

Poznań, 22.07.2020

Dr hab. Marta Olejniczak, prof. ICHB PAN

Zakład Inżynierii Genomowej

Recenzja rozprawy doktorskiej mgr Magdaleny Reginy Kubiak zatytułowanej

**„Udział retrokopii genów kodujących białka w kształtowaniu
genomu, transkryptomu i proteomu człowieka”**

Przedstawiona do oceny rozprawa doktorska została wykonana w Instytucie Biologii i Ewolucji Człowieka Wydziału Biologii Uniwersytetu im. Adama Mickiewicza w Poznaniu pod kierunkiem prof. dr hab. Izabeli Makałowskiej.

Rozprawa doktorska obejmuje zbiór pięciu spójnych tematycznie artykułów naukowych opublikowanych w latach 2015-2020, z których trzy to prace eksperymentalne a dwie przeglądowe. Na uwagę zasługuje fakt, że prace są opublikowane w dobrych czasopismach naukowych, a w dwóch artykułach Pani mgr Magdalena Kubiak jest pierwszą autorką. Tematyka badawcza rozprawy doktorskiej wpisuje się w zagadnienia realizowane z dużym powodzeniem w zespole Pani Promotor, dotyczące zjawiska retrotranspozycji i jej funkcji ewolucyjnej. Rozprawa doktorska składa się ze Spisu treści, listy publikacji, streszczenia, krótkiego 10 stronicowego wprowadzenia, oświadczenia Doktorantki i współautorów prac dotyczące ich udziału w poszczególnych publikacjach oraz kopii publikacji. Choć opracowanie zawiera zasadniczo wszystkie informacje niezbędne w rozprawie doktorskiej, to wyodrębnienie w rozdziale „Wprowadzenie” części: wstępu, celu badań, opisu badań i dyskusji poprawiłoby jasność opisu wyników pochodzących ze zbioru aż pięciu artykułów naukowych.

Mimo że od poznania sekwencji genomu człowieka minęło już 20 lat wciąż skrywa on wiele tajemnic. Naukowcy próbują zrozumieć znaczenie i pochodzenie wielu sekwencji, w tym retrokopii, uważanych kiedyś za „śmieciowy” DNA. Dziś wiadomo, że przynajmniej część z nich to funkcjonalne kopie genów rodzicielskich, niezwykle istotne z punktu widzenia ewolucji genomów. Głównym celem

badania realizowanych w ramach rozprawy doktorskiej była analiza wpływu retrokopii genów kodujących białka na genom, transkryptom i proteom człowieka. Cele szczegółowe obejmowały (i) analizę wpływu retrotranspozycji genów na kształtowanie różnic międzygatunkowych, (ii) badanie zmienności retroduplikacji w zróżnicowanych geograficznie populacjach ludzkich oraz (iii) funkcji pełnionych przez retrokopie w genomie człowieka.

Choć retrogeny posiadają pewne cechy charakterystyczne, takie jak brak intronów, obecność ogona poli-A i powtórzeń oskrzydlających region insercji, ich identyfikacja w genomach i potwierdzenie ekspresji nie jest łatwe. Poprzez porównanie sekwencji znanych genów zespół pod kierunkiem prof. Makałowskiej zidentyfikował ok. 90 tysięcy retrokopii w genomach 99 gatunków i stworzył bazę danych RetrogeneDB. Co ciekawe, ok. 30% zidentyfikowanych retrokopii to potencjalnie funkcjonalne retrogeny. W ramach dalszych badań autorzy potwierdzali ekspresję retrokopii kilkoma metodami: poprzez analizę dostępnych danych z baz RNAseq; mapowanie sekwencji EST (ang. Expressed Sequence Tags) do genomów referencyjnych, oraz dla wybranych sekwencji wykonano analizy eksperymentalne, takie jak PCR, badanie aktywności polimerazy II RNA (ChIP-Seq) i określenie potencjalnych miejsc startu transkrypcji (CAGE). Udział Pani mgr Magdaleny Kubiak w tych badaniach polegał na optymalizacji i przeprowadzeniu analiz PCR dla 51 ludzkich i 11 mysich retrokopii. Niestety zarówno w rozprawie doktorskiej jak i w publikacji wyniki tych badań oraz opis metodologii nie został przedstawiony. W związku z tym pojawiają się pytania dotyczące zastosowanych kontroli, jakości wyników? Czy był to tradycyjny PCR na matrycy genomowego DNA, mówiący jedynie o obecności określonej sekwencji w genomie, czy jak można się domyślić analiza ekspresji metodą RT-PCR (PCR na matrycy cDNA), czy może PCR w czasie rzeczywistym (qRT-PCR)? Ponieważ retrokopie ulegają ekspresji na niskim poziomie oraz są trudne do analizy ze względu na specyfikę tych sekwencji, proszę o komentarz dotyczący zalet i wad metod służących do analizy ich ekspresji. Proszę o porównanie wyników metod zastosowanych w publikacji, czy były zgodne i która z nich jest wg Autorki najlepsza?

Podsumowując tę część rozprawy, mgr Magdalena Kubiak potwierdziła eksperymentalnie ekspresję 62 retrokopii adnotowanych w genomie człowieka lub myszy. Obejmowały one zarówno znane jak i nowe retrokopie, wytypowane metodą przewidywań bioinformatycznych. Co ciekawe, część analizowanych retrokopii wykazywała specyficzność gatunkową. Wyniki tych analiz posłużyły do rozbudowy i aktualizacji bazy danych RetrogeneDB, i zostały opublikowane w czasopiśmie Database w pracy pt. „*RetrogeneDB- a database of plant and animal retrocopies*”.

Kolejnym etapem badań mgr Magdaleny Kubiak była analiza zmienności genetycznej w postaci utraty retrokopii w obrębie wybranych populacji ludzkich. Polimorfizm retrokopii jest zjawiskiem częstym ale równocześnie trudnym do badania. Szczególnie trudnym do analizy jest polimorfizm związany z utratą retrokopii, ponieważ w przypadku względnie dużej częstości jego występowania w obrębie różnych populacji może być on wynikiem zarówno utraty retrokopii jak i nowej retrotranspozycji. W ramach pracy doktorskiej mgr Magdalena Kubiak weryfikowała eksperymentalnie obecność 5 nowych retrokopii w genomach, pochodzących od zróżnicowanych geograficznie przedstawicieli populacji ludzkich. Wykazała zmienność międzypopulacyjną trzech z nich i w wyniku analiz porównawczych sekwencji człowieka, goryla i szympansa wykazała, że polimorfizm tych retrokopii wynika z ich utraty w niektórych populacjach. W ramach projektu przeprowadzono również analizy bioinformatyczne, które pozwoliły na identyfikację aktywnych transkrypcyjnie retrokopii, które ulegały delecji w niektórych populacjach ludzkich i były zachowane ewolucyjnie. Zadaniem Pani Mgr Magdaleny Kubiak była eksperymentalna walidacja ekspresji 11 z nich w 16 tkankach ludzkich. W tym celu wykorzystwała metodę RT-PCR oraz qRT-PCR, a produkty amplifikacji były sekwencjonowane w celu potwierdzenia specyficzności amplifikacji. Z powodów technicznych ilościowa metoda qRT-PCR była zastosowana jedynie w przypadku pięciu retrokopii. Co ciekawe, część wyników nie była zgodna, np. dlaczego dla retro_hsap_3468 w niektórych tkankach nie obserwuje się ekspresji w qRT-PCR? Zjawisko utraty retrokopii i jej funkcjonalne implikacje są równie ciekawe jak sama retrotranspozycja. Autorzy wykazali, że aktywne, zachowane ewolucyjnie retrokopie uległy delecji w bardzo małej frakcji alleli, co może świadczyć m.in. o ich ukierunkowanej selekcji i być może ważnej funkcji. Czy dziś wiemy więcej o funkcji tych konkretnych retrokopii?

Podsumowując tę część badań, mgr Magdalena Kubiak przyczyniła się do identyfikacji i głębszej charakterystyki polimorfizmu utraty retrokopii w obrębie populacji ludzkich. Wyniki badań opublikowano w czasopiśmie Plos Genetics w artykule pt. *"Inter-population differences in retrogene loss and expression in humans"* i włączono do bazy RetrogeneDB.

Analiza funkcjonalna retrokopii to kolejny ważny element rozprawy doktorskiej, a jego opracowanie zostało poprzedzone przygotowaniem ciekawej i wyczerpującej pracy przeglądowej pt. *"Protein-Coding Genes, Retrocopies and Their Functions"* opublikowanej w czasopiśmie Viruses. W kolejnym etapie Doktorantka wykorzystwała dane zdeponowane w bazie RetrogeneDB do poszukiwania funkcji retrokopii (wyrażanych na poziomie białka, RNA i DNA), takich jak kodowanie białek, krótkich peptydów, transkryptów antysensowych, czy kompetytywnych RNA (ceRNA). Zidentyfikowała wiele potencjalnie funkcjonalnych retrokopii. Proszę o omówienie podczas obrony w jaki sposób były przeprowadzone te analizy i jakie parametry decydowały o uznaniu, że retrokopia ma potencjalną funkcję lub nie. Powyższe badania ujawniły nową grupę 117 „niekodujących”

retrokopii, które mogą ulegać translacji; 181 retrokopii mogących funkcjonować jako ceRNA oraz setki potencjalnych transkryptów antysensowych. W wyniku analizy transkryptów fuzyjnych Doktorantka wysunęła hipotezę, że retrokopie, które nie ulegają ekspresji również mogą mieć istotny wpływ na kształt transkryptomu. Wyniki te stwarzają pole do ciekawych badań funkcji retrokopii w tworzeniu nowych mechanizmów regulatorowych, ścieżek oddziaływań zarówno w wymiarze organizmu jak i ewolucji gatunków. Ciekawym aspektem w kontekście funkcjonalności retrokopii jest ich tkankowo-specyficzna vs konstytutywna ekspresja (Tabela nr 1). Czy na formę ekspresji i jej siłę ma wpływ lokalizacja genomowa retrokopii? Czy istnieje korelacja między ekspresją retrokopii a ekspresją ich „gospodarzy” czy „sąsiadujących” genów? Czy retrokopie ulegające ekspresji np. w trzustce mają cechy wspólne? Wyniki tej części rozprawy zostały opublikowane w czasopiśmie *Genes* w pracy pt. „*Complex analysis of retroposed Genes’ Contribution to Human Genome, Proteome and Transcriptome*”. Z oświadczenia Doktorantki o jej procentowym udziale w przygotowaniu pracy wynika, że była ona głównym wykonawcą tych badań (udział oszacowany na 70%) i jest to kluczowa część rozprawy doktorskiej.

W ostatniej pracy, wchodzącej w skład rozprawy doktorskiej mgr Magdalena Kubiak podsumowała stan wiedzy na temat molekularnych podstaw działania naturalnych transkryptów antysensowych w chorobach człowieka. Zebrała również informacje na temat baz danych transkryptów antysensowych. Jest to ciekawe i wartościowe opracowanie, będące naturalnym uzupełnieniem badań funkcji retrokopii i świadczącym o szerokiej wiedzy Doktorantki z zakresu regulatorowych RNA. Wyniki tych analiz opublikowano w artykule przeglądowym pt. „*Natural antisense transcripts in diseases: from modes of action to targeted therapies*” w czasopiśmie *WIREs RNA*.

Wnioski końcowe:

Podsumowując, Doktorantka podjęła się bardzo ciekawego i trudnego tematu badawczego. Zrealizowała wszystkie postawione sobie cele. Połączyła w swoich badaniach zaawansowane analizy bioinformatyczne z metodami eksperymentalnymi, a takie umiejętności są niezwykle cenne przy coraz szerszym stosowaniu analiz wysokoprzepustowych. Choć zastosowane metody eksperymentalne zaliczyć można do podstawowych technik biologii molekularnej ich dobór jak i jakość wykonania nie budzi zastrzeżeń.

W związku z tym stwierdzam, że przedstawiona mi do oceny rozprawa doktorska pt. **„*Udział retrokopii genów kodujących białka w kształtowaniu genomu, transkryptomu i proteomu człowieka*”** w pełni odpowiada warunkom stawianym pracom doktorskim i zwracam się do Rady Naukowej Wydziału Biologii Uniwersytetu im. Adama Mickiewicza w Poznaniu o dopuszczenie Pani mgr Magdaleny Kubiak do dalszych etapów przewodu doktorskiego. Biorąc pod uwagę, że wyniki badań stanowią istotny wkład w rozwój badań nad funkcją retrokopii i zostały opublikowane w postaci pięciu artykułów naukowych, wnioskuję o wyróżnienie tej pracy doktorskiej.



Marta Olejniczak

