



UNIwersytet
Warszawski

Wydział Biologii
Instytut Genetyki i Biotechnologii
prof. dr hab. Paweł Golik



Warszawa, 26.09.2023

Recenzja rozprawy doktorskiej mgr Piotra Karasia pt. „Molekularne podstawy ewolucji małych białek szoku cieplnego w *Erwiniaceae*”

Duplikacje genów są jednym z kluczowych mechanizmów ewolucji genomów, prowadzącym do zwiększenia repertuaru kodowanych białek. Neofunkcjonalizacja lub subfunkcjonalizacja powstałych w wyniku duplikacji paralogów prowadzi do zwiększenia liczby różnych kodowanych przez genom funkcji. Uproszczeniem jest jednak stwierdzenie, że jest to proces w każdym przypadku związany z podlegającym dodatniemu doborowi zwiększeniem dostosowania. W myśl sformułowanej przez Stoltzfusa, Lyncha i innych koncepcji nieadaptacyjnego wzrostu złożoności (teoria konstruktywnej ewolucji neutralnej), zastąpienie pojedynczego białka przez układ złożony z dwóch lub więcej nie musi być związane z przewagą selekcyjną. Od czasu sformułowania tej teorii, kwestie znaczenia ewolucyjnego procesów duplikacji i subfunkcjonalizacji są przedmiotem ożywionej dyskusji, głównie jednak ograniczonej do rozważań teoretycznych. Wciąż niewiele jest danych eksperymentalnych porównujących funkcjonowanie pełniących podobną funkcję układów jedno- i dwubiałkowych.

Oceniana praca doktorska stanowi udaną próbę przeprowadzenia takiej właśnie analizy, wykorzystującej bardzo interesującą sytuację związaną z małymi białkami szoku cieplnego (sHsp) u bakterii z rzędu *Enterobacterales*. W grupie tej doszło do duplikacji i subfunkcjonalizacji białek sHsp, w wyniku czego powstał system złożony z dwóch białek IbpA i IbpB. Co jeszcze bardziej ciekawe, w jednej z rodzin *Enterobacterales* – *Erwiniaceae* sytuacja uległa odwróceniu: jeden z dwóch paralogów (IbpB) został utracony, a całość funkcji ponownie przejęło białko IbpA. Jest to bardzo interesujące zjawisko, gdyż w większości teoretycznych rozważań przyrost złożoności poprzez subfunkcjonalizację paralogów uznawany jest za generalnie nieodwracalny. Bliższe przyjrzenie się ewolucji dwu- i jednobiałkowych systemów sHsp *Enterobacterales*, którego w recenzowanej pracy podjął się Kandydat spełnia zatem najważniejsze moim zdaniem kryterium – podejmuje się rozwiązania naprawdę istotnego poznawczo problemu naukowego. Problematyka badawcza została zatem wybrana właściwie i dotyczy zagadnienia istotnego od strony poznawczej.

W analizie ewolucji wtórnie jednobiałkowego systemu sHsp *Erwiniaceae* Kandydat przeprowadził szereg eksperymentów i analiz teoretycznych, wykazując się opanowaniem bardzo rozległego wachlarza metod z zakresu bioinformatyki, biologii molekularnej i biofizyki. Bogactwo różnych podejść metodycznych jest jedną z istotnych zalet omawianej pracy i świadczy o dużej dojrzałości Kandydata.

Praca ma formę tradycyjnej rozprawy zawierającej wszystkie niezbędne elementy: obszerny wstęp teoretyczny, szczegółowo opisane metody i materiały w zakresie wystarczającym dla zweryfikowania poprawności technicznej i ewentualnej replikacji wyników, same wyniki opisane dogłębnie i opracowane graficznie, głównie w postaci wykresów podsumowujących dane ilościowe, a także analizę wyników w zwięzłej dyskusji. Praca prawidłowo cytuje właściwą literaturę i od strony formalnej spełnia wszystkie wymagania stawiane tego typu tekstom naukowym. Tytuł pracy jest precyzyjnie sformułowany i dobrze oddaje istotę jej treści.

Jak zawsze w przypadku nietrywialnych prac naukowych, w trakcie jej lektury nasuwają się pewne pytania, które bardziej szczegółowo przedstawiam poniżej. Pytania takie nieodłącznie towarzyszą dyskusjom nad nowymi wynikami badawczymi i w najmniejszym stopniu nie podważają ogólnie bardzo pozytywnej oceny wartości naukowej i warsztatu ocenianej rozprawy.

W pierwszym etapie badań Kandydat porównał aktywności współczesnych białek sHsp z systemu jednobiałkowego (*Erwiniaceae*) i dwubiałkowego (*E. coli* jako przedstawiciel *Enterobacteriaceae*). Zgodnie z przewidywaniami, białka z systemu dwubiałkowego nie są w stanie wydajnie działać pojedynczo w nieobecności partnera. Do tej części mam tylko jedną drobną uwagę: sformułowania takie, jak „[...] na poziomie porównywalnym z kontrolą” (s. 48) powinny być wsparte odpowiednią analizą statystyczną, wykazującą istotność różnicy lub jej brak.

W kolejnym etapie przeprowadzona została analiza filogenetyczna białek IbpA *Enterobacteriales*. Zasadniczym wynikiem jest tu zgodność z drzewem gatunków, co wyklucza udział zjawisk takich, jak poziomy transfer genów. Ważnym produktem tej analizy jest też uliniowanie sekwencji aminokwasowych i nukleotydowych IbpA, wykorzystywane następnie w dalszych analizach. W tekście rozprawy brakuje mi podsumowania wyników tego uliniowania, a zwłaszcza informacji o tym, jak silna jest konserwacja analizowanych sekwencji ortologicznych na poziomie aminokwasowym i nukleotydowym, co ma duże znaczenie dla interpretacji dalszych rezultatów.

Kolejny etap analizy *in silico*, niezwykle interesujący i ważny dla dalszych etapów pracy, to rekonstrukcja ancestralnych sekwencji kluczowych dla ewolucji badanego systemu: ostatniego wspólnego przodka siostrzanych kładów z systemem dwubiałkowym (*Enterobacteriaceae*) i jednobiałkowym (*Erwiniaceae*) – działającego wciąż w systemie dwubiałkowym (AncA0), oraz ostatniego wspólnego przodka jednobiałkowych sHsp z *Erwiniaceae*, czyli najstarszego białka

systemu wtórnie jednobiałkowego (AncA1). Białka o zrekonstruowanej sekwencji zostały następnie wyprodukowane i oczyszczone przy zastosowaniu technik biologii syntetycznej, co umożliwiło badanie ich aktywności za pomocą szeregu oznaczeń. Kandydat wykazał tu, że właściwości zrekonstruowanych białek w oznaczeniu deoligomeryzacji w warunkach szoku cieplnego są porównywalne z białkami współczesnymi, a pod kątem wiązania agregatów białkowych i sekwestracji substratu lokują się pomiędzy współczesnymi białkami systemów jedno- i dwubiałkowych.

Kluczowym dla tych analiz eksperymentem było porównanie zdolności rekonstruowanych i współczesnych białek do przywracania właściwej konformacji białek w teście reaktywacji lucyferazy. Zgodnie z wstępnymi wynikami, współczesne białka systemu jedno- i dwuskładnikowego (w komplecie) wykazują podobnie wysoką aktywność, natomiast rekonstruowany ostatni wspólny przodek systemu wtórnie jednobiałkowego (AncA1) ma aktywność wykrywalną, lecz zauważalnie niższą od białek współczesnych. Ponownie, brakuje tu statystycznej oceny istotności obserwowanych różnic w aktywności. W analizie, której wyniki przedstawiono na Ryc. 13 brakuje mi też porównania aktywności AncA1 z białkiem IbpA *E. coli* działającym pojedynczo. Uzyskana aktywność lucyferazy na poziomie ~40% na pierwszy rzut oka może przypominać to, co uzyskano dla pojedynczego IbpA z *E. coli* w eksperymencie przedstawionym na Ryc. 8A, nie wiem jednak, czy wyniki obu eksperymentów można porównywać. Jak wygląda różnica między aktywnością AncA1 a IbpA z *E. coli* bez IbpB? Wątpliwości nie budzi natomiast to, że ancestralne białko systemów dwubiałkowych (AncA0) pojedynczo nie jest w tym oznaczeniu w jakimkolwiek stopniu aktywne. W kontekście ewolucji współdziałania między białkami sHsp ciekawe (i stosunkowo łatwe) byłoby sprawdzenie, czy ancestralne białko IbpA systemu dwubiałkowego (AncA0) współpracuje ze współczesnym IbpB (np. z *E. coli*).

Ponieważ odtwarzanie sekwencji ancestralnych jest zawsze obciążone niepewnością, Kandydat przebadał też aktywność alternatywnych form ancestralnych A0 i A1, wykazując, że wspomniana niepewność nie wpływa na zasadnicze konkluzje doświadczenia. W wynikach przedstawionych na Ryc. 15 i 16 brakuje mi jedynie porównania alternatywnych form A0Alt i A1Alt z zasadniczymi formami A0 i A1 – na Ryc. 15 są porównane wyłącznie z białkami współczesnymi, a na Ryc. 16 brak jakiegokolwiek kontroli do porównań. Ciekawy wynik przedstawiono na Ryc. 17, gdzie alternatywne białko A1Alt działa dopiero w podwyższonym stężeniu Hsp70. Znowu nasuwa się pytanie, jak w warunkach zwiększonego stężenia Hsp70 zachowuje się pojedyncze IbpA z *E. coli* – wszystkie wcześniejsze oznaczenia (np. z Ryc. 8) prowadzone były jedynie w standardowym stężeniu. Jak się ma aktywność A1Alt do aktywności IbpA z *E. coli* pojedynczo?

Kolejnym, bardzo ważnym i pomysłowym etapem badania ewolucji białek sHsp było określenie z wykorzystaniem analiz *in silico* (program *codeml* z pakietu PAML) czy za różnice między białkami IbpA systemów jedno- i dwubiałkowych odpowiadał dobór naturalny, a następnie identyfikacja kluczowych dla ewolucji tych białek pozycji aminokwasowych. W interpretacji tych analiz brakuje mi nieco wspomnianych już wcześniej danych dotyczących stopnia podobieństwa porównywanych sekwencji. Analizy tego typu (wykorzystujące parametr ω) bywają bowiem wrażliwe na wysycenie mutacjami synonimicznymi w słabiej konserwowanych sekwencjach. Dogłębne analizy biochemiczne z wykorzystaniem białek oczyszczonych dzięki ekspresji syntetycznych sekwencji wykazały, że dla ewolucji zdolności reaktywacji zdenaturowanych białek w nieobecności białka IbpB wystarczają dwa podstawienia w pozycjach 66 i 109. Nasuwa się tu jedno pytanie: program *codeml* przy analizie modeli typu branch-site (stosowanych przez Kandydata) przewiduje, które pozycje w sekwencji podlegają istotnemu działaniu doboru dodatniego. Czy miejsca wskazane przez *codeml* jako podlegające doborowi z najwyższym prawdopodobieństwem *a posteriori* odpowiadały pozycjom zidentyfikowanym w sekcji 7.7 w analizach biochemicznych?

Zwieńczeniem badań było zbadanie znaczenia zidentyfikowanych podstawionych pozycji dla aktywności białka i odniesienie ich do modelu struktury przestrzennej białka. Podobnie jak poprzednie, ta część badań także imponuje szerokim wachlarzem wykorzystanych narzędzi teoretycznych i doświadczalnych.

Dyskusja podsumowująca wyniki jest zwięzła i bardzo konkretna. Osobiście chętnie poznałbym poglądy Kandydata na znaczenie wyników jego badań w szerszym, wspomnianym we wstępnym akapicie niniejszej recenzji kontekście teoretycznych rozważań o adaptacyjnym i nieadaptacyjnym przyroście złożoności w ewolucji, nie jest to jednak w żadnym razie zarzut wobec samej pracy.

W podsumowaniu stwierdzam, że przedstawiona mi do recenzji praca doktorska Pana mgr Piotra Karasia pt. „Molekularne podstawy ewolucji małych białek szoku cieplnego w *Erwiniaceae*” stanowi oryginalne rozwiązanie istotnego problemu naukowego i potwierdza ogólną wiedzę teoretyczną Kandydata oraz umiejętność samodzielnego prowadzenia pracy naukowej. Do rozprawy doktorskiej dołączone zostały streszczenia w języku polskim i angielskim. Na tej podstawie stwierdzam, że wszystkie wymagania prawne zapisane w Ustawie z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. 2018 poz. 1668 z późn. zm.) stawiane rozprawom doktorskim i Kandydatom ubiegającym się o nadania stopnia doktora zostały spełnione i może stanowić podstawę do nadania stopnia doktora. Zwracam się zatem do Rady Dyscypliny Biotechnologia Uniwersytetu Gdańskiego o dopuszczenie mgr Piotra Karasia do dalszych etapów postępowania o nadanie stopnia doktora w dziedzinie nauk ścisłych i przyrodniczych w dyscyplinie biotechnologia.

prof. dr hab. Paweł Golik