

Dr hab. Szymon Kaczanowski

Warszawa 14.02.2019

Recenzja rozprawy doktorskiej
magister Magdaleny Karoliny Biesiadeckiej
pt. „Eksperymentalny test hipotez wyjaśniający zależność między tempem ewolucji
molekularnej i poziomem ekspresji genów”

Magister Magdalena Karolina Biesiadecka przygotowała rozprawę doktorską pod tytułem „Eksperymentalny test hipotez wyjaśniający zależność między tempem ewolucji molekularnej i poziomem ekspresji genów” pod kierunkiem profesora Ryszarda Korony.

Celem przedstawionej rozprawy jest próba wyjaśnienia, dlaczego wysoka ekspresja genów kodujących białka sprzyja wolniejszej ich ewolucji. **Jest to problem fundamentalny i bardzo istotny dla rozwoju nauk biologicznych.**

To zaskakujące zjawisko było wielokrotnie wykazane u różnych gatunków przy pomocy zaawansowanych metod statystycznych. Wykazano, że obserwowany wpływ tempa ewolucji na poziom ekspresji genów jest nawet bardziej istotny, niż wpływ znaczenia genu mierzonego dostosowaniem i jego niezbędności dla życia organizmu.

Warto zwrócić uwagę, że analiza statystyczna tego interesującego zjawiska nie była prosta, ponieważ aktywność transkrypcyjna genów kodujących białko jest bardzo różna. Dlatego istnieją liczne wyjątki od podanej reguły. To znaczy, istnieją białka, których ekspresja jest wysoka, ale tempo ewolucji też jest szybkie. Z drugiej strony istnieją białka, których ekspresja jest niska i tempo ewolucji jest niskie. Jednak dzięki użyciu wyszukanych metod statystycznych, udało się wcześniej wykryć tę zależność, mimo istnienia takich wyjątków.

Jedną z najbardziej przekonujących opisanych w literaturze przedmiotu obserwacji dotyczy porównania par genów paralogicznych, które stosunkowo niedawno uległy duplikacji. Takie pary genów mają bardzo podobną funkcję, a gen, który był ich przodkiem pełnił funkcję obu tych genów w stanie pierwotnym, przed duplikacją. Białka kodowane przez takie pary genów posiadają podobne masy, podobne długości, sekwencje białkowe oraz wiele innych właściwości. Przy porównaniu aktywności transkrypcyjnych takich par genów, wpływ bardzo wiele czynników można więc wykluczyć, co jest istotne w analizie statystycznej. Okazało się jednak, że w większości przypadków w parach genów paralogicznych, gen szybciej ewoluujący ma niższą ekspresję.

Magister Biesiadecka w swoich bardzo pomysłowych badaniach skupiła się właśnie na porównaniu par genów paralogicznych, które niedawno uległy duplikacji. **Doktorantka postawiła hipotezę, że koszt produkcji białka kodowanego przez zduplikowane geny**

szybciej ewoluujące jest wyższy niż koszt produkcji białka kodowanego przez jego wolniej ewoluującego paralogicznego „odpowiednika”.

Postawioną hipotezę doktorantka przetestowała eksperymentalnie. **Warto od razu podkreślić, że ilość pracy, jaką włożyła w wykonanie doświadczeń jest naprawdę imponująca, budzi najwyższy szacunek i jest rzadko spotykana.**

Jako model badawczy użyto drożdży *Sacharomyces cerevisiae*. Jest to bardzo dobrze dobrany model do tego typu badań, ponieważ posiada on około 400 zduplikowanych genów głównie w wyniku duplikacji całego genomu. Istnieją także dobrze opracowane narzędzie molekularne dla tego organizmu. W badaniach użyto wcześniej opracowaną kolekcję plazmidów (tzw. MoRF) umożliwiających wysoką ekspresję wszystkich genów drożdżowych indukowanych przy pomocy promotora galaktozowego. Dzięki zastosowaniu tych plazmidów, doktorantka uzyskała wysoką ekspresję, aż 314 par zduplikowanych genów białek paralogicznych. Dzięki temu było możliwe było eksperymentalne porównanie kosztu wysokiej ekspresji genów w obrębie każdej pary paralogów i sprawdzenie, czy rzeczywiście organizm, w którym jest wysoka ekspresja paralogu szybko ewoluującego ma niższe dostosowanie, niż organizm w którym jest wysoka ekspresja paralogu wolno ewoluującego. Warto nadmienić na marginesie, że doktorantka wykonała badania wstępne, których celem było ustalenie, jakie podłoże genetyczne jest najbardziej właściwe do przeprowadzenia tych badań. Dostosowanie porównano przy pomocy dwóch metod. W pierwszej mierzono maksymalne tempo wzrostu drożdży przy pomocy metody fotometrycznej. W drugiej analizowano wynik konkurencji dwóch typów drożdży, w których zduplikowane geny z danej pary stosunkowo niedawno zduplikowanych paralogów były silnie nadeksprimowane. Wyniki konkurencji zbadano przy pomocy sekwencjonowania metodą NGS (Next Generation Sequencing) nadeksprimowanych genów, co pozwoliło na przybliżone ustalenie ilości drożdży zawierających oba typy plazmidów.

Wykonano także szereg dodatkowych eksperymentów kontrolnych. Porównano wyniki eksperymentów w sytuacji, w której transkrypcja genów z użytych plazmidów jest indukowana i w których nie ma takiej indukcji. Ilościowo oznaczono białka eksprimowane z plasmidu i procentowy ich udział wśród białek drożdżowych. Wykorzystano fakt, że geny w bibliotece MoRF kodują białka połączone z krótkim peptydem, nazywanym metką. Metka nie ma wpływu na właściwości białka i jest używana do oznaczeń ilościowych przy pomocy metody immunologicznej ELISA, które wykonała doktorantka. Zbadano także wpływ różnych pożywek na wynik eksperymentu. Warto tu podkreślić, że część zastosowanych podłoży zawierała czynniki sprzyjające denaturacji białek. Zastosowano takie podłoża ponieważ istnieją obserwacje wskazujące, że wysoka ekspresja białek może prowadzić do postania agregatów zdenaturowanych białek i że często takie agregaty są toksyczne. Przy opracowaniu wyników wzięto pod uwagę liczne parametry analizowanych genów i białek, które dostępne są w literaturze.

W wyniku przeprowadzonych badań doktorantka nie potwierdziła swojej hipotezy. Okazało się, że nadekspresja genów wolno ewoluujących jest bardziej szkodliwa niż nadekspresja genów szybko ewoluujących.

Jest to więc bardzo interesująca obserwacja z punktu widzenia rozwoju nauk biologicznych, ponieważ przyczynia się do wyjaśnienia bardzo istotnego zjawiska jakim jest mechanizm ewolucji genów.

Po za tym liczne analizy wskazały, że koszt „nadprodukcji” białka jest proporcjonalny do energii, jakiej wymaga. Tym samym doktorantka nie znalazła potwierdzenia hipotez dotyczących innych czynników, który mają wpływ na szkodliwy efekt nadprodukcji określonego białka.

Odnosnie przedstawionych wyników eksperymentalnych mam dwie uwagi.

Pierwsza z nich dotyczy szczegółów metodycznych i zastosowania sekwencjonowania fragmentów plazmidów do szacowania wielkości populacji konkurujących drożdży. Jeśli dobrze zrozumiałem z mieszaniny konkurujących drożdży izolowane było plazmidowe DNA i namnażane metodą PCR były sekwencje DNA kodujące geny plazmidów użytych w doświadczeniu. Celem sekwencjonowania było sprawdzenie jaka jest ilość plazmidów kodujących te nadprodukowane geny. Jest to ciekawy pomysł. Zaletą tego pomysłu jest to, że pozwala jednocześnie na sprawdzenie, czy pojawiają się mutacje w genach kodujących nadprodukowane białka. Jednak autorka nie podała informacji na temat takich mutacji. Brakuje także kontroli potwierdzających prawidłowość całej metody. Pokazanie takiej kontroli uczyniłoby wynik bardziej przekonującym. Trzeba jednak zaznaczyć, że cały ten eksperyment jest kontrolą, która potwierdza wyniki uzyskane inną metodą tzn. przy pomocy analizy maksymalnego tempa wzrostu.

Druga moja uwaga jest bardziej ogólna. W rozprawie przyjęto, że miarą znaczenia genu jest liczba środowisk w których delecja danego genu dawała defekt wzrostu. Moim zdaniem uzyskane wyniki będą ciekawsze, jeśli zostaną uzupełnione jednak o informacje na temat znaczenia genu mierzonego utratą dostosowania spowodowanego jego delecją. Tego typu dane można znaleźć w literaturze. Oczywiście, podobne analizy były już wcześniej robione, ale posiadając tego typu informacje w kontekście opisanych eksperymentów można pokusić się o lepszą interpretację wyników.

O ile poziom badań przedstawionych w rozprawie oceniam bardzo wysoko, to mam pewne krytyczne uwagi dotyczące sposobu napisania rozprawy doktorskiej. Przedstawiona do oceny rozprawa doktorska napisana jest w języku polskim i zawiera 126 stron. Rozprawa ma typowy układ i składa się z sześciu rozdziałów: Wstęp, Materiały i Metody, Wyniki, Dyskusja, Literatura i Suplementu. Rozprawa zawiera wszystkie potrzebne elementy: opis hipotezy badawczej, wprowadzenie do tematyki rozprawy doktoratu i szczegółowy opis wykonanych eksperymentów, oraz literaturę.

Moim zdaniem hipoteza badawcza nie jest opisana w sposób wystarczająco jasny. Zabrakło mi tu jasnego wyjaśnienia, że można sobie wyobrazić dwa podstawowe wyjaśnienia badanego efektu:

- 1) Tempo ewolucji genu kodującego białko ma wpływ na poziom jego ekspresji
- 2) Poziom ekspresji genu kodującego białko ma wpływ na tempo jego ewolucji

Wszystkie przeprowadzone badania dotyczą weryfikacji pierwszego hipotetycznego wyjaśnienia. Jaśniejsze wyjaśnienie tutaj hipotezy badawczej, ułatwiłoby czytelnikowi zrozumienie tekstu.

Uważam też, że doktorantka także nie podała w sposób wystarczająco jasny kluczowych dla zrozumienia rozprawy informacji na temat wcześniej wykonanych badań dotyczących wpływu ekspresji na ewolucję w przypadku par zduplikowanych genów paralogicznych. Czytelnik zmuszony jest szukać tych informacji we wskazanej przez doktorantkę literaturze przedmiotu.

W rozprawie pojawiały się też błędy literowe. Część z nich nie miała poważnego znaczenia np. na stronie 25 brakuje końcówki i litery m w słowie zmianom, albo na stronie 85 zamiast słowa często jest słowo częstya. Problem jednak poważniejszy sprawia nagminne umieszczanie przez doktorantkę dodatniego wykładnika wartości P (ang. P-value), czyli prawdopodobieństwa przypadkowego pojawienia się danego wydarzenia. Na przykład na 64 stronie wartość ta wacha się od 10^5 do 10^4 , a prawdopodobnie powinno być 10^{-5} i 10^{-4} . Ten literowy błąd utrudnia czytelnikowi zrozumienie tekstu. Czytelnik musi zgadywać, że został popełniony błąd literowy i że autorka ma na myśli rzeczywiście prawdopodobieństwo określone angielskim terminem P-value.

Natomiast literatura jest starannie dobrana.

Podsumowując redakcja i sposób napisania samej rozprawy zawiera pewne drobne niedociągnięcia, które jednak podaję z recenzenckiego obowiązku. Opisane w rozprawie badania są niezwykle ciekawe, dotyczą bardzo istotnego zjawiska z punktu widzenia rozwoju nauki. Doktorantka włożyła olbrzymią ilość pracy w wykonanie badań.

Stwierdzam więc, że oceniana rozprawa doktorska spełnia wszystkie wymagania zwyczajowe oraz stawiane przez ustawodawcę. Wnoszę o dopuszczenie doktorantki do dalszych etapów obrony doktoratu.

Szymon Kuczanowski