott törzsekben. Kérdés ugyanakkor, hogy ezek a mutációk az adaptív kompenzáló evolúció során jöttek létre, így pozitív szelekció termékei-e, vagy neutrális, esetleg enyhén káros hatású mutációk. Ennek eldöntésére két tesztet végeztünk. A nem-szinonim mutációk feldúsulásának tesztelését Barrick és munkatársai [Barrick et al., 2009] módszere alapján végeztük és azt kaptuk, hogy a nem-szinonim mutációk szignifikánsan gyakoribbak a kódoló régiókat érintő pontmutációk között, mint azt várnánk tisztán neutrális mutációk esetén ($p = 0{,}003$).

Adaptív mutációk esetén azt is várjuk, hogy a felgyülemlett pontmutációk a véletlennél gyakrabban olyan fehérjéket kódoló génekben történjenek, amelyek valamilyen funkcionális kapcsolatban állnak azzal a génnel, amelynek elvesztésére a kompenzáló evolúció zajlik. Két gén összetartozását mutatja, ha funkcionális besorolásuk megegyezik, ha ugyanazon fehérje komplex tagjai, ha genetikai interakciós profiljuk hasonló, illetve ha koexpressziót mutatnak különböző környezetekben. Randomizációs teszt alapján az előbbi funkcionális kapcsolatokat a véletlen vártnál gyakrabban tapasztaljuk a deletált gén és az adott géndeléciós genotípusból kiinduló evolvált vonalban a kódoló régióban pontmutációt szenvedett gén között (18. ábra). A közös fehérje komplexben való jelenlét a véletlen várthoz képest több, mint tízszeres gyakoriságot mutat.



17. ábra. **A laborevolúció alatt felhalmozódott mutációk eloszlása genomonként.**
Az ábra a 41 genomszekvenálással megvizsgált evolúciós vonal *de novo* mutációt mutatja az ősi géndeléciós genotípushoz képest. A mutációkat három csoportba osztottuk: pontmutáció (SNP), kisebb inzerció/deléció (INDEL), illetve nagy skálájú (teljes kromoszóma, vagy kromoszómaszegmens) duplikáció (CNV). Az ábra jobb felső sarkában a genomonkénti átlagot is feltüntettük.

Tehát a laborevolúció alatt végbement pontmutációk, legalábbis részben, adaptív evolúció következményei. A mutáción átesett gének jelentős része funkcionális kapcsolatban áll az evolúció előtt inaktivált génnel. Vajon az azonos genotípusból kiinduló független evolúciós vonalakban azonos (vagy legalábbis azonos gént érintő) mutációkat, azaz parallel evolúciót milyen gyakran figyelhetünk meg? Azokban az esetekben, ahol genotípusonként több független evolúciós vonalat megszekvenáltunk, nagyon kis arányban találtuk nyomát parallel evolúciónak. Az azonos genotípusból indított független vonalakban felhalmozódott pontmutációk átlagosan csak 5,3%-ban fedtek át a gének szintjén (azaz történt mutáció ugyanazon gén szekvenciájában). Ha egymástól különböző genotípusokból indított evolvált vonalak között nézzük meg, még ritkábban, a mutációknak mindössze 0,1%-ában találunk átfedést.

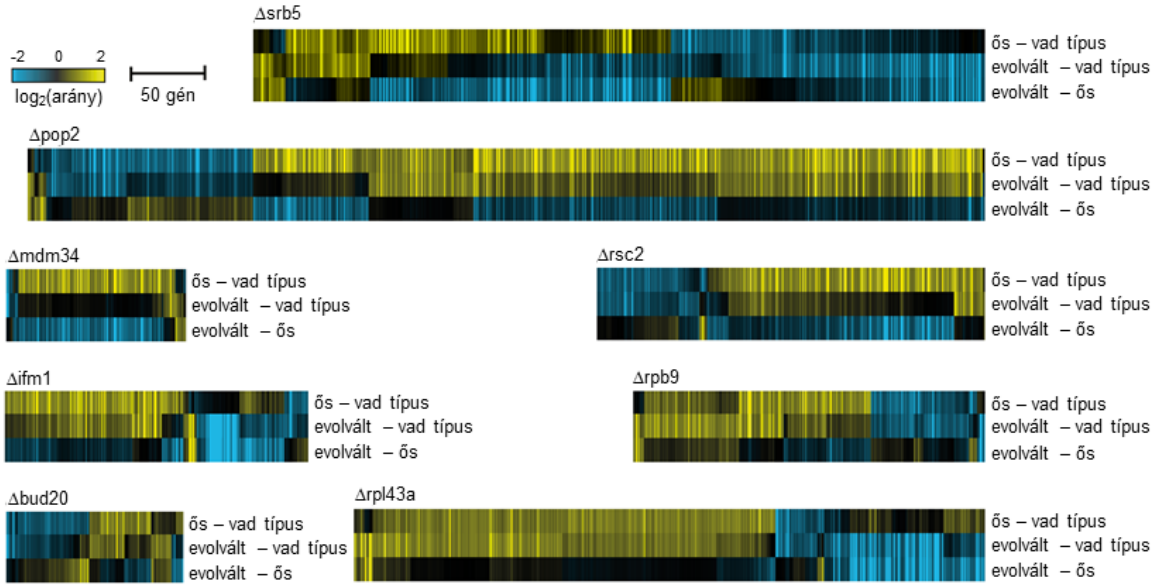


18. ábra. **A laborevolúció során mutálódott gén és az evolvált vonal ősében kiütött gén funkcionális kapcsolatban áll egymással.**
Az oszlopok a mutálódott gén és az ősben inaktivált gén között véletlenszerűen várt funkcionális átfedéséhez képest mutatják a megfigyelt átfedés feldúsulási arányát. A szaggatott vonal a véletlenszerűen várt átfedést mutatja. A csillagok a szignifikancia szintet jelzik: *: $P < 0{,}05$; **: $P < 0{,}01$; ***: $P < 0{,}001$. A feldúsulási arány tengely logaritmikus.

4.8. A kompenzáló evolúció nem állítja vissza a gének expresszióját a vad típus szintjére

A kompenzáció alatt bekövetkezett változások a fitnesz szintjén emelkedést okoztak, kisebb, vagy nagyobb mértékben, de a vad típushoz közelítették az evolvált vonal fitneszét a géndeléciós őshöz képest. Nem egyértelmű ugyanakkor, hogy ezt a kompenzációt a vad típusú fiziológia (teljes, vagy részleges) visszaállításával éri el a szervezet, vagy új fiziológiai állapotba kerül az organizmus és ezáltal javítja növekedési rátáját. Mindkét mód által végbemehet a kompenzáció. A két eltérő kompenzációs mechanizmus arányát kívántuk vizsgálni vonalaink expressziós profiljának meghatározásával.

Nyolc evolválódó vonalat választottunk ki transzkriptomikai analízisre. Ezen vonalak ősei nagy fitneszcsökkenést mutattak az adott gén inaktivációjakor és jelentős fitneszemelkedést a kompenzáló evolúció során. Minden egyes vonal genom-szintű génexpressziós profilját lemértük, emellett minden evolvált vonalhoz tartozó ős, és a vad típus expressziós

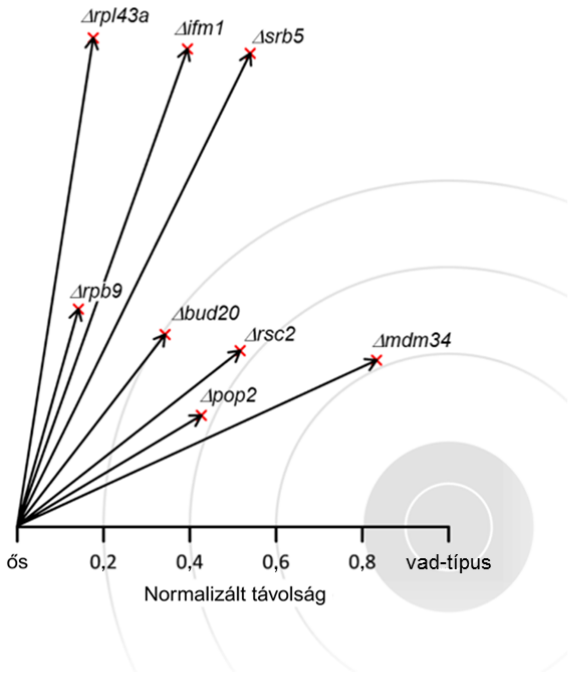


19. ábra. **Az evolvált vonalak, a kiinduló ősi genotípusok és a vad típus expressziós profiljának összehasonlítása.**
A sorok színskálával jelölik a nyolc, génexpressziós vizsgálaton átesett vonalak $\log_2$ expresszió arányát háromféle összehasonlításban: ős - vad típus; evolvált vonal - vad típus; evolvált vonal - ős. Minden egyes alábránál csak azon gének expressziója van bemutatva, amely a három összehasonlítás legalább egyikében szignifikáns eltérést mutatott (az expresszió relatív változása nagyobb, mint 1,7 és $p < 0{,}05$). A vad típusú expressziós állapot kompenzáló evolúció általi visszaállítása esetén azt kellene látnunk, hogy az "evolvált - vad típus" sorok túlnyomóan fekete színűek (nincs különbség a vad típus és az evolvált között), ugyanakkor az "ős - vad típus" és "evolvált - ős" sorok ellentétes expressziós változásokat mutatnak. Ilyen jellegű mintázatot egyedül a Δ*mdm34* törzsben figyelhetünk meg. Az ábrát Patrick Kemmeren és Marian J. A. Groot Koerkamp készítette.

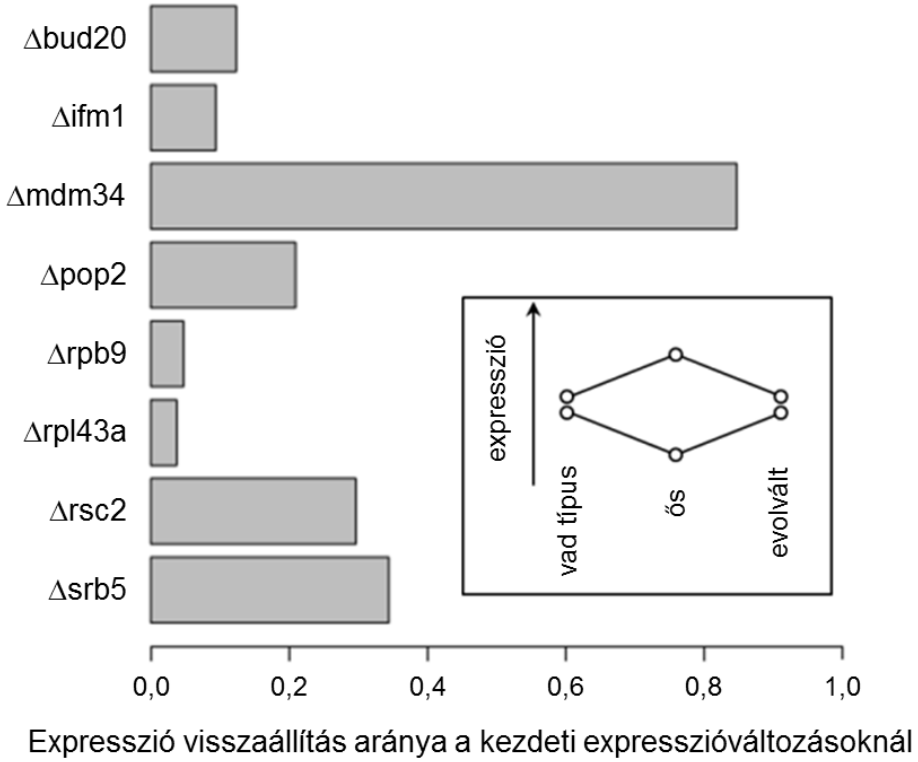
profilját is meghatároztuk. Így minden vonal esetén három expressziós profilt hasonlíthatunk össze (19. ábra). Az ősi genotípus és a vad típus összehasonlításából kiderül, hogy a génvesztés mely más gének expresszióját változtatta meg. Az evolvált és a vad típus, illetve az evolvált és az ős összehasonlításából pedig megtudhatjuk, hogy az evolvált vonalban az ősben talált változások közül hány eliminálódott. Azt találtuk, hogy a vonalak többségében a kompenzáló evolúció által okozott expressziós változások nem állították vissza a vad típusú szintet (19. ábra), egyedüli kivételt jelenthet a Δ*mdm34* törzs.

Ezt a jelenséget szemléletesebben be tudjuk mutatni, ha minden egyes törzset expressziós profilja alapján egy hipertérben helyezünk el és a vad típus, a géndeléciós ősi genotípus és az abból származó evolvált vonal expressziós profiljainak a hipertérben mért távolságait ábrázoljuk a síkon. Azt figyelhetjük meg, hogy a legtöbb evolvált vonal nem

közeledik lényegesen a vad típushoz (20. ábra). Ez alól csak a Δ*mdm34* vonal kivétel (a Δ*mdm34* különállása e tekintetben a 19. ábrán is megfigyelhető). Ha vizsgálatunkat csak az expressziójukat szignifikánsan megváltoztató génekre korlátozzuk, és vonalanként megnézzük, hogy az ősben a vad típushoz képest történt szignifikáns expresszió változások hány százaléka változott szignifikánsan az ellenkező irányba a kompenzáló evolúció során (ezt a jelenséget expresszió visszaállításnak nevezzük), azt tapasztaljuk, hogy a Δ*mdm34* törzs kivételével nagyon alacsony százaléka állítódik vissza a génvesztés okozta kezdeti expresszióváltozásoknak (21. ábra). Eredményeink változatlanok maradnak, ha analízisünkből eltávolítjuk azon géneket, amelyek kromoszomális régiója duplikáción esett át a kompenzáció alatt, azon géneket, amelyek aneuploid törzsekben expresszióváltozást mutatnak [Torres et al., 2007] illetve azon géneket, amelyek expressziós szintje a növekedési rátával együtt változik [Airoldi et al., 2009]. Ez utóbbi azért fontos, mert a kompenzációt mutató vonalakban a növekedési rátához kapcsolt expressziójú gének esetén a megváltozott expresszió helyreállítását tapasztalhatjuk akkor is, ha az expresszió visszaállítása nem oka, hanem okozata a megemelkedett fitnesznek.



20. ábra. **A kompenzáló evolúció során az expressziós állapot nem áll vissza a vad típusra.**
Minden evolvált vonal - géndeléciós ős - vad típus hármas esetén kiszámoljuk a három törzs euklideszi távolságát transzkriptomikai méréseik alapján. A három távolságot ábrázoljuk az ábrán úgy, hogy az ősi géndeléciós törzs és a vad típus közötti távolságra normálunk. A vad típus körüli szürke zóna és koncentrikus körök mutatják, hogy a kompenzáció alatt milyen mértékben állítódott vissza az expressziós állapot.

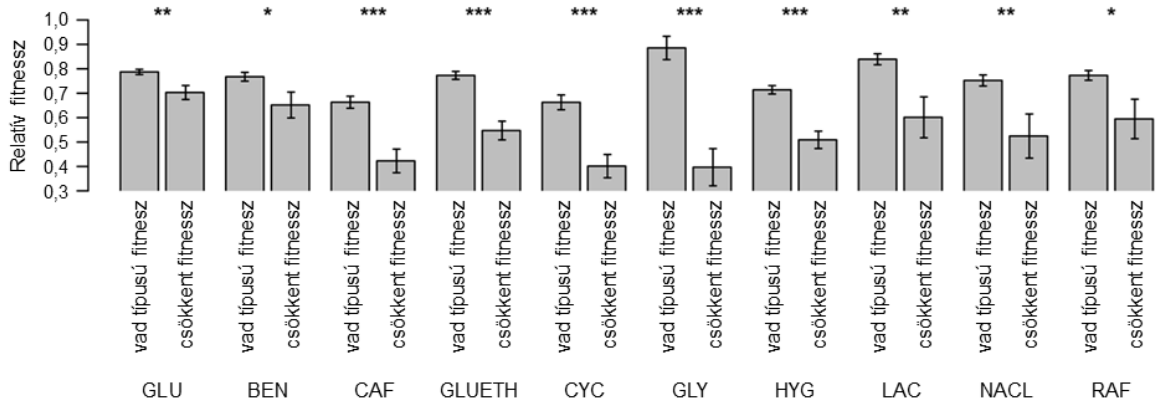


21. ábra. **A génvesztés hatására beálló szignifikáns génexpresszió változások nem állítódnak vissza a kompenzáció folyamán.**
Az ábra a transzkriptomikai vizsgálatnak alávetett evolúciós vonalakban mutatja, hogy a génvesztés által az ősben okozott szignifikáns expresszió változások milyen hányada változott szignifikánsan az ellenkező irányba a laborevolúció alatt. Az ábrát Dr. Kovács Károly készítette.

4.9. A kompenzáló evolúció diverzifikációs hatása változatos környezetekben felderíthető

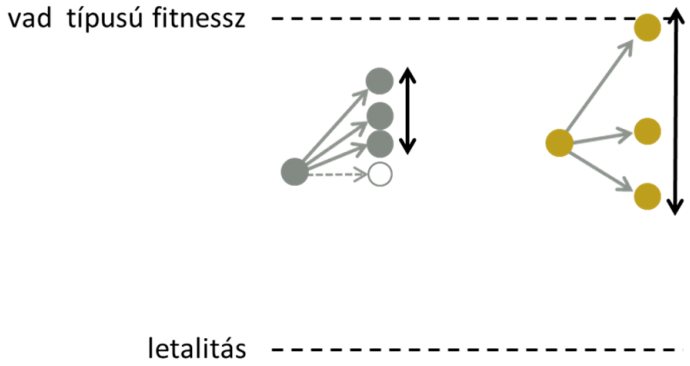
A szekvenálási adatokon láthattuk, hogy az egy ősből származó evolválódó vonalaink is divergenciát mutattak, nagyon alacsony volt a parallel evolúció aránya. Ugyanígy, divergenciát figyelhettünk meg a gének expressziója szintjén, a vad típusú expressziós állapot visszaállítását alig mutatták a kompenzáló törzseink. Felmerül a kérdés, hogy bár a kompenzáció alatt történő változások genotípus szintjén divergenciát okoznak, vajon a kísérletben evolválódott, a vad típustól különböző megoldások a fitnesz növelésére fenotípusosan egyenértékűek-e. Ennek a kérdésnek a megválaszolására meghatároztuk 237 kompenzációt mutató evolvált vonal és a hozzájuk tartozó ősi törzsek fitneszét tizennégy különböző környezetben. Módszerünk validálásaként először megvizsgáltuk az ősi géndeléciós vonalak által mutatott fitneszértékeket. Gének elvesztésének fitneszhatását (laborevolúció nélkül) már korábban vizsgálták Dudley és munkatársai [Dudley et al., 2005]. Mivel a kísérletben szereplő környezeteink egy részét is az előbbi publikáció alapján

választottuk ki, eredményeinket össze tudtuk vetni Dudley és munkatársai [Dudley et al., 2005] által kapott fitneszértékekkel és szignifikáns egyezést találtunk (22. ábra).



22. ábra. **A különböző környezetekben a géndeléciós ős által mutatott fitnesz egyezik korábbi irodalmi adatokkal.**
Az ősi géndeléciós törzseink fitnesz eredményeit összehasonlítottuk Dudley és munkatársai [Dudley et al., 2005] által mérttel. Minden genotípus minden környezetben kategorizálva volt Dudley és mtsai. munkájában: egyes génkiütések nem okoztak fitneszhatást az adott környezetben (vad típusú fitnesz), mások pedig csökkentették a fitneszt, vagy akár letalitást okoztak (csökkent fitnesz). E kategorizálás szerint összehasonlítottuk az általunk mért folytonos, a vad típusúra normált fitnesz adatokat, és szignifikáns eltérést találtunk a várt irányba minden átfedő környezetben (egyoldalú Mann Whitney próba, a csillagok a szignifikancia szintet jelzik: *: $P < 0{,}05$; **: $P < 0{,}01$; ***: $P < 0{,}001$.).

Kísérletünkben az eredeti tápoldatban (amelyben a szelekció is zajlott) kompenzációt mutatott evolvált vonalak fitneszértékeit határoztuk meg más környezetekben. Ugyanazon gén elvesztését két egymástól független módon kompenzáló evolvált vonal fenotípusos különbségét esetleg nem tapasztaljuk gazdag tápoldatban, de egyes környezetekben egymástól igen különböző növekedési rátát mutathatnak (23. és 10. ábra).

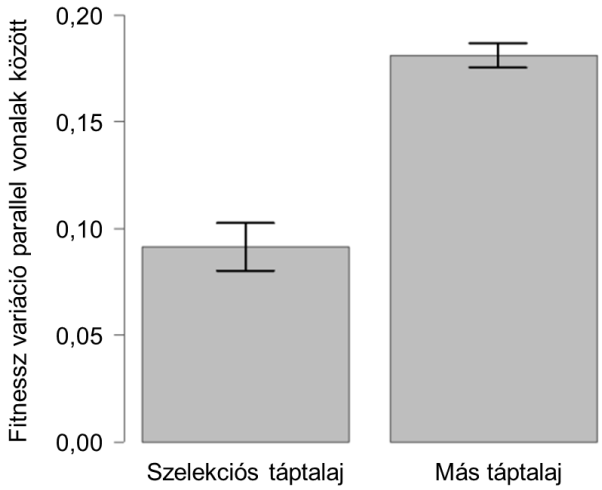


23. ábra. **A környezet megváltoztatásával a független evolúciós vonalak fenotípusos különbözőségei kimutathatók.**
Sematikus ábránk azt szemlélteti, hogy míg gazdag táptalajon (szürke szín), ahol az evolúció zajlott, a fitneszértékekben esetleg nem látunk nagy különbséget az egy ősből evolvált vonalaink között, addig a környezet megváltoztatásával (sárga szín) az evolvált vonalak addig rejtve maradt különbsége felszínre hozható. A bal oldali kör az őst, az onnan nyíllal leszármazó körök az evolvált vonalakat jelölik. A szaggatott nyíl és üres kör azt mutatja, hogy az eredeti gazdag táptalajon kompenzációt nem mutató vonalakat nem mértük más környezetekben.

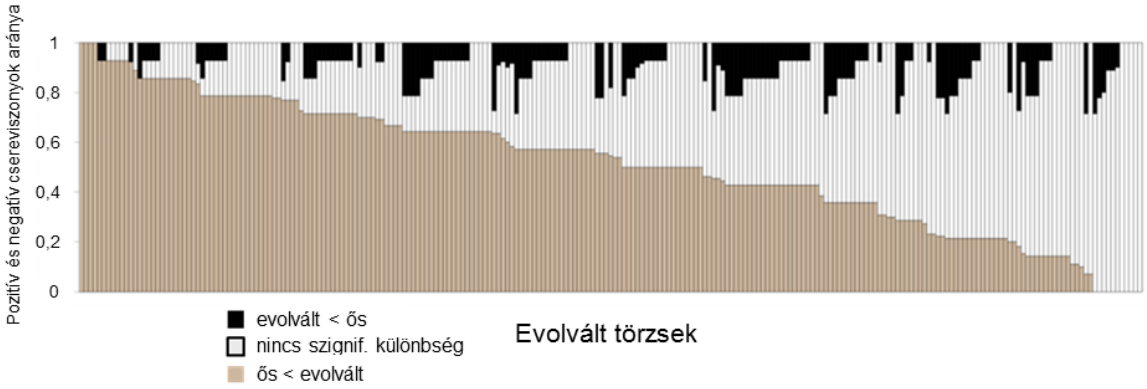
Az egy ősből kiinduló evolvált vonalak varianciája megnőtt a környezet megváltozásával (24. ábra, (Mann-Whitney próba, $p < 10^{-}7$)). Emellett a más környezetekben mutatott variancia pozitívan korrelált az eredeti környezetben mutatott fitneszemelkedéssel (Spearman korreláció, $\rho = 0{,}36$, $p = 0{,}0001$). A változatosságot itt is a variációs koefficienssel mértük, ami az átlagra normálva van, ezért a varianciában mutatott különbség nem az átlag fitneszek közötti különbségekből eredő műtermék. Tehát minél jelentősebb kompenzáció történt az evolúció során, ez annál erősebb fenotípusos diverzifikációt eredményezett más környezetekben. Ha részleteiben is megvizsgáljuk az egyes vonalak különböző környezetekben mutatott fitneszét a vad típushoz képest (25. ábra), azt találjuk, hogy a vonal - környezet párosítások 52%-ában a kompenzált vonal az új környezetben is növelte a fitneszét, azaz ezekben az esetekben a kompenzáció a külső feltételek megváltozása esetén is működött. Ugyanakkor, a vonal - környezet párosítások egy elég nagy hányadában, 8%-ában, az eredeti táptalajon kompenzációt mutató vonal az új környezetben szignifikánsan rosszabbul nőtt, mint az ősi genotípus, amelyből kifejlődött a laborevolúció során. Ezekben az esetekben a kompenzáció olyannyira specifikus volt,

hogy a környezet megváltozásával nemcsak, hogy nem javult az evolvált vonal fitnesze az őshöz képest, de egyenes romlott (antagonisztikus pleiotrópia). Az esetek maradék 40%-ában nem találtunk szignifikáns eltérést az eredeti és az új táptalajon mutatott fitnesz között (25. ábra).

Néhány olyan esetet is találtunk, ahol az evolvált vonal egy adott környezetben gyorsabb növekedést mutatott a vad típusnál. Ezeket az eseteket független mérésekkel igazoltuk és arra mutatnak példát, hogy kompenzáló evolúció segítségével a génvesztés elősegítheti új környezetekhez való adaptációt. A jelenség kivizsgálásra feltétlenül érdemes, de túlmutat doktori munkám keretein.



24. ábra. **A környezet megváltoztatásával a független evolúciós vonalak fitnesz varianciája megnő.**
Az egy ősből származó evolvált vonalak fitneszértékeiben mutatott variációs koefficiensét (szórás/átlag) mutatja az ábra a szelekciónál használt (gazdag) tápoldatban és egyéb környezetekben. A variációs koefficiensek átlagát és a standard hibát ábrázoltuk.



25. ábra. **A laborevolúció alatt létrejött fitneszemelkedés gyakran környezetfüggő.**
Az ábra az evolvált vonalakban mutatja, hogy a kompenzáció a vizsgált környezetek mekkora arányában áll fenn (evolvált fitnesz > ős fitnesz, barna szín), nem áll fenn (evolvált fitnesz ≈ ős fitnesz, fehér szín), vagy éppen fordul meg (evolvált fitnesz < ős fitnesz, fekete szín).

5. Diszkusszió

Doktori munkám témája a kompenzáló evolúció, amely egy adaptációs folyamat. Az adaptív evolúció során az organizmus egy optimum irányába mozog, amelynek legáltalánosabban használt szemléltetése egy absztrakt fitnesztájképen történő hegymászás. A hegy csúcsa az optimális, adaptált állapot. Ez a klasszikus darwini nézet szerint előnyös változatok (mutációk) elterjedésével történik. Ugyanakkor a populációgenetikai megfontolások (genetikai sodródás) és modern molekuláris biológiai kutatások egyértelműen azt mutatják, hogy káros hatású mutációk is gyakran jelen vannak és elterjednek a fajokban [Doniger et al., 2008, Fay et al., 2002, Fay et al., 2001, Sunyaev et al., 2001, Ng and Henikoff, 2002, Chasman and Adams, 2001, MacArthur et al., 2012, Lang et al., 2013, Elena and Lenski, 2003, Blank et al., 2014, Qian et al., 2012]. Ez a mutációs teher az, ami a kompenzáló evolúciót indukálja, abban az esetben, ha nem a mutáció reverziója következik be, hanem a genom egyéb területein kompenzáló mutációk semlegesítik az eredeti káros mutációk hatását.

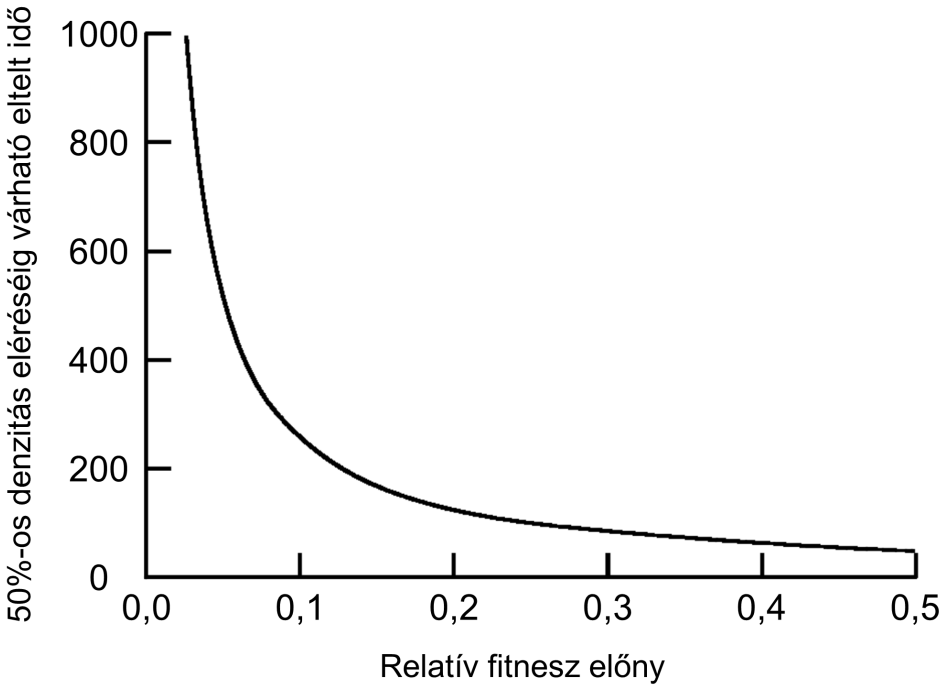
Célom munkám során a gének elvesztésére mutatott kompenzáló evolúció vizsgálata volt, melyet rendszerszinten végeztünk az élesztő (*Saccharomyces cerevisiae*) organizmusban. Teljes genomszinten vizsgálva összesen 187 olyan géndeléciós törzset találtunk, amelyek esetén az adott gén elvesztésének volt ugyan detektálható fitneszcsökkentő hatása, ugyanakkor a géndeléciós törzs még életképes a laboratóriumi körülmények között. Ilyen skálájú átfogó vizsgálat keretében eddig még nem vizsgálták a kompenzációt semmilyen élőlényben. Az egyetlen, a mi kísérletünket megközelítő spektrumú kutatás néhány hónappal cikkünk megjelenése előtt látott napvilágot [Blank et al., 2014], melyben *Escherichia coli* baktériumban 87 metabolikus gén deléciós törzsén végeztek laborevolúciós kísérletet. A kiütött gének funkciójának függvényében a 87 törzsből 22 mutatott kompenzációt. Reményeink szerint munkánk eredményeképpen általános érvényű konklúziók vonhatók le a kompenzáló evolúcióról.

A laborevolúción átesett törzsek vizsgálatakor azt találtuk, hogy a gének többségének elvesztése kompenzálható. A 187 vizsgált géndeléciós törzs 68%-a mutatta kompenzáció jelét. Ez igen magas arány, főleg, ha figyelembe vesszük, hogy kísérletünk evolúciós mércével mérve igen rövid, hozzávetőlegesen 400 generáció hosszú volt. A sörélesztő széleskörűen vizsgált modellorganizmus, amely egyrészről a laboratóriumban való könnyen kezelhetősége és kidolgozott metodológiai háttere miatt ideális választás kísérletünkhöz. Másrészt előnyt jelent az is, hogy széleskörű genomikai ismeretekkel rendelkezünk korábbi kutatások eredményeképpen. Doktori munkám során összeállítottam egy genomléptékű adatbázist, amely több, mint 200 genetikai változót tartalmaz. Ennek az adatbázisnak a segítségével olyan összefüggéseket kerestünk, amelyek megmutatják mi befolyásolja egy adott génvesztés kompenzációját. Melyek azok a gének, melyek elvesztését követően

a kompenzáció gyakran végbemegy, és mely gének delécióját nem képes kompenzálni az organizmus? Talán a legkézenfekvőbb hipotézis, amit megvizsgáltunk, hogy vajon a paralógokon keresztül történik-e a kompenzáció. Feltételezhetnénk, hogy egy gén elvesztése után annak duplikátumában történik a kompenzáció. Meglepő módon nem mutattak gyakoribb vagy erősebb kompenzációt azok a géndeléciós törzsek, ahol az inaktivált gén korábban duplikáción esett át. Egyrészről lehetséges, hogy a deletált gén paralóg párja kisebb mutációs célpontot jelent, mint sok más, potenciálisan kompenzáló gén együttesen, ezért nem találjuk gyakran kompenzáció nyomát a paralógokban. Másrészről egy redundáns paralóg pár megléte a gén elvesztésének káros hatását evolúció nélkül is gyakran csökkenti [DeLuna et al., 2008], amely kompenzációt nem egyértelmű, hogy mennyire könnyen lehet növelni adaptív evolúció által. A genomszekvenált törzseink közül is csak egyetlen egyben történt kompenzáló mutáció a paralóg párjában. Másik hipotézisünk szerint a pleiotróp gének elvesztését könnyebb kompenzálni. A pleiotróp fehérjék széleskörű hatást fejtenek ki, sok más fehérjével lépnek interakcióba, sok funkció ellátásában vesznek részt, ezáltal több élethelyzetben fontos meglétük az organizmusnak. Azt gondolhatjuk, hogy egy ilyen fehérje génjének elvesztése után a kompenzációs célpontok száma magas lesz, ezáltal nagyobb esély van rá, hogy a genomon véletlen pozícióiban megjelenő mutációk közül valamelyik kompenzáló legyen és elterjedhessen a populációban. Ennek ellenére azt találtuk, hogy a pleiotrópia sem befolyásolja egy gén kompenzálhatósági hajlamát.

A gének kompenzálhatóságának egyetlen, ám igen erős összefüggését a géndeléciós törzs fitneszével találtuk. Azon géndeléciós törzsek mutattak gyakrabban kompenzációt kísérletünk során, amelyek fitnesze alacsonyabb volt. Minél nagyobb hatása van egy gén elvesztésének, annál könnyebb ezt a hatást enyhíteni. Korábbi kis skálájú vizsgálatok is találtak ilyen összefüggést [Moore et al., 2000, Estes and Lynch, 2003, Poon and Chao, 2005, Rokyta et al., 2009], és nemrég napvilágot látott több publikáció is, melyek nagyobb skálán és más organizmusok esetén is megerősítették eredményünket [Couce and Tenaillon, 2015, Blank et al., 2014]. Több okból is kialakulhat a megfigyelt mintázat. Egyrészt a kis fitneszcsökkenést okozó génvesztések esetén azt várjuk, hogy a potenciális kompenzáló mutációk fitneszt növelő hatása is kisebb. Ez ugyanakkor azzal jár, hogy az adaptív mutáns relatív fitneszelőnye a génvesztett őshöz képest alacsony lesz. És minél kisebb egy megjelenő adaptív mutáció fitneszelőnye, annál nagyobb valószínűséggel válik a genetikai sodródás áldozatává és szorul ki egy kis populációból, vagy lassabban terjed el, ami rövid kísérletünk alatt korlátozhatja a kompenzáció detektálását (26. ábra) [Moore et al., 2000]. Mérésünk pontossága is korlátozza természetesen a nagyon kis fitneszhatások kimutatását, ugyanakkor a tapasztalt összefüggés a géndeléciós fitnesz és a kompenzálhatóság aránya között a detekciós küszöbünknél sokkal nagyobb fitnesz intervallumban fennáll, ezért ez a kísérleti limitáció nem okozhatja a talált összefüggést. Egy másik oka lehet a nagyobb hatású mutációk könnyebb kompenzálhatóságának az, hogy egy erősen káros kezdeti

mutáció hatását enyhítő kompenzáló változás káros mellékhatásaira kevésbé lesz érzékeny az organizmus. Fisher geometriai modelljének (1.1.1. fejezet) egyik fő konklúziója volt, hogy minél nagyobb fitneszesést okoz a gén elvesztése a vad típushoz képest, minél távolabb kerül az organizmus az optimumtól, annál nagyobb az esélye, hogy egy véletlen mutáció előnyös legyen, még akkor is, ha pleiotróp negatív hatása is van a fitneszre. Az elv a modell nélkül is könnyen érthető: *Ha nagy a baj, egy járulékos veszteségekkel járó megoldás is előnyös lehet.* Hogy a két magyarázat közül melyik igaz, nem tudjuk eldönteni. Valószínűleg mindkettő, különböző mértékben. Az első populációgenetikai magyarázat tulajdonképpen egy matematikai trivialitás, ami bizonyos mértékig szinte biztosan okozója a leírt mintázatnak. Ugyanakkor a különböző környezetekben tapasztalt, a kompenzált törzsek által mutatott változatos fitneszértékek (25. ábra) arra engednek következtetni, hogy a kísérletünkben lezajlott kompenzáló változásoknak gyakran voltak pleiotróp hatásai.



26. ábra. **A mutáció terjedése a szelekciós előny függvényében.**
Az ábra egy adott mutáció elterjedési sebességét mutatja a mutáció által okozott relatív fitneszelőny függvényében. A modell [Lenski et al., 1991, Gerrish and Lenski, 1998] figyelembe veszi a genetikai sodródást és az előnyös allél terjedési sebességét, 3×10^7 effektív populációméretet és 10^{-8} mutációs rátát feltételezve. Ábra módosítva [Moore et al., 2000] alapján.

Tehát az egyetlen változó, amivel a gén elvesztésének kompenzációja összefüggést mutatott, a géndeléciós törzs fitnesze. Ugyanakkor azt találtuk, hogy a különböző geno-

típusok egymástól eltérő kompenzálhatóságot mutatnak és ez az eredményünk akkor is megmarad, ha kontrollálunk arra, hogy eltérő genotípusok fitnesze is eltérő (27. ábra). Tehát a különböző gének fitneszhatásuktól függetlenül is eltérnek egymástól kompenzálhatóságukban. Úgy tűnik, hiába kerestük ennek meghatározóit az élesztőről elérhető széleskörű adatbázisokban, nem leltük nyomát azoknak a génjellemzőknek, amelyek ezt meghatározhatnák. További kutatások témája lehet az ezzel kapcsolatos hipotézisek felállítása és tesztelése.

A kompenzáció kísérletes kutatásában ismétlődően felmerülő kérdés, hogy a laborban végbemenő evolúciós változások valódi kompenzációt jelentenek-e. A géndeléció hatására lecsökkent rátermettségű törzseink növelték ugyan növekedési rátájukat, de ez a növekedés nem jelent feltétlenül kompenzációt, lehet a környezethez történő egyszerű adaptáció eredménye is. Több eredményünk is alátámasztja, hogy valódi kompenzáció zajlott kísérletünkben: A géndeléciós genotípusok mellett vad típusú kontroll törzseket is alávetettünk a laborevolúciónak (lásd Anyag és Módszer fejezet). Ezek a törzsek is mutattak fitneszemelkedést a kísérlet során, de szignifikánsan kisebb mértékűt, mint a géndeléciós kompenzált törzseink (11. ábra). A kompenzációs ráta (lásd előző bekezdés) deléció-specifikussága is valódi kompenzációra utal, közvetlenebb bizonyítékot szolgáltatnak ugyanakkor genomszekvenálási eredményeink. Az evolválódó vonalakban bekövetkező mutációk között feldúsultak azok, amelyek funkcionális kapcsolatban állnak a deletált génnel (18. ábra). Emellett konvergens evolúció nyomát is nagyon ritkán találtuk, ami tápanyagadaptációra utalna [Lang et al., 2013]. A legdirektebb bizonyítéka a kompenzációnak a kompenzáló mutáció specificitása a genetikai háttérre, amelyet néhány eset részletes elemzésekor mutattunk ki. A specificitás azt jelenti, hogy a géndeléciós törzsben fitneszelőnyt biztosító mutáció vad háttéren semleges, vagy akár káros hatást fejt ki. Ilyen eseteket mutattunk ki a Δ*mdm34* törzs *mga2-1* mutáció általi kompenzációjakor, illetve a Δ*rpb9* genotípus egyik evolvált vonalában, melyben a *WHI2* gén inaktiválódott (részletekért lásd: [Szamecz et al., 2014]).

Kísérletünk tervezésekor kíváncsiak voltunk, vajon a kompenzáló evolúció során a gén inaktivációt követően milyen mértékben állítódik vissza a vad típusú állapot. Vajon okoz-e divergenciát a kompenzáció (lásd 10. ábra)? További kísérletekkel próbáltuk feltárni, hogy a génvesztésen átesett törzsek, melyek rátermettségüket megnövelték a laborevolúció során, tehát kompenzálták a kezdeti káros mutációt, a vad típusú állapot visszaállításával tették-e ezt, vagy az eredeti génfunkció helyettesítésével. Több eredményünk is azt a hipotézist támasztja alá, miszerint a kompenzáló vonalakban jelentős divergencia történt. A genomszekvenálási eredményeinkben nagyon kevés parallel mutációt találtunk. Minden géndeléciós ősből négy függetlenül evolválódó vonalat indítottunk. Ezekben a parallel vonalakban nagyon ritkán figyeltünk meg mutációt a laborevolúció során ugyanazon génekben. Ezzel szemben egy nemrég megjelent munka *Escherichia coli* baktériumban

gyakran találta parallel evolúció nyomát [Blank et al., 2014]. A szekvenálás mellett egyes ősi és evolvált törzsek transzkriptomikai analízisét is elvégeztük. Ezen vizsgálatok is magasfokú divergenciát mutattak (19., 20., 21. ábra). A vad típustól különböző ősi géndeléciós expressziós profil a kompenzáló evolúció alatt nem, vagy csak kis mértékben közelített a vad típusú expressziós állapothoz. A nyolc vizsgált törzsből egyetlen egy közelítette meg jelentős mértékben a vad típusú expressziót. A kompenzáció során bekövetkező divergencia kimutatására talán legjobb kísérleti teszt az evolválódott vonalak rátermettségét különböző környezetekben vizsgálni. Ha a kompenzáció során az inaktivált gén funkciójának visszaállítása történik [Hall and Hartl, 1974, Campbell et al., 1973], azt várjuk, hogy a kompenzált vonal a géndeléciós őshöz viszonyított fitneszemelkedése a környezettől független legyen. Ugyanakkor, ha az eredeti génfunkció helyettesítése által valósult meg a kompenzáció, akkor ezen adaptív mutációk pleiotróp jellege változatos környezetekben felderíthetőek (lásd 10. ábra). Tizennégy különböző környezetben megvizsgálva a szignifikáns kompenzációt mutató vonalak fitneszét azt tapasztaltuk, hogy az eredeti környezetben mutatott kompenzáció nagyban környezetspecifikus volt (25. ábra). A környezetek felében nem tapasztaltunk gyorsabb növekedést az evolvált vonalaknál az őshöz képest. Az esetek 8%-ában ráadásul az evolvált vonal szignifikánsan rosszabb növekedést mutatott az új környezetben, mint a géndeléciós ősi genotípus. Az eredeti génfunkció helyettesítését támasztja alá az is, a visszaállítással szemben, hogy az elvesztett gén paralógjában csupán egyetlen egyszer találtunk mutációt (lásd 4.5. fejezet).

Eredményeink alapján feltételezhetjük, hogy a mai fajokban a gének jelenlétében illetve hiányában megfigyelhető különbségekért, legalábbis részben, a kompenzáló evolúció felelős. A törzsfejlődés során bekövetkező génvesztések természetesen lehetnek a megváltozott környezet következményei. Egy más környezetben az adott gén elvesztése lehet neutrális, vagy akár előnyös. Ugyanakkor megfigyelésünk, miszerint kísérletünk során a legtöbb gén elvesztésekor mutatott fitneszromlás nagy mértékben és igen rövid időn belül kompenzálódott, arra enged következtetni, hogy a törzsfejlődés során is történhettek olyan, eredetileg káros génvesztések, melyek káros hatása kompenzáló evolúció által enyhült. Ezt támasztja alá a különböző *Saccharomyces cerevisiae* törzsek között talált eltérés az esszenciális génekben, amely genetikai interakcióra és valójában kompenzációra utal [Dowell et al., 2010]. Emellett egy *Escherichia coli* baktériumon végzett kutatásban [Bergmiller et al., 2012] azt találták, hogy azon eszenciális gének vesztek el a törzsfa egyéb ágain, melyek funkcióvesztésének káros hatását más gének túlexpressziója enyhíteni tudta. Ellentmondásos módon saját adatainkon végzett analíziskor nem találtunk összefüggést az evolóciós kísérletünk során mutatott kompenzálhatóság és az élesztő génfa más ágain az adott gén által mutatott génvesztési gyakoriság között. Ez a negatív eredmény nem jelenti azt, hogy az evolúció során a génvesztések hátterében nem állhat kompenzáló evolúció, ugyanakkor megmutatja, hogy a kísérletünkben körülbelül 400 generáción át tartó laborevolúció és az

evolúciós időskálán végbemenő törzsfejlődés nem feltétlenül ugyanazon mechanizmusokon keresztül vagy ugyanazon génekre hat.

Mint a bevezetőben említettem, érdekes evolúcióbiológiai kérdés, hogy a káros mutációk fixációja milyen mértékben járul hozzá az adaptáció folyamatához. Több számítógépes szimulációs munka is azt mutatja, hogy egy káros mutáció fixációja és az azt követő kompenzáló evolúciója elősegítheti új adaptációk megjelenését [Covert et al., 2013, Lenski et al., 2003]. Valóban, evolúciós vonalaink között találtunk olyan populációkat, melyek méréseink alapján a vad típusnál magasabb rátermettséget mutattak. Ugyanakkor jelenleg ilyen irányú következtetéseket levonni merészség lenne. Ezen törzsek részletes vizsgálata és a kompenzáló evolúció innovációt elősegítő hatásának felderítése további kutatások tárgyát képezheti.

6. Köszönetnyilvánítás

Elsősorban szeretném megköszönni témavezetőmnek, Dr. Papp Balázsnak az évek során nyújtott irányítását, segítségét, mindig pozitív és segítőkész hozzáállását. Azt, hogy bízott bennem, még akkor is, amikor én sem bíztam magamban.

Nem lehetek eléggé hálás munkatársamnak, Dr. Szamecz Bélának a rengeteg erőfeszítésért, meglátásaiért és nem utolsósorban a jó hangulatért, amit az élesztő labor vezetőjeként mindig teremtett. Várunk vissza, Béla!

Köszönöm Papp Balázs csoportjában dolgozó minden munkatársamnak segítségüket, a jó munkakörnyezetet, a csocsómeccseket. Továbbá köszönöm Pál Csaba kísérletes csoportjának is a laboros segítséget, a közös projekteket. Remélem lesz még sok belőlük!

Kimondhatatlanul hálás vagyok szüleimnek, családomnak, barátaimnak, korábbi és jelenlegi mentoraimnak, Dr. Berend Mihálynak, Besnyőné Titter Beátának, Vajda Istvánnak, Csík Zoltánnak, Orosz Ferencnek, Dr. Németh Istvánnak és Dr. Farkas Illésnek. Az ő segítségük és támogatásuk nélkül biztosan nem írhattam volna le ezt a sort.

Köszönöm, Orsi!

7. Hivatkozások

[Airoldi et al., 2009] Airoldi, E. M., Huttenhower, C., Gresham, D., Lu, C., Caudy, A. A., Dunham, M. J., Broach, J. R., Botstein, D., and Troyanskaya, O. G. (2009). Predicting cellular growth from gene expression signatures. *PLoS computational biology*, 5(1):e1000257.

[Andersson and Hughes, 2010] Andersson, D. I. and Hughes, D. (2010). Antibiotic resistance and its cost: is it possible to reverse resistance? *Nature Reviews. Microbiology*, 8(4):260–271.

[Andersson and Levin, 1999] Andersson, D. I. and Levin, B. R. (1999). The biological cost of antibiotic resistance. *Current Opinion in Microbiology*, 2(5):489–493.

[Barrick et al., 2009] Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., Lenski, R. E., and Kim, J. F. (2009). Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature*, 461(7268):1243–1247.

[Barton and Keightley, 2002] Barton, N. H. and Keightley, P. D. (2002). Understanding quantitative genetic variation. *Nature Reviews. Genetics*, 3(1):11–21.

[Bell and Gonzalez, 2009] Bell, G. and Gonzalez, A. (2009). Evolutionary rescue can prevent extinction following environmental change. *Ecology Letters*, 12(9):942–948.

[Bergmiller et al., 2012] Bergmiller, T., Ackermann, M., and Silander, O. K. (2012). Patterns of Evolutionary Conservation of Essential Genes Correlate with Their Compensability. *PLoS Genetics*, 8(6):e1002803.

[Bjorkman et al., 1999] Bjorkman, J., Samuelsson, P., Andersson, D. I., and Hughes, D. (1999). Novel ribosomal mutations affecting translational accuracy, antibiotic resistance and virulence of Salmonella typhimurium. *Molecular Microbiology*, 31(1):53–58.

[Björkman et al., 1998] Björkman, J., Hughes, D., and Andersson, D. I. (1998). Virulence of antibiotic-resistant Salmonella typhimurium. *Proceedings of the National Academy of Sciences of the United States of America*, 95(7):3949–3953.

[Blank et al., 2014] Blank, D., Wolf, L., Ackermann, M., and Silander, O. K. (2014). The predictability of molecular evolution during functional innovation. *Proceedings of the National Academy of Sciences of the United States of America*, 111(8):3044–3049.

[Boeva et al., 2012] Boeva, V., Popova, T., Bleakley, K., Chiche, P., Cappo, J., Schleiermacher, G., Janoueix-Lerosey, I., Delattre, O., and Barillot, E. (2012).

Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics (Oxford, England)*, 28(3):423–425.

[Brauer et al., 2008] Brauer, M. J., Huttenhower, C., Airoldi, E. M., Rosenstein, R., Matese, J. C., Gresham, D., Boer, V. M., Troyanskaya, O. G., and Botstein, D. (2008). Coordination of Growth Rate, Cell Cycle, Stress Response, and Metabolic Activity in Yeast. *Molecular Biology of the Cell*, 19(1):352–367.

[Breslow et al., 2008] Breslow, D. K., Cameron, D. M., Collins, S. R., Schuldiner, M., Stewart-Ornstein, J., Newman, H. W., Braun, S., Madhani, H. D., Krogan, N. J., and Weissman, J. S. (2008). A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nature methods*, 5(8):711–718.

[Burch and Chao, 1999] Burch, C. L. and Chao, L. (1999). Evolution by small steps and rugged landscapes in the RNA virus phi6. *Genetics*, 151(3):921–927.

[Burger and Lynch, 1995] Burger, R. and Lynch, M. (1995). Evolution and Extinction in a Changing Environment: A Quantitative-Genetic Analysis. *Evolution*, 49(1):151.

[Campbell et al., 1973] Campbell, J. H., Lengyel, J. A., and Langridge, J. (1973). Evolution of a second gene for beta-galactosidase in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 70(6):1841–1845.

[Chasman and Adams, 2001] Chasman, D. and Adams, R. M. (2001). Predicting the functional consequences of non-synonymous single nucleotide polymorphisms: structure-based assessment of amino acid variation. *Journal of Molecular Biology*, 307(2):683–706.

[Chen et al., 1999] Chen, Y., Carlini, D. B., Baines, J. F., Parsch, J., Braverman, J. M., Tanda, S., and Stephan, W. (1999). RNA secondary structure and compensatory evolution. *Genes & Genetic Systems*, 74(6):271–286.

[Costanzo et al., 2010] Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E. D., Sevier, C. S., Ding, H., Koh, J. L., Toufighi, K., Mostafavi, S., Prinz, J., St. Onge, R. P., VanderSluis, B., Makhnevych, T., Vizeacoumar, F. J., Alizadeh, S., Bahr, S., Brost, R. L., Chen, Y., Cokol, M., Deshpande, R., Li, Z., Lin, Z.-Y., Liang, W., Marback, M., Paw, J., San Luis, B.-J., Shuteriqi, E., Tong, A. H. Y., van Dyk, N., Wallace, I. M., Whitney, J. A., Weirauch, M. T., Zhong, G., Zhu, H., Houry, W. A., Brudno, M., Ragibizadeh, S., Papp, B., Pál, C., Roth, F. P., Giaever, G., Nislow, C., Troyanskaya, O. G., Bussey, H., Bader, G. D., Gingras, A.-C., Morris, Q. D., Kim, P. M., Kaiser, C. A., Myers, C. L., Andrews, B. J., and Boone, C. (2010). The Genetic Landscape of a Cell. *Science*, 327(5964):425 –431.

[Couce and Tenaillon, 2015] Couce, A. and Tenaillon, O. A. (2015). The rule of declining adaptability in microbial evolution experiments. *Frontiers in Genetics*, 6:99.

[Covert et al., 2013] Covert, A. W., Lenski, R. E., Wilke, C. O., and Ofria, C. (2013). Experiments on the role of deleterious mutations as stepping stones in adaptive evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 110(34):E3171–3178.

[Cowperthwaite et al., 2006] Cowperthwaite, M. C., Bull, J. J., and Meyers, L. A. (2006). From bad to good: Fitness reversals and the ascent of deleterious mutations. *PLoS computational biology*, 2(10):e141.

[Darwin, 1859] Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* John Murray, London.

[DeLuna et al., 2008] DeLuna, A., Vetsigian, K., Shoresh, N., Hegreness, M., Colón-González, M., Chao, S., and Kishony, R. (2008). Exposing the fitness contribution of duplicated genes. *Nature Genetics*, 40(5):676–681.

[Denver et al., 2010] Denver, D. R., Howe, D. K., Wilhelm, L. J., Palmer, C. A., Anderson, J. L., Stein, K. C., Phillips, P. C., and Estes, S. (2010). Selective sweeps and parallel mutation in the adaptive recovery from deleterious mutation in Caenorhabditis elegans. *Genome Research*, 20(12):1663–1671.

[Deutschbauer et al., 2005] Deutschbauer, A. M., Jaramillo, D. F., Proctor, M., Kumm, J., Hillenmeyer, M. E., Davis, R. W., Nislow, C., and Giaever, G. (2005). Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics*, 169(4):1915–1925.

[Dixon et al., 2009] Dixon, S. J., Costanzo, M., Baryshnikova, A., Andrews, B., and Boone, C. (2009). Systematic mapping of genetic interaction networks. *Annual Review of Genetics*, 43:601–625.

[Doniger and Fay, 2007] Doniger, S. W. and Fay, J. C. (2007). Frequent gain and loss of functional transcription factor binding sites. *PLoS computational biology*, 3(5):e99.

[Doniger et al., 2008] Doniger, S. W., Kim, H. S., Swain, D., Corcuera, D., Williams, M., Yang, S.-P., and Fay, J. C. (2008). A catalog of neutral and deleterious polymorphism in yeast. *PLoS genetics*, 4(8):e1000183.

[Dowell et al., 2010] Dowell, R. D., Ryan, O., Jansen, A., Cheung, D., Agarwala, S., Danford, T., Bernstein, D. A., Rolfe, P. A., Heisler, L. E., Chin, B., Nislow, C.,

Giaever, G., Phillips, P. C., Fink, G. R., Gifford, D. K., and Boone, C. (2010). Genotype to phenotype: a complex problem. *Science (New York, N.Y.)*, 328(5977):469.

[Dudley et al., 2005] Dudley, A. M., Janse, D. M., Tanay, A., Shamir, R., and Church, G. M. (2005). A global view of pleiotropy and phenotypically derived gene function in yeast. *Molecular Systems Biology*, 1(1):E1–E11.

[Dyson and Hurst, 2004] Dyson, E. A. and Hurst, G. D. D. (2004). Persistence of an extreme sex-ratio bias in a natural population. *Proceedings of the National Academy of Sciences*, 101(17):6520–6523.

[Elena and Lenski, 2003] Elena, S. F. and Lenski, R. E. (2003). Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews. Genetics*, 4(6):457–469.

[Estes and Lynch, 2003] Estes, S. and Lynch, M. (2003). Rapid fitness recovery in mutationally degraded lines of Caenorhabditis elegans. *Evolution; International Journal of Organic Evolution*, 57(5):1022–1030.

[Estes et al., 2011] Estes, S., Phillips, P. C., and Denver, D. R. (2011). Fitness recovery and compensatory evolution in natural mutant lines of C. elegans. *Evolution; International Journal of Organic Evolution*, 65(8):2335–2344.

[Faith et al., 2008] Faith, J. J., Driscoll, M. E., Fusaro, V. A., Cosgrove, E. J., Hayete, B., Juhn, F. S., Schneider, S. J., and Gardner, T. S. (2008). Many Microbe Microarrays Database: uniformly normalized Affymetrix compendia with structured experimental metadata. *Nucleic Acids Research*, 36(Database issue):D866–870.

[Fay et al., 2001] Fay, J. C., Wyckoff, G. J., and Wu, C. I. (2001). Positive and negative selection on the human genome. *Genetics*, 158(3):1227–1234.

[Fay et al., 2002] Fay, J. C., Wyckoff, G. J., and Wu, C.-I. (2002). Testing the neutral theory of molecular evolution with genomic data from Drosophila. *Nature*, 415(6875):1024–1026.

[Fisher, 1918] Fisher, R. A. (1918). The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh*, 52(02):399–433.

[Fisher, 1930] Fisher, R. A. (1930). *The genetical theory of natural selection.* Oxford University Press, Oxford, a complete variorum ed edition.

[Gerrish and Lenski, 1998] Gerrish, P. J. and Lenski, R. E. (1998). The fate of competing beneficial mutations in an asexual population. *Genetica*, 102-103(1-6):127–144.

[Giaever et al., 2002] Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., Arkin, A. P., Astromoff, A., El-Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K.-D., Flaherty, P., Foury, F., Garfinkel, D. J., Gerstein, M., Gotte, D., Güldener, U., Hegemann, J. H., Hempel, S., Herman, Z., Jaramillo, D. F., Kelly, D. E., Kelly, S. L., Kötter, P., LaBonte, D., Lamb, D. C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S. L., Revuelta, J. L., Roberts, C. J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D. D., Sookhai-Mahadeo, S., Storms, R. K., Strathern, J. N., Valle, G., Voet, M., Volckaert, G., Wang, C.-y., Ward, T. R., Wilhelmy, J., Winzeler, E. A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J. D., Snyder, M., Philippsen, P., Davis, R. W., and Johnston, M. (2002). Functional profiling of the Saccharomyces cerevisiae genome. *Nature*, 418(6896):387–391.

[Gillespie, 1984] Gillespie, J. H. (1984). Molecular Evolution Over the Mutational Landscape. *Evolution*, 38(5):1116.

[Gillespie, 2002] Gillespie, J. H. (2002). *The causes of molecular evolution*. Number 2 in Oxford series in ecology and evolution. Oxford Univ. Press, New York, reprinted edition.

[Gokhale et al., 2009] Gokhale, C. S., Iwasa, Y., Nowak, M. A., and Traulsen, A. (2009). The pace of evolution across fitness valleys. *Journal of Theoretical Biology*, 259(3):613–620.

[Gomulkiewicz and Holt, 1995] Gomulkiewicz, R. and Holt, R. D. (1995). When does Evolution by Natural Selection Prevent Extinction? *Evolution*, 49(1):201.

[Goncalves et al., 2012] Goncalves, A., Leigh-Brown, S., Thybert, D., Stefflova, K., Turro, E., Flicek, P., Brazma, A., Odom, D. T., and Marioni, J. C. (2012). Extensive compensatory cis-trans regulation in the evolution of mouse gene expression. *Genome Research*, 22(12):2376–2384.

[Gray et al., 2010] Gray, M. W., Lukes, J., Archibald, J. M., Keeling, P. J., and Doolittle, W. F. (2010). Cell biology. Irremediable complexity? *Science (New York, N.Y.)*, 330(6006):920–921.

[Hall and Hartl, 1974] Hall, B. G. and Hartl, D. L. (1974). Regulation of newly evolved enzymes. I. Selection of a novel lactase regulated by lactose in Escherichia coli. *Genetics*, 76(3):391–400.

[Harcombe et al., 2009] Harcombe, W. R., Springman, R., and Bull, J. J. (2009). Compensatory evolution for a gene deletion is not limited to its immediate functional network. *BMC evolutionary biology*, 9:106.

[Heineman et al., 2005] Heineman, R. H., Molineux, I. J., and Bull, J. J. (2005). Evolutionary robustness of an optimal phenotype: re-evolution of lysis in a bacteriophage deleted for its lysin gene. *Journal of Molecular Evolution*, 61(2):181–191.

[Hurst, 1992] Hurst, L. D. (1992). Intragenomic Conflict as an Evolutionary Force. *Proceedings of the Royal Society B: Biological Sciences*, 248(1322):135–140.

[Ivankov et al., 2014] Ivankov, D. N., Finkelstein, A. V., and Kondrashov, F. A. (2014). A structural perspective of compensatory evolution. *Current Opinion in Structural Biology*, 26:104–112.

[Johanson et al., 1996] Johanson, U., Ævarsson, A., Liljas, A., and Hughes, D. (1996). The Dynamic Structure of EF-G Studied by Fusidic Acid Resistance and Internal Revertants. *Journal of Molecular Biology*, 258(3):420–432.

[Kauffman and Levin, 1987] Kauffman, S. and Levin, S. (1987). Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology*, 128(1):11–45.

[Kenigsberg et al., 2010] Kenigsberg, E., Bar, A., Segal, E., and Tanay, A. (2010). Widespread compensatory evolution conserves DNA-encoded nucleosome organization in yeast. *PLoS computational biology*, 6(12):e1001039.

[Kern and Kondrashov, 2004] Kern, A. D. and Kondrashov, F. A. (2004). Mechanisms and convergence of compensatory evolution in mammalian mitochondrial tRNAs. *Nature Genetics*, 36(11):1207–1212.

[Kimura, 1968] Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature*, 217(5129):624–626.

[Kimura, 1983] Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.

[Kimura, 1985] Kimura, M. (1985). The role of compensatory neutral mutations in molecular evolution. *Journal of Genetics*, 64(1):7–19.

[Kirby et al., 1995] Kirby, D. A., Muse, S. V., and Stephan, W. (1995). Maintenance of pre-mRNA secondary structure by epistatic selection. *Proceedings of the National Academy of Sciences of the United States of America*, 92(20):9047–9051.

[Kohany et al., 2006] Kohany, O., Gentles, A. J., Hankus, L., and Jurka, J. (2006). Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC bioinformatics*, 7:474.

[Kondrashov et al., 2002] Kondrashov, A. S., Sunyaev, S., and Kondrashov, F. A. (2002). Dobzhansky-Muller incompatibilities in protein evolution. *Proceedings of the National Academy of Sciences*, 99(23):14878–14883.

[Landry et al., 2005] Landry, C. R., Wittkopp, P. J., Taubes, C. H., Ranz, J. M., Clark, A. G., and Hartl, D. L. (2005). Compensatory cis-trans evolution and the dysregulation of gene expression in interspecific hybrids of Drosophila. *Genetics*, 171(4):1813–1822.

[Lang et al., 2013] Lang, G. I., Rice, D. P., Hickman, M. J., Sodergren, E., Weinstock, G. M., Botstein, D., and Desai, M. M. (2013). Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature*, 500(7464):571–574.

[Lenski et al., 2003] Lenski, R. E., Ofria, C., Pennock, R. T., and Adami, C. (2003). The evolutionary origin of complex features. *Nature*, 423(6936):139–144.

[Lenski et al., 1991] Lenski, R. E., Rose, M. R., Simpson, S. C., and Tadler, S. C. (1991). Long-term experimental evolution in Escherichia coli. I. Adaptation and divergence during 2,000 generations. *American Naturalist*, 138(6):1315–1341.

[Li and Durbin, 2010] Li, H. and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, 26(5):589–595.

[Lynch et al., 1991] Lynch, M., Gabriel, W., and Wood, A. M. (1991). Adaptive and demographic responses of plankton populations to environmental change. *Limnology and Oceanography*, 36(7):1301–1312.

[MacArthur et al., 2012] MacArthur, D. G., Balasubramanian, S., Frankish, A., Huang, N., Morris, J., Walter, K., Jostins, L., Habegger, L., Pickrell, J. K., Montgomery, S. B., Albers, C. A., Zhang, Z. D., Conrad, D. F., Lunter, G., Zheng, H., Ayub, Q., DePristo, M. A., Banks, E., Hu, M., Handsaker, R. E., Rosenfeld, J. A., Fromer, M., Jin, M., Mu, X. J., Khurana, E., Ye, K., Kay, M., Saunders, G. I., Suner, M.-M., Hunt, T., Barnes, I. H. A., Amid, C., Carvalho-Silva, D. R., Bignell, A. H., Snow, C., Yngvadottir, B., Bumpstead, S., Cooper, D. N., Xue, Y., Romero, I. G., 1000 Genomes Project

Consortium, Wang, J., Li, Y., Gibbs, R. A., McCarroll, S. A., Dermitzakis, E. T., Pritchard, J. K., Barrett, J. C., Harrow, J., Hurles, M. E., Gerstein, M. B., and Tyler-Smith, C. (2012). A systematic survey of loss-of-function variants in human protein-coding genes. *Science (New York, N.Y.)*, 335(6070):823–828.

[Maynard Smith, 1970] Maynard Smith, J. (1970). Natural Selection and the Concept of a Protein Space. *Nature*, 225(5232):563–564.

[McKenna et al., 2010] McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9):1297–1303.

[Meer et al., 2010] Meer, M. V., Kondrashov, A. S., Artzy-Randrup, Y., and Kondrashov, F. A. (2010). Compensatory evolution in mitochondrial tRNAs navigates valleys of low fitness. *Nature*, 464(7286):279–282.

[Moore et al., 2000] Moore, F. B., Rozen, D. E., and Lenski, R. E. (2000). Pervasive compensatory adaptation in Escherichia coli. *Proceedings of the Royal Society B: Biological Sciences*, 267(1442):515–522.

[Muller, 1964] Muller, H. J. (1964). THE RELATION OF RECOMBINATION TO MUTATIONAL ADVANCE. *Mutation Research*, 106:2–9.

[Myers et al., 2006] Myers, C., Barrett, D., Hibbs, M., Huttenhower, C., and Troyanskaya, O. (2006). Finding function: evaluation methods for functional genomic data. *BMC Genomics*, 7(1):187.

[Natarajan et al., 2013] Natarajan, C., Inoguchi, N., Weber, R. E., Fago, A., Moriyama, H., and Storz, J. F. (2013). Epistasis Among Adaptive Mutations in Deer Mouse Hemoglobin. *Science*, 340(6138):1324–1327.

[Ng and Henikoff, 2002] Ng, P. C. and Henikoff, S. (2002). Accounting for human polymorphisms predicted to affect protein function. *Genome Research*, 12(3):436–446.

[Ohta, 1973] Ohta, T. (1973). Slightly deleterious mutant substitutions in evolution. *Nature*, 246(5428):96–98.

[Ohta, 1992] Ohta, T. (1992). The Nearly Neutral Theory of Molecular Evolution. *Annual Review of Ecology and Systematics*, 23(1):263–286.

[Orr, 1998] Orr, H. A. (1998). The Population Genetics of Adaptation: The Distribution of Factors Fixed during Adaptive Evolution. *Evolution*, 52(4):935.

[Orr, 1999] Orr, H. A. (1999). The evolutionary genetics of adaptation: a simulation study. *Genetical Research*, 74(3):207–214.

[Orr, 2000] Orr, H. A. (2000). Adaptation and the cost of complexity. *Evolution; International Journal of Organic Evolution*, 54(1):13–20.

[Orr, 2002] Orr, H. A. (2002). The population genetics of adaptation: the adaptation of DNA sequences. *Evolution; International Journal of Organic Evolution*, 56(7):1317–1330.

[Orr, 2005a] Orr, H. A. (2005a). The probability of parallel evolution. *Evolution; International Journal of Organic Evolution*, 59(1):216–220.

[Orr, 2005b] Orr, H. A. (2005b). Theories of adaptation: what they do and don?t say. *Genetica*, 123(1-2):3–13.

[Orr and Unckless, 2008] Orr, H. A. and Unckless, R. L. (2008). Population extinction and the genetics of adaptation. *The American Naturalist*, 172(2):160–169.

[Pazos and Valencia, 2008] Pazos, F. and Valencia, A. (2008). Protein co-evolution, co-adaptation and interactions. *The EMBO journal*, 27(20):2648–2655.

[Pease et al., 1989] Pease, C. M., Lande, R., and Bull, J. J. (1989). A Model of Population Growth, Dispersal and Evolution in a Changing Environment. *Ecology*, 70(6):1657–1664.

[Poon and Chao, 2005] Poon, A. and Chao, L. (2005). The rate of compensatory mutation in the DNA bacteriophage phiX174. *Genetics*, 170(3):989–999.

[Poon and Chao, 2006] Poon, A. F. Y. and Chao, L. (2006). Functional origins of fitness effect-sizes of compensatory mutations in the DNA bacteriophage phiX174. *Evolution; International Journal of Organic Evolution*, 60(10):2032–2043.

[Pu et al., 2009] Pu, S., Wong, J., Turner, B., Cho, E., and Wodak, S. J. (2009). Up-to-date catalogues of yeast protein complexes. *Nucleic Acids Research*, 37(3):825 –831.

[Qian et al., 2012] Qian, W., Ma, D., Xiao, C., Wang, Z., and Zhang, J. (2012). The genomic landscape and evolutionary resolution of antagonistic pleiotropy in yeast. *Cell Reports*, 2(5):1399–1410.

[Rancati et al., 2008] Rancati, G., Pavelka, N., Fleharty, B., Noll, A., Trimble, R., Walton, K., Perera, A., Staehling-Hampton, K., Seidel, C. W., and Li, R. (2008).

Aneuploidy underlies rapid adaptive evolution of yeast cells deprived of a conserved cytokinesis motor. *Cell*, 135(5):879–893.

[Rokyta et al., 2002] Rokyta, D., Badgett, M. R., Molineux, I. J., and Bull, J. J. (2002). Experimental genomic evolution: extensive compensation for loss of DNA ligase activity in a virus. *Molecular Biology and Evolution*, 19(3):230–238.

[Rokyta et al., 2009] Rokyta, D. R., Abdo, Z., and Wichman, H. A. (2009). The Genetics of Adaptation for Eight Microvirid Bacteriophages. *Journal of Molecular Evolution*, 69(3):229–239.

[Ruepp et al., 2004] Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Güldener, U., Mannhaupt, G., Münsterkötter, M., and Mewes, H. W. (2004). The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Research*, 32(18):5539–5545.

[Schrag and Perrot, 1996] Schrag, S. J. and Perrot, V. (1996). Reducing antibiotic resistance. *Nature*, 381(6578):120–121.

[Schrag et al., 1997] Schrag, S. J., Perrot, V., and Levin, B. R. (1997). Adaptation to the fitness costs of antibiotic resistance in Escherichia coli. *Proceedings. Biological Sciences / The Royal Society*, 264(1386):1287–1291.

[Sherman et al., 1996] Sherman, D. R., Mdluli, K., Hickey, M. J., Arain, T. M., Morris, S. L., Barry, C. E., and Stover, C. K. (1996). Compensatory ahpC Gene Expression in Isoniazid-Resistant Mycobacterium tuberculosis. *Science*, 272(5268):1641–1643.

[Stark, 2006] Stark, C. (2006). BioGRID: a general repository for interaction datasets. *Nucleic Acids Research*, 34(90001):D535–D539.

[Stephan and Kirby, 1993] Stephan, W. and Kirby, D. A. (1993). RNA folding in Drosophila shows a distance effect for compensatory fitness interactions. *Genetics*, 135(1):97–103.

[Sunyaev et al., 2001] Sunyaev, S., Ramensky, V., Koch, I., Lathe, W., Kondrashov, A. S., and Bork, P. (2001). Prediction of deleterious human alleles. *Human Molecular Genetics*, 10(6):591–597.

[Szamecz et al., 2014] Szamecz, B., Boross, G., Kalapis, D., Kovács, K., Fekete, G., Farkas, Z., Lázár, V., Hrtyan, M., Kemmeren, P., Groot Koerkamp, M. J. A., Rutkai, E., Holstege, F. C. P., Papp, B., and Pál, C. (2014). The genomic landscape of compensatory evolution. *PLoS biology*, 12(8):e1001935.

[Teng et al., 2013] Teng, X., Dayhoff-Brannigan, M., Cheng, W.-C., Gilbert, C. E., Sing, C. N., Diny, N. L., Wheelan, S. J., Dunham, M. J., Boeke, J. D., Pineda, F. J., and Hardwick, J. M. (2013). Genome-wide consequences of deleting any single gene. *Molecular cell*, 52(4):485–494.

[Torres et al., 2007] Torres, E. M., Sokolsky, T., Tucker, C. M., Chan, L. Y., Boselli, M., Dunham, M. J., and Amon, A. (2007). Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science (New York, N.Y.)*, 317(5840):916–924.

[Wang et al., 2010] Wang, Z., Liao, B.-Y., and Zhang, J. (2010). Genomic patterns of pleiotropy and the evolution of complexity. *Proceedings of the National Academy of Sciences*.

[Warringer and Blomberg, 2003] Warringer, J. and Blomberg, A. (2003). Automated screening in environmental arrays allows analysis of quantitative phenotypic profiles in Saccharomyces cerevisiae. *Yeast (Chichester, England)*, 20(1):53–67.

[Weinreich and Chao, 2005] Weinreich, D. M. and Chao, L. (2005). Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. *Evolution; International Journal of Organic Evolution*, 59(6):1175–1182.

[Weinreich et al., 2005] Weinreich, D. M., Watson, R. A., and Chao, L. (2005). Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution; International Journal of Organic Evolution*, 59(6):1165–1174.

[Weissman et al., 2009] Weissman, D. B., Desai, M. M., Fisher, D. S., and Feldman, M. W. (2009). The rate at which asexual populations cross fitness valleys. *Theoretical Population Biology*, 75(4):286–300.

[Wittkopp et al., 2004] Wittkopp, P. J., Haerum, B. K., and Clark, A. G. (2004). Evolutionary changes in cis and trans gene regulation. *Nature*, 430(6995):85–88.

[Wolfe and Shields, 1997] Wolfe, K. H. and Shields, D. C. (1997). Molecular evidence for an ancient duplication of the entire yeast genome. *Nature*, 387(6634):708–713.

[Wright, 1932] Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. 1:356–366.

[Wright, 1970] Wright, S. (1970). Random Drift and the Shifting Balance Theory of Evolution. In Kojima, K.-i., editor, *Mathematical Topics in Population Genetics*, pages 1–31. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Xie and Tammi, 2009] Xie, C. and Tammi, M. T. (2009). CNV-seq, a new method to detect copy number variation using high-throughput sequencing. *BMC bioinformatics*, 10:80.

[Zhang and Hill, 2005] Zhang, X.-S. and Hill, W. G. (2005). Evolution of the environmental component of the phenotypic variance: stabilizing selection in changing environments and the cost of homogeneity. *Evolution; International Journal of Organic Evolution*, 59(6):1237–1244.

8. Összefoglaló

A klasszikus darwini evolúciós nézet szerint az adaptáció előnyös változatok egymás utáni elterjedésével, a populációban történő fixálódásával történik. Ugyanakkor káros mutációk is elterjedhetnek a genomban különböző mechanizmusokon keresztül, úgymint genetikai sodródás, "hitchhiking" vagy antagonisztikus pleiotrópia (egy mutáció adott környezetben előnyös hatású, de a környezet megváltozásával hatása épphogy károssá változik). Miután egy káros mutáció fixálódott egy populációban, annak negatív hatása az organizmus rátermettségére enyhülhet kompenzáló evolúció által. Tehát a kompenzáló evolúció egy adaptív folyamat, ahol a kezdeti káros mutációra specifikus genetikai változások terjednek el a populációban a természetes szelekció által.

Doktori munkám során élesztő modellorganizmusban (*Saccharomyces cerevisiae*) rendszerszinten vizsgáltuk a kompenzáló evolúciót. Ehhez egy haploid géndeléciós könytárban kerestük meg azokat az életképes törzseket, amelyek szignifikáns fitneszcsökkenést mutattak a vad típushoz képest. Összesen 187 ilyen genotípust azonosítottunk, melyekből laboratóriumi evolúciót indítottunk, genotípusonként négy, egymástól független vonalban. Az ősi és evolvált vonalak rátermettségét növekedési rátaként mértük, amelyet folyékony tápoldatban egy robotizált rendszer segítségével felvett növekedési görbéből számoltunk ki. Kompenzált vonalnak azokat a populációkat neveztük, amelyek szignifikáns fitneszjavulást mutattak az őshöz képest és ennek a javulásnak a mértéke nagyobb volt, mint az evolválódó kontroll vad típusú vonalak által mutatott javulás. Ez utóbbi feltétel a kompenzáló evolúció kutatásában általánosan alkalmazott. Célja, hogy megbizonyosodjunk róla, hogy az evolvált vonalakban tapasztalt fitneszemelkedés valódi kompenzáció (az ősben meglévő káros mutációra specifikus adaptáció) és nem egyszerűen a tápoldathoz való adaptáció. Evolvált populációnk vizsgálatakor azt találtuk, hogy a legtöbb gén elvesztését könnyen kompenzálni lehetett. Az ősi géndeléciós törzseink 68%-a mutatott kompenzációt a belőle indított legalább egyik független evolvált vonalban. A kompenzáció mértéke is jelentős volt. A legtöbb esetben a génvesztés által okozott fitneszesés legalább fele kijavításra került a kompenzáló evolúció által, és a törzsek egy jelentős hányadában visszaállítódott a vad típusú növekedési ráta.

Tehát azt találtuk, hogy a kompenzáció gyakori, gyorsan végbemegy, sokféle gén elvesztéséhez képes alkalmazkodni a sejt. A továbbiakban arra voltunk kíváncsiak, hogy vajon a géneknek van-e olyan jellemzője, ami meghatározza, hogy deléciójuk kompenzálható-e. Ennek eldöntésére létrehoztunk egy genomléptékű adatbázist a sörélesztő szerteágazó genomikai adataival. Néhány specifikus hipotézist állítottunk fel, melyeket tesztelni akartunk. Először kíváncsiak voltunk, hogy a kompenzáció gyakran történik-e a deletált gén paralógjain keresztül, feltéve, hogy nem szingleton génről van szó. Naivan azt képzelhetnénk, amit korábbi kisléptékű vizsgálatok is alátámasztanak, hogy egy gén elvesztése esetén dup-

likátuma pótolhatja annak funkcióját, ezáltal kompenzálva a gén elvesztését. Másodszor azt is tesztelni akartuk, hogy vajon egy pleiotróp gén (amelynek magas a fehérje-fehérje, vagy genetikai kölcsönhatás fokszáma, sok környezetben mutat fitneszhatást) elvesztését könnyebb-e kompenzálni. Emögött az a gondolatmenet áll, hogy ha egy sokféle funkcióval rendelkező fehérje hiányzik a sejtből, a mutációs célpontok száma, ami által a kompenzáció végbemehet, nagyobb lesz. Tehát a hipotézis szerint egy pleiotróp gén elvesztése után a véletlen mutációk nagyobb aránya lehet előnyös. Meglepetésünkre sem a paralógok jelenléte, sem a deletált gén pleiotrópia mérőszámai nem befolyásolták a kompenzálhatóságot kísérletünkben. Az egyetlen erős prediktor, amit találtunk, az ősi géndeletált törzs fitnesze. Minél alacsonyabb a rátermettsége az ősi genotípusnak, annál könnyebben kompenzálható a génvesztés. Ennek megértésére két magyarázattal is szolgálhatunk. Egyrészt, ha az ősi törzs fitnesze alacsony, akkor egy adott fitnesznövekmény magasabb szelektív előnyt fog okozni egy adaptív mutánsnak, ezáltal az könnyebben el fog tudni terjedni a populációban. És fordítva, minél kisebb a kezdeti fitneszesés, annál kisebb lesz a szelekciós koefficiens és annál nagyobb lesz az esélye, hogy egy adaptív mutáció a genetika sodródás áldozatává váljon. Másrészt, ha a kompenzáció során nem az elvesztett gén eredeti funkciója állítódik helyre, hanem a kompenzációnak pleiotróp, negatív mellékhatásai is vannak, abban az esetben minél nagyobb a génvesztés által okozott kezdeti fitneszesés, annál kisebb hatása lesz ezeknek a mellékhatásoknak a fitneszre. Ezt a gondolatmenetet legegyszerűbben a Fisher-féle geometriai adaptációs modell segítségével érthetjük meg. Minél messzebb van egy populáció az optimumtól, annál kisebb lesz a relatív nagysága a kompenzáció pleiotrópikus, káros hatásainak. A modell kimondja, minél távolabb áll egy populáció az optimumtól, annál nagyobb a valószínűsége, hogy egy véletlen irányú mutáció előnyös legyen.

A kompenzáló evolúció genetikai mechanizmusának kivizsgálása céljából néhány kiválasztott törzset genomszekvenálásnak vetettünk alá. Genomonként átlagosan 0.5 kisméretű inzerciót, illetve deléciót, 0.34 szegmens, vagy teljes kromoszóma duplikációt és 6 pontmutációt találtunk. A pontmutációk további analízise során pozitív szelekció jelét találtuk (a neutrális evolúció alapján vártnál magasabb dN/dS arány) és a mutációk feldúsultak olyan génekben, amelyek funkcionális kapcsolatban vannak az ősben inaktivált génnel. Gyakrabban történt mutáció olyan génben, amely a deletált génnel ugyanazon funkcionális besorolásban van, hasonló a genetikai interakciós profiljuk, koexpresszáltak vagy éppen egy fehérje komplex tagjai. Ezek az eredmények mind azt mutatják, hogy a laborevolúció alatt felhalmozott genetikai változásoknak legalábbis egy része specifikus az ősi géndelécióra. Mivel egy adott ősből több független evolúciós vonal genomját is megszekvenáltuk, így eredményeink között tudtunk parallel mutációkat keresni. Parallelnek tekintünk két mutációt, ha ugyanazon ősi genotípusból származó két független evolúciós vonalban ugyanazon gén szekvenciájában történt mutáció. Meglepően kevés esetben találtunk ilyet, a mutáci-

óknak mindössze 5,3%-a volt parallel mutáció. Ez azt jelenti, hogy míg a kompenzáció alatt felhalmozódó mutációk specifikusak az elvesztett génre, addig a kompenzáló evolúció nem determinisztikus, az evolválódó vonalak divergenciáját okozza.

A kompenzáció mechanizmusának további vizsgálata céljából nyolc evolvált vonalon és ősi törzseiken transzkriptomikai analízist végeztünk. Meg akartuk vizsgálni a kompenzáció során történő fiziológiai változásokat. Különösen kíváncsiak voltunk, hogy a géndeléciós ősi törzsben a vad típushoz képest megváltozott expressziójú gének expressziója visszaáll-e a vad típus szintjére a kompenzált vonalakban, vagy ellenkezőleg, a kompenzáció során felgyülemlő genetikai változások az evolvált vonalak új fiziológiai állapotok felé történő divergenciáját okozzák-e? Azt találtuk, hogy az evolvált vonalakban általában nem állt helyre a vad típusú génexpressziós állapot, amely további bizonyítékot szolgáltat a kompenzáló evolúciónak genetikai divergenciát okozó hatásáról.

A kompenzáció által okozott rejtett különbségek kimutatásának egy másik módja, ha kompenzációt mutató, függetlenül evolválódó vonalak rátermettségét egyéb környezetekben is megvizsgáljuk. Amíg sok evolvált vonalban a génvesztés által okozott kezdeti fitneszesés után a fitnesz javult a gazdag tápoldatban, melyen az evolúciós kísérlet zajlott, addig más környezetben mérve a törzsek rátermettséget, különböző eredményeket kaphatunk. Ha az eredeti gén funkciója helyreállt a kompenzáció során, ha a vad típusú fiziológia újra fennáll, akkor azt várjuk, hogy az evolvált vonal fitneszemelkedés általános jellegű legyen, ne függjön a külső környezettől. Ugyanakkor, ha az elvesztett gén funkciója nem állt helyre, a funkcióra való igény csökkent csak a sejtben, akkor azt várjuk, hogy az evolvált vonalak rátermettsége jelentősen különbözzön a különböző környezetekben. Hogy a két szcenárió között különbséget tudjunk tenni, 237 evolvált vonalat, melyek a gazdag tápoldatban kompenzációt mutattak, 14 különböző környezetben fitneszmérésnek vetettünk alá. Míg az evolvált vonal - környezeti párosítások közel felében az evolvált vonal az új környezetben is jobban nőtt, mint az ősi genotípus, addig sok esetben az evolvált vonal nem mutatott előnyösebb növekedést, sőt, a vonal - környezet párok 8%-ában az evolvált vonal szignifikánsan rosszabb fitneszt mutatott az új környezetben, mint a géndeléciós ősi genotípus. Emellett, az egy ősből leszármazó független vonalak által mutatott variancia a rátermettségben jóval nagyobbnak bizonyult az új környezetek esetén, mint a gazdag tápoldatban. Ezek az eredmények, a genomszekvenálási és transzkriptomikai eredményekkel együtt, azt mutatják, hogy a kompenzáló evolúció a sörélesztő géndeléciós törzseiben jelentős genetikai divergenciához vezetett.

Végül, eredményeink arra engednek következtetni, hogy a kompenzáló evolúció szerepet játszhatott a mai fajok génállománya között megfigyelhető különbségek kialakításában. A kompenzáció genetikai divergenciához vezethet egy populációban és megváltoztathatja, potenciálisan akár elő is segítheti annak adaptációját.

9. Summary

In the original Darwinian view of evolution, adaptation happens through sequential fixation of adaptive changes in the population. However, deleterious mutations can fix in populations through different mechanisms, like genetic drift or draft (hitchhiking) or antagonistic pleiotropy. Once a deleterious mutation is fixed in the population, its negative effect on the organism's fitness can be mitigated through compensatory evolution. Thus, compensatory evolution is an adaptive process, where specific mutational changes are driven to fixation by positive selection to suppress the effect of a harmful mutation.

In this study we set out to systematically study compensatory evolution in the fungi model organism, *Saccharomyces cerevisiae*. To do so, we screened a haploid gene deletion library for viable mutant strains that exhibit a significant fitness decrease compared to the wild-type due to loss of a single gene. We identified 187 gene deletion genotypes from which we started lab evolution in four independent evolutionary lines for each ancestor genotype. We approximated fitness for the ancestor and evolved strains as growth rate, calculated from the growth curves that were monitored in rich liquid medium with a robotic system. We identified "compensated lines" as evolved strains with a significant fitness increase compared to their ancestor strain and a significantly higher fitness increment during lab evolution than what the evolving control wild-type strains showed. This latter criteria is generally used in compensatory evolution studies to ensure that the fitness increment observed in the evolving populations is due to compensation (adaptation specific to the deleterious fixed mutation in the ancestor) and not just due to adaptation to the medium. We found that the loss of most of the genes can be easily compensated for; 68% of the gene deletion genotypes showed compensatory evolution in our experiment in at least one of the four parallel evolving lineages. The extent of compensation was also substantial, meaning that in most cases at least half of the fitness decrease caused by the deletion of a gene in the ancestor was mitigated through compensatory evolution and many compensated lines approximated wild-type fitness.

So, it seems that compensatory evolution is pervasive, it happens fast and can mitigate the loss of many different genes. Next, we tried to identify gene properties that determine whether the loss of the given gene can be compensated for. To this end, we created a genome-scale database of genomic and functional genomic gene properties covering many aspects of budding yeast genetics. We had specific hypotheses to test with the available data. First, we wanted to see whether compensation generally happens through duplicates of the lost gene, when they are present in the genome. One would intuitively expect, as it was found in previous small-scale studies, that in case of loss of a gene, its duplicate could step in to execute the function of the lost gene. Second, we wanted to test whether loss of a gene with high pleiotropy (i.e., many protein-protein interactions, many

genetic interaction, many environments where the loss of the gene incurs a fitness defect) is easier to compensate for. The rationale behind this is that if a protein with widespread functions is missing from the cell, the mutational target size for compensatory evolution to work with will be greater. It means, that after the loss of a pleiotropic gene, a higher fraction of mutations can be advantageous. To our surprise, none of these hypotheses were supported by our data. However, we identified fitness of the ancestral strain as the main determinant of compensatory potential. The lower the fitness of the deletion strain, the easier to compensate for the loss of the gene. There are two intuitive explanations for this phenomenon. On one hand, if the ancestor strain has a lower fitness, a given increase in fitness will cause a higher selective advantage for an adaptive mutant, making it easier to spread and ultimately fix in the population. Basically, the smaller the initial fitness defect to compensate for, the lower the selective coefficient will be, and the the higher the chance that an adaptive mutation will be subject to genetic drift. On the other hand, if compensation not just replaces the lost function but has pleiotropic, negative side-effects, then the higher the initial fitness defect, the smaller effect these side-effects will have on fitness. This concept is easiest to understand through the geometrical model of adaptation by Fisher. The further away a population is from the optimum, the smaller the relative size of the pleiotropic negative effects of a compensatory mutation. As Fisher's model tells us, the further away a population is from the optimum, the more likely that a mutation will be advantageous.

To explore the genetic basis of compensatory evolution, we sequenced the genome of a few chosen compensated lines and their ancestor strains. We found on average 0.5 small insertions/deletions, 0.34 copy number variations and 6 single nucleotide polymorphisms (SNP) per genome. Analyzing the SNPs further, we found signs of positive selection (high dN/dS ratio) and mutations were enriched in genes that are functionally related to the gene that was deleted in the ancestor. They were more often in the same functional category, had more similar genetic interaction profiles, were more often co-expressed and were more often present in the same protein complex than expected by chance. These results show that the genetic changes that accumulated in the evolved lines are, at least partially, specific to the gene deletion in the ancestor. As we sequenced the genome of several independent evolving lines per ancestor strains, we could search our results for parallel evolutionary events, meaning that a mutation happened in the same gene in two independently evolving lines starting from the same ancestor. We found surprisingly few parallel mutations. This means that while the evolutionary changes during adaptation are specific to the gene that was deleted in the ancestor strain, compensatory evolution is not deterministic, but can lead to genetic divergence of the evolving strains.

To further examine the possible divergence behind compensation, we have performed transcriptome analysis on 8 evolving strains and their corresponding ancestors. We

wanted to gain an understanding of the physiological changes during compensation. More specifically, we were curios whether compensatory evolution restores the expression level of those genes for which gene expression substantially changed due to the deletion of a single gene in the ancestor, or on the contrary, whether compensatory genetic changes generate divergence towards new physiological states. We found that evolved lines generally did not approximate wild-type expression state, a result that further supports the divergence generating effect of compensatory evolution in yeast knock-out strains.

Another experimental approach to reveal hidden differences caused by compensation is to measure the fitness of independently evolved lines that show compensation (i.e., an increased fitness in the medium of the evolutionary experiment) in different environments. While in many evolved lines the initial fitness defect caused by the loss of a single gene got mitigated in the environment of the batch selection experiment, if we test these compensated lines in other environments, we can expect different scenarios. If the function of the lost gene got restored during compensation, if the wild-type physiology is re-established, then we expect the fitness increment showed by the evolved line to be general, not to be specific to the environment. However, if the function of the lost gene is only substituted (i.e., the need for the function is reduced) then we expect the evolved lines to show substantial variation in fitness in different environments. To be able to distinguish between the two scenarios, we measured the fitness of 237 evolved lines that showed compensation in the medium of the selection experiment in 14 different environments, media with different carbon sources or stress factors. While around half of the evolved line - environment pairs showed compensation (evolved line showing higher fitness than ancestor in new environment), in many environments the evolved line did not have a significantly higher fitness than its ancestor and surprisingly, in 8% of the evolved line - environment pairs, the evolved line showed significantly lower fitness than its ancestor strain. Also, the variance in fitness among independently evolving lines descended from the same ancestor was generally higher in other media than in the medium of the selection experiment. These results, in line with the transcriptome and genome sequencing data of the selected lines, support the hypothesis that compensatory evolution following loss of a single gene resulted in substantial genetic divergence of the strains.

Finally, our results suggest that compensatory evolution might have a role in contributing to the gene content variation that is observed between extant species. Compensation can lead to genetic divergence in a population and might facilitate and modify the route of adaptation.

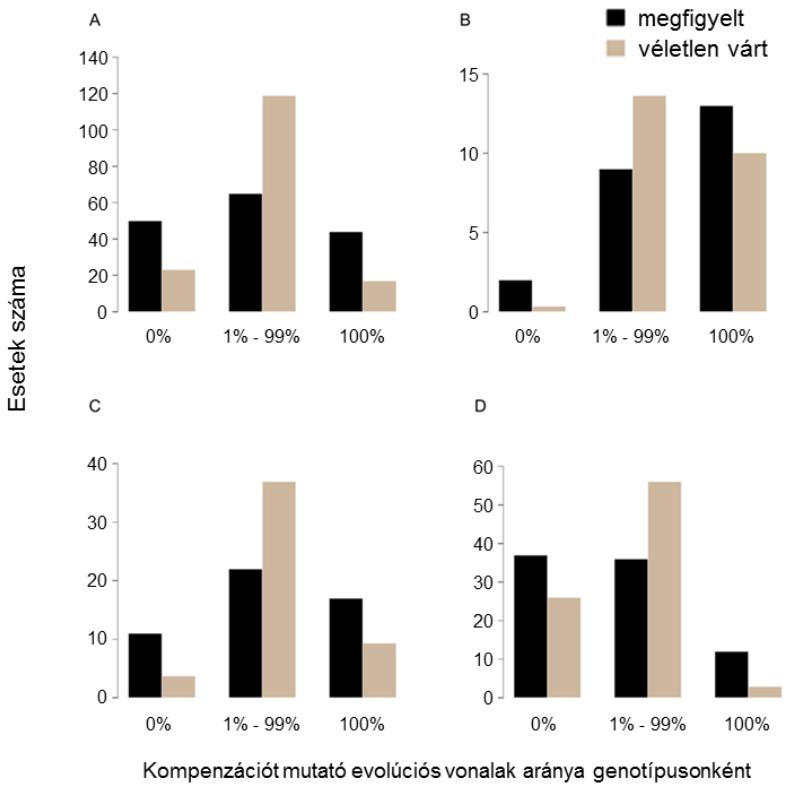
10. Függelék

10.1. A genomszekvenálásra és transzkriptomikai vizsgálatra kiválasztott evolúciós vonalak

Evolúciós vonal	Növekedési ráta (ős)	Növekedési ráta (evolvált)	Genomszekvenálás	Transzkriptomikai analízis
YBR200W 2. vonal	0,728	0,943	1	0
YBR200W 3. vonal	0,728	0,975	1	0
YBR267W 2. vonal	0,739	0,829	1	0
YBR267W 3. vonal	0,739	1,341	1	0
YER016W 1. vonal	0,853	1,035	1	0
YER016W 3. vonal	0,853	1,046	1	0
YER016W 4. vonal	0,853	1,211	1	0
YER083C 1. vonal	0,847	1,050	1	0
YER083C 3. vonal	0,847	1,036	1	0
YER083C 4. vonal	0,847	1,254	1	0
YGL070C 1. vonal	0,530	0,809	1	1
YGL070C 2. vonal	0,530	0,942	1	0
YGL219C 2. vonal	0,788	1,016	1	0
YGL219C 3. vonal	0,788	0,992	1	0
YGL219C 4. vonal	0,788	1,012	1	0
YGL219C 1. vonal	0,788	1,058	1	1
YGR104C 2. vonal	0,623	0,762	1	0
YGR104C 1. vonal	0,623	1,027	1	1
YGR104C 4. vonal	0,623	0,784	1	0
YJL180C 2. vonal	0,539	0,664	1	0
YJL180C 3. vonal	0,539	1,409	1	0
YJL180C 4. vonal	0,539	0,703	1	0
YLR074C 2. vonal	0,716	0,930	1	1
YLR074C 4. vonal	0,716	0,928	1	0
YLR074C 3. vonal	0,716	0,881	1	0
YLR074C 1. vonal	0,716	0,967	1	0
YLR357W 2. vonal	0,768	1,035	1	0
YLR357W 1. vonal	0,768	1,055	1	1
YLR357W 4. vonal	0,768	1,034	1	0
YLR448W 1. vonal	0,847	1,060	1	0
YLR448W 3. vonal	0,847	1,203	1	0
YLR448W 2. vonal	0,847	1,175	1	0
YNR052C 3. vonal	0,723	0,778	1	0
YNR052C 1. vonal	0,723	0,946	1	1
YOL023W 2. vonal	0,613	1,433	1	0
YOL023W 1. vonal	0,613	1,205	1	1
YOL023W 3. vonal	0,613	1,507	1	0
YOR026W 1. vonal	0,817	1,042	1	0
YOR026W 2. vonal	0,817	1,039	1	0
YOR026W 3. vonal	0,817	1,389	1	0
YOR026W 4. vonal	0,817	1,008	1	0
YPR043W 1. vonal	0,586	1,183	0	1

2. táblázat. **A genomszekvenálásra és transzkriptomikai vizsgálatra kiválasztott evolúciós vonalak.**
A táblázat tartalmazza azon vonalakat, amelyeket kiválasztottunk genomszekvenálásra, illetve genom skálájú génexpressziós vizsgálatra. A gazdag táptalajon mért ősi és evolvált fitneszüket is mutatjuk. A többi vizsgált vonal fitnesz adatait és a részletes szekvenálási és expressziós adatsorokat lásd: [Szamecz et al., 2014].

10.2. A gének kompenzálhatóságban mutatott különbözősége nem pusztán a géndeléció fitneszhatásától függ



27. ábra. **Különböző gének elvesztése eltérő mértékben kompenzálható, függetlenül a génvesztés fitneszhatásától.**
Minden géndeléciós ősből négy független vonalon indítottuk a laborevolúciót. A géneket csoportosítottuk aszerint, hogy a belőlük indított evolúciós vonalak egyike sem (0%)/ legalább egyike, de nem az összes (1% - 99%) / mindegyike (100%) mutatott-e kompenzációt. A kompenzálódást mutató vonalakat a genotípusok között véletlenszerűen megkeverve (randomizációs teszt) vizsgáltuk, hogy a különböző gének eltérnek-e kompenzálhatóságukban. 1000-szer randomizáltuk a kompenzációt mutató vonalakat a genotípusok között, ezáltal kaptunk egy eloszlást, amelyet akkor várnánk, ha a genotípusok nem különböznének kompenzálhatóságukban (szürke oszlopok). Az általunk megfigyelt valódi eloszlás (fekete oszlopok) ettől eltér, a különböző gének inherensen különböznek kompenzálhatóságukban. Ezt az összefüggést nem pusztán az okozza, hogy a nagyobb fitneszhatású deléciók könnyebben kompenzálhatóak: A 14. ábrát módosítottuk úgy, hogy a géneket csoportosítottuk a géndeléciós fitneszük alapján. Az A alábra az összes génre mutatja az összefüggést és megegyezik a 14. ábrával. B/C/D alábrák a <0,7 / 0,7-0,8 / >0,8 fitnesz intervallumokra mutatják ugyanazt az összefüggést. A valós és a véletlen várt eloszlások mind a négy esetben szignifikáns eltérést mutattak (χ^2 próba, p = < 10^{-20} / 0,013 / 7×10^{-6} / 10^{-8} az A/B/C/D alábrák esetében.) Tehát a gének különbözőségét kompenzálhatóságukban adott fitneszhatású génvesztések csoportján belül is kimutattuk, ezért valószínűtlen, hogy a kompenzálhatóságban talált különbséget egyedül a génkiütés fitneszköltsége határozza okozza.