

## Recenzja rozprawy doktorskiej

pana mgra Piotra Sławomira Karasia

zatytułowanej

„Molekularne podstawy ewolucji małych białek szoku cieplnego w *Erwiniaceae*”

Przedstawiona do recenzji rozprawa doktorska została wykonana pod kierunkiem promotora prof. dra hab. Krzysztofa Liberka i promotora pomocniczego dra inż. Bartłomieja Tomiczka na Międzyuczelnianym Wydziale Biotechnologii Uniwersytetu Gdańskiego i Gdańskiego Uniwersytetu Medycznego a finansowana była przez Narodowe Centrum Nauki w ramach grantu OPUS 17 2019/33/B/NZ1/00352.

### Problem badawczy

Cel pracy jest bardzo ściśle zdefiniowany jako próba zrozumienia jak przebiegła droga i jakie mechanizmy dominowały w ewolucji małych białek szoku cieplnego (sHsp) w bakteriach rodziny *Erwiniaceae*. Wybrano ten obiekt badań, ponieważ w *Erwiniaceae* występuje tylko jedno małe białko szoku cieplnego, lbpA, podczas gdy w siostrzanej rodzinie *Enterobacteriaceae* jest zarówno ortolog białka lbpA jak i jego paralog lbpB. Oznacza to utratę lbpB w *Erwiniaceae* i powstanie systemu wtórnie jednobiałkowego z ancestralnego systemu dwubiałkowego przy, jak się wydaje, dobrym zachowaniu właściwej mu funkcji. Już to jest samo w sobie ciekawe, ponieważ jesteśmy przyzwyczajeni do myślenia o ewolucji jako generalnie zwiększającej liczbę genów poprzez ich amplifikowanie i specjalizację powstałych paralogów, natomiast zmniejszenie ich liczby wydaje się nam możliwe wtedy, gdy ich rola staje się zbędna. Badanie zjawisk innych niż tych występujących najczęściej wcale nie jest działalnością o drugorzędym znaczeniu, ponieważ dostarcza okazji do rewizji powszechnych opinii i tym samym może pomóc w

zrozumieniu różnych wzorców w ewolucji, także tych, które uważane są za oczywiste, a nawet jedynie możliwe, tylko dlatego, że napotykamy je najczęściej. Dlatego temat tej pracy jest dla biologa ewolucyjnego niewątpliwie ważny. Jego wybór jest też zapewne zrozumiały dla biochemików i biologów komórki, ponieważ dotyczy białek, które są stosunkowo mniej poznane niż inne systemy białek opiekuńczych, a które składają się z molekuł większych, często współpracujących w zespołach dużych i funkcjonalnie zróżnicowanych. Fakt, że stosunkowo proste białka sHsp z ich prawdopodobnie stosunkowo prostą i raczej uniwersalną w różnych organizmach funkcją pozostają niedostatecznie rozpoznane jest oczywistym wyzwaniem dla badaczy poszukujących tematów wartych niezwłocznego zwrócenia uwagi. Dlatego wybrane pytanie badawcze i konsekwentne dążenie do znalezienia odpowiedzi na nie, o czym piszę dalej, uważam za jak najbardziej właściwe dla przygotowania wartościowej rozprawy doktorskiej.

#### Forma rozprawy

Praca ma format maszynopisu przygotowanego w języku polskim o ogólnym kształcie typowym dla artykułów naukowych. Zawiera wprowadzenie do zagadnienia, następnie przedstawia stosowane metody i uzyskane wyniki, a kończy się dyskusją uzyskanych rezultatów i ich porównaniem z wynikami innych badaczy. Poszczególne części wydają się być odpowiednio obszerne, czyli dostarczają treści w ilościach dostatecznie kompletnych a przy tym nie nadmiernych. Tak na przykład Wstęp bardzo skrótowo traktuje główne znane systemy białek opiekuńczych z wyjątkiem Hsp40/70, który to system współpracuje z białkami sHsp. Same małe białka szoku cieplnego są dobrze scharakteryzowane jako klasa występująca w bardzo różnych organizmach, a szczególnie wyczerpująco potraktowane są białka rodzin *Erwiniaceae* i *Enterobacteriaceae*. Dotyczy to aspektów filogenezy, struktury i roli komórkowej. Wszystko to sprawia, że czytelnik jest dobrze przygotowany do oceny wagi zagadnienia i zrozumienia celów zamierzonych badań. Opisy metody empirycznych są stosunkowo szczegółowe, pozwoliłyby na odtworzenie zastosowanych procedur, ale autorowi udało się uniknąć sformułowania ich jako nużących protokołów laboratoryjnych. Skromniejsze jest przedstawienie technik bioinformatycznych i obliczeniowych, ale jest to praktyka normalna, ponieważ stosowane narzędzia są najczęściej standardowymi programami opracowanymi przez specjalistów i zaakceptowanymi przez wielu badaczy do stosowania przy określonych zadaniach. Wyniki są przedstawione klarownie, zawierają odrobinę narracji przewodniej, co jest pomocne bo praca podaje sprawozdania z wielu eksperymentów a niektóre z nich podejmowano ze względu na wyniki uzyskane w eksperymentach poprzednich. Ryciny są czytelne, estetyczne i poprawne (wyjątkiem jest Ryc. 23 z błędnymi identyfikatorami podstawień aminokwasów). Dyskusja jest stosunkowo krótkim tekstem bez wydzielonych części, co jest jak najbardziej akceptowalne.

Przy tej ogólnie pozytywnej ocenie formy rozprawy, muszę zwrócić uwagę na pewne niedociągnięcia. Chodzi mi o organizację tekstu w akapity. Po pierwsze, są one wyróżniane zarówno przez wcięcie jak i poszerzoną interlinię, co jest zbędne a nawet nieprawidłowe. Poważniejszą sprawą jest ogromna liczba akapitów bardzo krótkich, często zawierających tylko jedno zdanie. Typowy akapit powinien zawierać zapowiedź jego tematu, możliwie pełną treść do niego się odnoszącą i podsumowanie, które pozwoliłoby czytelnikowi sprawdzić czy go zrozumiał albo uświadomić sobie czego jeszcze nie rozumie i na razie musi przyjąć zapewnienia autora. Artykuły naukowe są najczęściej trudne lub bardzo trudne i dlatego ich właściwa organizacja jest sprawą o dużej wadze. W dalszych częściach tekstu akapity się wydłużają, choć nie zawsze mają właściwą strukturę. Doktorant powinien popracować nad tym aspektem w czasie pisania kolejnych tekstów naukowych.

Podsumowując, przedstawiona praca jest typowym i zasadniczo dobrym maszynopisem przygotowanym w całości przez samego doktoranta, w którym opisuje on wykonane przez siebie prace i jednoznacznie wskazuje, które elementy eksperymentów były wniesione przez inne osoby.

#### Ocena dokonania naukowego

W przedstawionych badaniach można wydzielić kilka głównych etapów, z których pierwszym była analiza filogenetyczna. Wcześniejsze badania jednoznacznie wskazywały, że wspólny przodek *Enterobacteriaceae* i *Erwiniaceae* (AncA<sub>0</sub>) musiał mieć zarówno lbpA jak i lbpB, z których to drugie białko zostało utracone we wspólnym przodku wszystkich *Erwiniaceae* (AncA<sub>1</sub>). Dlatego Doktorant uznał, że ewolucja pomiędzy sekwencją (AncA<sub>0</sub>) a (AncA<sub>1</sub>) musiała być odpowiedzią na utratę lbpB, lub przygotowaniem do takiej utraty, tak by AncA<sub>1</sub> był już funkcjonalnie sprawnym systemem sHsp wtórnie jednobiałkowym. Ta hipoteza oparta jest na nie wypowiedzianym założeniu, że we wczesnej historii *Erwiniaceae* nie było okresu, w którym brak lub niedobór funkcji sHsp był w przybliżeniu selekcyjnie neutralny. Rozumiem, że takie założenie jest na tyle bezpieczne, że nie potrzebuje rozpatrywania. Doktorant najpierw zbudował drzewo filogenetyczne lbpA dla *Enterobacteriales* i znalazł, że jest ono bardzo zbliżone do drzew otrzymywanych dla innych białek co jest okolicznością pozytywną, jako że potwierdza zasadniczo klonalną historię tej grupy i tym samym wspiera wiarygodność głównego przedsięwzięcia tych badań, czyli zaproponowania sekwencji AncA<sub>0</sub> i AncA<sub>1</sub>. Posługując się istniejącymi narzędziami, Doktorant takie sekwencje podał i znalazł 10 podstawień aminokwasowych pomiędzy nimi. Dla porównania, już tylko dwaj wybrani współcześni przedstawiciele, jeden z *Enterobacteriaceae* i jeden z *Erwiniaceae*, różnią się na 30 pozycjach. Tak wydatne zawężenie pola poszukiwań zostało przez autora pracy jak najśluszej uznane za sukces tego etapu badań.

Kolejny krok miał miejsce w laboratorium i polegał na pozyskaniu sztucznie zsyntetyzowanych genów kodujących AncA<sub>0</sub> i AncA<sub>1</sub>, ich nadekspresji, oczyszczeniu i poddaniu badaniom w warunkach *in vitro*. Do porównania wzięto współczesne lbpA z *E. coli* i *E. amylovora*. Pomiarów obejmowały zdolność do odwracalnej oligomeryzacji, wiązanie zagregowanego substratu, sekwestracja zagregowanej lucyferazy przez same sHsps i zdolność do reaktywacji lucyferazy przy współdziałaniu z systemem Hsp70. Porównania pomiędzy białkami ancestralnymi i porównania względem białek współczesnych doprowadziły autora do konkluzji, że ewoluujące wtórnie pojedyncze sHsp nabyło wyższej zdolności do reaktywacji uszkodzonych białek przy udziale Hsp70 a towarzyszyło temu zmniejszenie siły sekwestracji agregatów przez samo sHsp. To potwierdzałoby istnienie kompromisu funkcjonalnego, w którym silne oddziaływanie sHsp z substratem przeszkadzałoby w dopuszczeniu do działania Hsp70. To ważne dla tej pracy stwierdzenie zostało dodatkowo potwierdzone poprzez analogiczne analizy dla sekwencji ancestralnych posiadających podstawienia alternatywne w miejscach, gdzie rekonstrukcja była najmniej wiarygodna. Bardzo podoba mi się ta ostrożność, bo autor narażał się na podważenie korzystnych dla siebie wyników uzyskanych dla sekwencji wcześniejszych. Całość tego etapu badań uważam za dobrze zaprojektowaną, przeprowadzoną i zinterpretowaną. Nie rozumiem w zasadzie tylko jednego stwierdzenia dotyczącego braku zdolności AncA<sub>0</sub> do stymulowania aktywności lucyferazy (Ryc. 13). Autor tłumaczy to „względnie silnym oddziaływaniem AncA<sub>0</sub> z substratem”. Tymczasem, AncA<sub>1</sub> radzi sobie dość dobrze z reaktywacją, a do substratu ma powinowactwo podobne do AncA<sub>0</sub> (Ryc. 11). Owszem, AncA<sub>0</sub> silniej sekwestruje (Ryc. 12). Nie będąc biochemikiem, nie potrafię powiedzieć, który z obu pomiarów siły oddziaływania z substratem należy tutaj uznać za bardziej miarodajny, a nawet domyśleć się po co stosowano oba. Liczę na wyjaśnienia Doktoranta.

Autor następnie wraca do analiz obliczeniowych by wykazać istnienie selekcji naturalnej pomiędzy węzłami AncA<sub>0</sub> a AncA<sub>1</sub>. Zrobił to poprzez obliczenie stosunku substytucji niesynonimowych do synonimowych i sprawdzeniu statystycznej wiarygodności wykrytego pozytywnego odstępstwa. Przyjmuję podany przez niego wynik, przy dwóch komentarzach. Po pierwsze, wnioskowanie o wystąpieniu selekcji pozytywnej byłoby wzmocnione, gdyby podobne analizy w innych odcinkach ewolucji dawały wyniki odmienne. Po drugie, i w sekcji 7.6 i w Dyskusji (str. 88), autor uznaje ten wynik za wskazujący, że przejście od AncA<sub>0</sub> do AncA<sub>1</sub> było wynikiem adaptacji i nie mogło być wynikiem dryfu. Wśród biologów ewolucyjnych jest jednak wielu takich, którzy ślady selekcji naturalnej w rejonie kodującym całe białko lub jego domenę wiążą z wystąpieniem intensywnej kompensacji. Najpierw dryf utrwałaby zmianę niekorzystną, którą następnie selekcja łagodziłaby w wielu nawet krokach częściowych kompensacji, bo pojedyncza kompensacja jest często trudna, a rezultatem byłby lokalny duży nadmiar podstawień dający istotny statystycznie sygnał. Oczywiście nie mam żadnych argumentów za

tym scenariuszem, podaję go tylko by wskazać, że wykrycie doboru pozytywnego objawiającego się jako nadmiar postawień niesynonimowych jest trudne do interpretacji.

Kolejnym zadaniem jakie sobie postawił doktorant było identyfikowanie kluczowych podstawień odpowiadających za różnice w aktywności AncA<sub>0</sub> i AncA<sub>1</sub>. Skupił się na siedmiu podstawieniach wprowadzając do AncA<sub>0</sub> warianty występujące w AncA<sub>1</sub>. Wynikowe białko aktywowało lucyferazę tak jak AncA<sub>1</sub>. Następnie wyłączał po jednym podstawieniu, czyli badał warianty z sześcioma podstawieniami. Dwa z nich wykazywały mniejszą aktywność, sugerując ich znaczącą rolę. Następnie Doktorant wykazał przekonująco, że te dwa podstawienia są istotnie krytyczne i działają synergistycznie, a przy tym tylko wtedy, gdy są wprowadzone w tło AncA<sub>0</sub>. Udowodnił następnie, że te dwa podstawienia wystarczają do wytłumaczenia wzorców oligomeryzacji i siły sekwestracji stwierdzonych dla wariantu z siedmioma podstawieniami. Pozwoliło mu to na powiązanie wspomnianego wcześniej kompromisu funkcjonalności sHsp pojedynczego i we współpracy z Hsp70 z tymi właśnie dwoma pozycjami. Ten ciąg eksperymentów można jedynie pochwalić.

Doktorant nie poprzestał na tym i w kolejnych eksperymentach zabrał się za badania roli tych dwóch pozycji w oddziaływaniach na poziomie molekularnym. Wykazał, że występując w domenie ACD, te dwie pozycje są ważne dla oddziaływań z rejonem C białka IbpA, a takie oddziaływania mogłyby być kluczowe dla jego odwracalnej oligomeryzacji. Autor pisze, że eksperymenty wykazują „nieznaczne zmniejszenie tworzonych oligomerów”. Ja nie potrafię tego stwierdzenia, nawet z pomocą Ryc. 27, zinterpretować ani jako potwierdzenia ani odrzucenia tej hipotezy. Podobnie nie jestem pewien, czy domysł autora, iż domena ACD wiąże się z amorficznymi agregatami poprzez kieszeń między kartkami  $\beta$ 4 i  $\beta$ 8 został zanegowany przez wykrycie braku sekwestracji substratu przez samą domenę ACD. Brakuje mi doświadczenia w takich interpretacjach i liczę na uproszczone wyjaśnienia. Tym niemniej, pozostaję przekonany, że Doktorant rzeczywiście wykrył kluczową rolę obu pozycji i dostarczył ważnych danych o ich roli w oddziaływaniach ze strukturami samych sHsps i ich substratów. Dostrzegam też to, że zakończył swoje analizy poprzez zbadanie roli obu pozycji w białkach współczesnych pokazując, że zachowują one swoją ważną rolę a tym samym najprawdopodobniej pozostawały kluczowe w całej ewolucji sHsps, przynajmniej w *Enterobacteriales*.

Dyskusja obejmuje wszystkie ważniejsze wyniki sprawozdane wcześniej przez Doktoranta. Jej najważniejszym motywem jest przegląd argumentów za kompromisem ewolucyjnym uniemożliwiającym białkom sHsps równoczesny rozwój zarówno silnego powinowactwa do substratu jak i efektywnej współpracy z Hsp70. Ciekawe i nie budzące sprzeciwu są rozważania o przyczynach i przebiegu ewolucji IbpA w kontekście utraty IbpB. Muszę jednak zwrócić uwagę na pewien element wywodów autora, który chyba powinien być przez niego zrewidowany. Na stronie 89, doktorant dyskutuje niską efektywność AncA<sub>0</sub> w

aktywacji lucyferazy i wnioskuje, że została ona potem poprawiona w toku ewolucji przez dodatkowe podstawienia w pozycjach innych niż w tych dwóch kluczowych. Wcześniej, na stronie 58, postuluje, że podobna adaptacja pomogła białku AncA<sub>1</sub> zniwelować jego słabą zdolność do sekwestracji. Ja nie odczuwam potrzeby takich tłumaczeń, bo oba białka nazwałbym raczej hAncA<sub>0</sub> i hAncA<sub>1</sub>, gdzie h oznacza *hypothetical*. Autor nie podaje prawdopodobieństw właściwych poszczególnym najbardziej wiarygodnym aminokwasom w rekonstruowanych białkach, ale podejrzewam, że ich iloczyn, czyli prawdopodobieństwo otrzymania ściśle poprawnej sekwencji byłby raczej niski. Dlatego uważam, że faktycznie istniejące dawne białka AncA<sub>0</sub> i AncA<sub>1</sub> najprawdopodobniej już miały podstawienia poprawiające ich pracę w porównaniu do postulowanych hAncA<sub>0</sub> i hAncA<sub>1</sub>. Rekonstrukcje ewolucji nie mają przecież prowadzić do jej odtworzenia w detalach, ale do uchwycenia jej kluczowych elementów, które potem można zweryfikować, co Doktorantowi znakomicie się udało.

#### Wniosek o wyróżnienie

Doktorant zidentyfikował i rozwiązał ważny problem badawczy postępując się szerokim wachlarzem metod obliczeniowych i laboratoryjnych. Jego praca pokazuje w jaki sposób nastąpić mogła ewolucja wtórnie jednobiałkowego systemu sHsp w rodzinie *Erwiniaceae* i dodała istotne wiadomości o mechanizmach działania tego typu białek.

#### Wniosek końcowy

W posumowaniu stwierdzam, że przedłożona do recenzji rozprawa pana mgra Piotra Sławomira Karasia spełnia wymogi stawiane pracom doktorskim opisane w ustawie z dnia 20 lipca 2018 roku i wnoszę do Rady Naukowej Dyscypliny Biotechnologia MWB UG i GUM o dopuszczeniu doktoranta do dalszych etapów tego przewodu.

Kraków 2 września 2023



Ryszard Korona