

Bírálati vélemény

Dr. Papp Balázs

Evolutionary genetics in the era of genome-scale modelling

c. doktori disszertációjáról

Általános megjegyzések:

A doktori mű egy – jelentős részben a Jelölt vezetésével folyó – több, mint két évtizedes, szisztematikus kutatási program kiemelkedő eredményeinek rövid, angol nyelvű összefoglalása. A törzs-szöveg terjedelme a 7 oldalnyi irodalomjegyzékkel együtt 52 oldal, amely kiegészül a doktori mű alapját képező 5 szakkikk másolatait tartalmazó függelékkel. A kiemelt publikációk közül 3 a Nature-ben, 2 a PNAS-ben jelent meg; ez a tény önmagában híven illusztrálja a Jelölt eddigi tudományos munkájának színvonalát és a nemzetközi tudományos közösség értékítéletét a témáról és a Jelölt szerepéről annak művelésében. A doktori mű azonban mindentől eltekintve is imponálóan világos, lényegre törő, a téma diszciplináris sokszínűségét is híven bemutató és élvezetes stílusban fogalmazott szöveg, mely 6 fejezetből áll: a Bevezetésen és a Kitekintésen kívül a genom-szintű rendszerbiológiai modellezés és az ahhoz kapcsolódó empirikus vizsgálatok négy nagyobb témakörét tárgyalja 4 fejezetben, formai szempontból egy-két elütésen és egy kimaradt hivatkozáson (V. fejezet, 31. oldal alján) kívül szintén minden igényt kielégítően. A következőkben fejezetenként térek ki röviden a disszertáció tartalmi elemeinek értékelésére.

I. Introduction:

A Bevezetés a disszertáció tágabb témakörének, a metabolikus reakcióhálózatokat katalizáló enzimek mutációs-szelekciós változásának genom-szintű, evolúciós szempontú értékelésének kereteit vázolja. A rendszerbiológiai módszer lehetővé teszi a vizsgált szervezetek metabolikus hálózatait érintő genetikai változásoknak, illetve azok kombinációinak fenotípusos, végső soron fitness-értékekben mérhető hatásának magyarázatát, illetve predikcióját a környezeti kényszer-feltételek figyelembevételével. A korábbi, csak a metabolizmus egyes részeire fókuszáló vizsgálatoknál lényegesen holisztikusabb, rendszer-szintű kutatásokat a teljes genomokra kiterjedő szekvencia-adatbázisok megjelenése, illetve azokban az egyes metabolikus enzimeket kódoló gének azonosítása tette lehetővé. Az ebben rejlő óriási tudományos potenciál kiaknázásában a Jelölt és munkatársai a nemzetközi élmezőnybe tartoznak. A jelen disszertációban szereplő négy kutatási irány az evolúciogenetika rendszer-szemléletű kiterjesztésének egy-egy mérföldköve.

II. Genetic interactions in metabolic networks

A metabolikus hálózatok működését lehetővé tevő transzport-fehérjék és enzimek nyilvánvaló funkcionális kapcsolatban állnak egymással a hálózaton belüli reakciókon keresztül. Az első genom-szintű vizsgálat az ezen enzimek és/vagy transzport-fehérjék génjei közötti episztatikus

hatásokra kérdez rá a *Saccharomyces cerevisiae* metabolikus hálózatában, kölcsönhatásokat kizárólag a metabolikus reakciók sztöchiometriáján és a tápanyagok membrántranszportján keresztül véve figyelembe (vagyis az enzimkinetikai különbségeket és az esetleges allosztérikus vagy más genetikai szabályozási mechanizmusokat kizárva). A vizsgálatban alkalmazott genom-szintű flux balance analysis (FBA) a metabolikus hálózat enzimeinek páronkénti ki-be kapcsolásával a kinetikai és a szabályozási szempontok elhanyagolása ellenére a modell által jósolt pozitív és negatív episztatikus hatásokat is elég jó arányban igazolták az ismert kísérleti eredmények, illetve – megfelelő cutoff-értékek mellett – a kísérletileg kimutatott episztázisok majdnem mindegyike szerepelt a modell által jósoltak között. A pontosság (accuracy) két komponense (precision és recall) nagyon erős trade-off viszonyban állnak, amin gépi tanulós módszerrel lehetett sokat javítani a kísérleti és a modell által jósolt eredmények eltéréseinek, a téves predikciók számának csökkentésével az ismert kísérleti eredményekből kiindulva. Ezzel a módszerrel sikerült pl. az élesztő genomjában az irodalomban tévesen szereplő bakteriális NAD-szintézisutat is kiszűrni.

III. Predicting genome reduction

A disszertációban szereplő második nagyobb téma a radikálisan megváltozó környezetekhez történő alkalmazkodás genom-szintű eseményeinek a jósolhatóságával foglalkozik. Ennek a vizsgálatnak a rendkívül jól megválasztott alanya egy olyan baktériumfaj-pár, amelynek tagjai a szabadon élő *E.coli* és az abból evolvál, endoparazita életmódra áttért *Buchnera aphidicola*. Utóbbi a legtöbb tápanyagot és köztes metabolitot a gazdasejtől kapja, ezért genomjában az *E.coli* közel 4.5 ezer génjéből csak kb. 600-at őrzött meg. A Jelölt és munkatársai megmutatták, hogy az endoparazitikus életmód által redukált genom elemei, sőt a gén-vesztések sorrendje is megdöbbenően megbízhatóan jósolható a kiindulási genomból a flux-balance analysis segítségével a gazdasejt citoplazmájának összetétele ismeretében. A génvesztések szimulált trajektóriái közötti kevés különbség az FBA szerint fitness-szempontról közel neutrális, ami jól magyarázza a *Buchnera*-törzsek között megfigyelhető genom-variációk létezését. Ez a kutatási program is számos szempontból úttörő: egyrészt a metabolikus hálózatok evolúciójának jósolhatósága, másrészt az ezt színező történeti esetlegességek hatásának koherens elemzését teszi lehetővé, empirikus adatokkal szilárdan alátámasztva az elméleti eredményeket.

IV. Underground metabolism and the predictability of adaptive evolution

A két előző témakörhöz metodológiai szorosán kapcsolódó harmadik az *E.coli* metabolizmusát működtető enzimek szubsztát-promiszkuitására vonatkozó empirikus ismeretek szakirodalmi adatai alapján a világon elsőként rekonstruálja egy, a natív metabolikus hálózat mellett működő ún. „underground metabolizmus” topológiáját és tanulmányozza annak evolúciós szerepét. A vizsgált 262 mellékreakció majdnem fele a natív metabolikus hálózat elemeit köti össze, gyakran alternatív reakcióutakat képezve, ami a natív hálózathoz adott redundancián keresztül nyilvánvaló evolúciós előnyt jelenthet. Ez a vizsgálat kiterjed az underground reakciók adaptív evolúciós szerepére is számos alternatív szénforrás felhasználása felé. A szimulációs rendszermodell kísérletekkel igazolt eredményei szerint a

natív metabolikus hálózatba integrált mellékreakciók 15%-a önmagában, és további 5% más mellékreakciókkal együtt növeli a baktérium fitnessét.

V. *Simple paths to complex adaptations*

A metabolikus hálózatok kontextusában komplex adaptációk azok az evolúciósan stabil változások, amelyek a metabolizmust működtető fehérjék génjein egynél több mutáció rögzülésével érhetők el. A Jelölt és munkatársai megmutatták FBA-szimulációkkal elméletileg, illetve *E.coli* kísérleti rendszeren empirikusan is, hogy a metabolikus környezet megfelelő változtatásaival egyes komplex adaptációk viszonylag gyorsan és egyszerűen indukálhatók csak adaptív lépéseken át – lényegesen hatékonyabban, mint a korábban domináns elmélet szerinti neutrális mutációk sorozatával. Ez az eredmény legalább három szempontból jelentős: egyrészt újabb bizonyíték az „irreducibilis komplexitás” fogalmával érvelő „intelligent design” pszeudo-elmélettel szemben, másrészt egyszerű és empirikusan tesztelhető (és tesztelt) elméleti magyarázattal szolgál a metabolikus hálózatokon zajló gyors evolúció mechanizmusára; végül, de nem utolsósorban a disszertáció többi témaköréhez hasonlóan gyakorlati szempontból jelentős biotechnológiai eljárások elvi alapjául szolgálhat.

VII. *Outlook: genome-scale modelling meets machine learning*

A disszertációnak és az annak magját képező publikációknak közös módszertani alapja a flux-balance analysis, ami jelen állapotában a metabolizmus egyes lépéseinek (az azokat katalizáló enzimeknek) a meglétét vagy hiányát tételezi fel, azok enzimkinetikai tulajdonságainak és a génjeik közötti direkt (nem a metabolikus lépéseken keresztül megvalósuló) szabályozási kölcsönhatásoknak a teljes elhanyagolásával. Az a tény, hogy még ezzel a dinamikai felbontással is ennyire jelentős, empirikusan igazolt eredményeket lehetett elérni, önmagában imponáló. Az eddig elhanyagolt dinamikai hatások beépítése a meglévő FBA modell-keretbe a konkrét kinetikai paraméterek és a komplex génregulációs hálózatokra vonatkozó tudás hézagossága miatt azonban egyelőre nem reális cél. Ehelyett a közeljövőben a II. fejezetben már alkalmazott gépi tanulás megfelelő adaptálása lehet alkalmas a predikciók pontosságának a fokozására.

Nyilatkozat:

A fent részletezett (és a disszertáció VI. fejezetében tételesen is felsorolt) eredmények mindegyikét új tudományos eredményként fogadom el, és a doktori művet nyilvános vitára alkalmasnak tartom, a Jelöltnek pedig szívből gratulálok a nemzetközi mezőnyben is kiemelkedő tudományos teljesítményéhez.

Kérdések:

II. fejezet:

Technikai jellegű kérdésem, hogy a pozitív és negatív episztatikus kölcsönhatások pontosságának (accuracy) meghatározásánál az alkalmazott cutoff-értékek mire vonatkoztak? Gondolom, a 2. ábra plotjain az egyes pontok egy-egy cutoff-értékhez tartoznak, de azok konkrét értéke sem derül ki az ábrából (talán lehetett volna szín-kódolni őket).

III. fejezet:

Nyilvánvaló, hogy az endoparazita genomjából hiányzó, a szabadon élő ősből esszenciális metabolikus enzimek termékeit a gazdasejt szolgáltatja a parazita számára. Ez viszont azt teszi szükségessé, hogy a parazita által nem termelt esszenciális metabolitok átjussanak annak sejtmembránján. Vannak-e olyan transzport-fehérjéket kódoló gének a redukált *Buchnera*-genomban, amelyek ezt a funkciót látják el, és az *E.coli* genomjában nincsenek jelen?

Ezzel összefügg a genom-redukció predikciójára vonatkozó eredeti kérdés megfordítása: mennyire kellett az *E.coli* és a *Buchnera* genomok különbségei, vagyis az utóbbiból hiányzó gének, valamint a gazda metabolikus reakcióhálózatának ismeretére támaszkodni a parazita környezeti inputjának (a gazdasejtek által a parazitának szolgáltatott metabolitoknak) az azonosításához?

IV. fejezet:

Lehet-e tudni valamit arról, hogy a natív metabolizmusba nem integrált, pl. egzotikus termékeket produkáló, ill. esetleg a natív hálózattól teljesen izolált mellékreakciók hasznosulnak-e valamely környezetben a baktérium rátermettsége szempontjából? Másrészt: van-e ezeknek már ismert, igazolt szerepe a biotechnológiában?

V. fejezet:

Sikerült-e a vonatkozó cikkek megjelenése óta azonosítani (saját kísérletekkel, vagy irodalmi adatok alapján) más komplex adaptív reakcióutakat, amelyeket változó környezetekkel indukáltak?