

Recenzja
rozprawy doktorskiej mgr Joanny Ciomborowskiej pt.
„Funkcjonalne retrogeny w genomie człowieka”
[Functional retrogenes in the human genome]

Praca doktorska mgr Joanny Ciomborowskiej dotyczy niezwykle interesujących zagadnień związanych z występowaniem i ewolucją funkcjonalnych retrogenów w genomie człowieka. Prowadzone przez minione dwadzieścia lat badania zmieniły pierwotny pogląd na kopie genów powstające w wyniku retropozycji, tj. odwrotnej transkrypcji mRNA i wbudowania cDNA w sekwencję genomową. Retrokopie, przez wiele lat nazywane „śmieciowym” DNA, obecnie uznawane są za „ziarna ewolucji” i intensywnie badane. Wiedza na temat powszechności ich występowania, ewolucji i roli w genomie jest jednak nadal niewystarczająca. Praca mgr Ciomborowskiej doskonale wpisuje się w nurt badań dotyczących procesów ewolucyjnych związanych z powstawaniem funkcjonalnych retrogenów i ich roli w genomie człowieka.

Przedłożona do oceny dysertacja jest, w rozumieniu nowej procedury przewodu doktorskiego, osiągnięciem naukowym, na które składa się spójny tematycznie zbiór trzech anglojęzycznych artykułów opublikowanych w czasopismach naukowych. Jedną z publikacji to praca przeglądowa przedstawiająca dotychczasowy stan wiedzy na temat funkcjonalnych retrokopii (retrogenów) w genomach zwierzęcych; dwie pozostałe to prace oryginalne opisujące nowatorskie analizy i wynikające z nich odkrycia dotyczące dwóch słabo poznanych zjawisk: powstawania intronów w retrogenach oraz występowania w genomie funkcjonalnych retrokopii pozbawionych swoich „rodzicielskich” genów wieloeksonowych (tzw. osieroconych retrogenów).

W dysertacji przedstawione zostały wymagane ustawowo oświadczenia współautorów, określające ich rolę w przygotowaniu trzech publikacji stanowiących podstawę osiągnięcia naukowego. Pozycja p. Ciomborowskiej jako pierwszego lub równorzędnego autora w tych publikacjach pozwala wnioskować, że Jej wkład w wykonanie prac i przygotowanie publikacji był znaczący. Ciekawa byłaby dodatkowa informacja „ilościowa”, na ile Doktorantka określa w przybliżeniu swój procentowy udział w powstaniu tych prac, zarówno na etapie analiz jak i redagowania manuskryptów.

Praca przeglądowa [opublikowana w 2012 r. jako rozdział w *Evolutionary Biology – Mechanisms and Trends*], w której Doktorantka jest pierwszym autorem, stanowi rzetelne omówienie stanu wiedzy w zakresie funkcjonalnych retrogenów w genomach zwierzęcych, oparte o aktualną i trafnie dobraną literaturę (ponad 70 pozycji, od prac „historycznych” do prac z 2011 r.). We wstępie przedstawiona została zwięzła charakterystyka procesu duplikacji genów

na drodze retropozycji, omówione cechy charakteryzujące retrokopie genów rodzicielskich oraz podane przykłady funkcjonalnych retrokopii (retrogenów) w genomach różnych organizmów. W rozdziale traktującym o liczbie retrogenów w genomach zwierzęcych wyjaśnione zostały przyczyny rozbieżności wyników uzyskiwanych przez różnych autorów. Osobny rozdział poświęcony jest metodom identyfikacji retrokopii i retrogenów. Słusznie podkreślono słabe strony niektórych kryteriów stosowanych w takich analizach. Na przykład, zachowanie nietkniętej otwartej ramki odczytu (ORF) nie jest konieczne jeżeli retrogen pełni nową funkcję (np. gąbki mikroRNA); z kolei w przypadku sekwencji EST (expressed equence tag) trudność może stanowić różnicowe przypisanie ich do genu rodzicielskiego albo retrogenu.

W kolejnym rozdziale, dotyczącym roli retrogenów w tworzeniu różnic międzygatunkowych, podkreślona została, wynikająca z badań ogólnogenomowych i z analiz poszczególnych retrogenów, gatunkowa specyficzność większości retrokopii, w tym funkcjonalnych retrogenów. W rozdziale dotyczącym ewolucji retrogenów, omówiono różne scenariusze nabywania przez retrokopie zdolności do ekspresji i różne typy funkcjonalizacji retrokopii. Na końcu pracy przeglądowej przedstawione zostały różne hipotezy wyjaśniające odmienną niż w genach rodzicielskich tkankową specyficzność ekspresji funkcjonalnych retrokopii.

Praca przeglądowa świadczy o dobrej znajomości tematu funkcjonalnych retrogenów i umiejętności logicznego przedstawienia złożonych zagadnień. W świetle wykorzystania tej pracy jako rodzaju teoretycznego wstępu do dysertacji, zabrakło jednak szerszego opisu aspektów tematyki wymagających wyjaśnienia i dalszych badań, opisu mogącego stanowić zapowiedź prac oryginalnych przedstawionych w cyklu publikacji. Takiego akapitu, stanowiącego rozwinięcie przesłanek podjęcia badań, brakuje także w streszczeniu poprzedzającym załączone publikacje.

Dwie prace oryginalne [opublikowane w *Mol. Biol. Evol.* w latach 2011 i 2013] mają układ typowy dla artykułów w czasopismach naukowych. Każda zawiera krótki wstęp, opis stosowanej metodyki badań, obszerne przedstawienie uzyskanych wyników i dyskusję stanowiącą omówienie wyników w odniesieniu do szerokiej, właściwie dobranej i aktualnej literatury.

Pierwsza z prac, w której Doktorantka dzieli pierwsze autorstwo z dr Szcześniakiem, poświęcona jest zjawisku intronizacji sekwencji eksonowych w ludzkich retrogenach. Zjawisko takie nie było dotąd udokumentowane i opisane w genomach ssaków. W badaniach dotyczących retropozycji sekwencji rodzicielskich genów *RNF113* i *DCAF12*, autorzy znaleźli – poza typowymi jednoeksonowymi retrogenami (*RNF113A* i *DCAF12L1*) – funkcjonalne retrokopie zawierające introny (*RNF113B* i *DCAF12L2*). Analiza genomu wielu gatunków pozwoliła na ustalenie ewolucyjnego pochodzenia tych retrogenów w różnych liniach. Zidentyfikowano jeden

przypadek intronizacji w specyficznym dla naczelnych retrogenie *RNF113B* i dwie niezależne intronizacje w retrogenie *DCAF12L2* – jedną wspólną dla naczelnych i gryzoni, drugą obecną tylko u gryzoni. Na podstawie analizy sekwencji *RNF113B* zaproponowano mechanizm intronizacji, tj. mutacje tworzące miejsca splicingowe w sekwencji eksonowej powiązane z włączeniem sekwencji LINE1 i Alu na końcach retrogeny. W przypadku retrogeny *DCAF2L1* duża dywergencja sekwencji retrogeny i genu macierzystego nie pozwoliła na wyjaśnienie mechanizmu intronizacji. Autorzy podają, że wykryte warianty są pierwszym opisanym w literaturze przykładem intronizacji w retrogenie u ssaków. W przypadku retrogeny *RNF113B*, powiązanie pojawienia się intronu z eksonizacją sekwencji LINE1 wbudowanej na końcu 3' retrogeny budzi niejaką wątpliwość co do autentyczności zjawiska intronizacji definiowanego jako przekształcenie w intron zakonserwowanej sekwencji eksonowej; nie można bowiem wykluczyć, że przypadek ten reprezentuje nabycie intronu poprzez przyłączenie nowego eksonu, a ten mechanizm jest znany i był wcześniej opisywany.

Badanie profilu ekspresji pozwoliło dodatkowo stwierdzić, że retrogeny *RNF113B* i *DCAF12L2* ulegają alternatywnemu splicingowi; warianty bez intronów ulegają ekspresji w wielu tkankach, natomiast warianty z intronem (młodsze ewolucyjnie) wykazują tkankowo-specyficzną ekspresję w jądrach, zgodną z wzorem ekspresji wielu retrogenów. Dodatkową interesującą obserwacją jest kierunek retropozycji omawianych sekwencji, z autosomów na chromosom X; kierunek ten świadczy o tym, że preferencyjna ekspresja retrogenów w jądrach może wynikać z hipertranskrypcji w komórkach gametogenicznych a nie z zabezpieczenia przed mejozytyczną inaktywacją chromosomu X w spermatogenezie.

Pewne zastrzeżenia budzi jakość Ryc. 4. Według opisu w tekście, forma *RNF113B* z zachowanym intronem ulega ekspresji we wszystkich badanych tkankach; tymczasem na dolnym panelu produkty o długości 295 pz są niewidoczne w ścieżkach reprezentujących mózg, mięśnie, jądra i leukocyty.

Na str. 286 publikacji (\$1, l.9) wystąpił błąd w nazwie retrogeny (Rps23r1 > Rps23rg1).

Druga publikacja oryginalna, w której mgr Ciomborowska jest pierwszym autorem, opisuje badania, które pozwoliły na zidentyfikowanie w ludzkim genomie retrogenów nie posiadających swoich odpowiedników w postaci wieloeksonowych genów rodzicielskich. W wyniku zastosowania kilku równoległych metod analitycznych (porównanie podobieństwa sekwencji genów człowieka i kury; analiza położenia porównywanych sekwencji w genomie; analiza struktury ortologów człowieka i *C. elegans*) zidentyfikowanych zostało 25 takich „osieroconych” retrogenów, przy czym cztery z tych genów wykryto każdą z opisanych metod. Rycina przedstawiona jako materiał dodatkowy, na której przedstawiono porównanie struktury osieroconych retrogenów i ortologów ich wieloeksonowych odpowiedników, dowodzi, że w

retrogenach została zachowana struktura większości eksonów obecnych w genach rodzicielskich. W rycinie tej brakuje wyjaśnienia kolorów (kolor czarny/czerwony – nowe sekwencje kodujące dołączone do końca lub w środku sekwencji pochodzącej z genu rodzicielskiego; co oznacza czarny odcinek w środku genu KTI12?); nie wyjaśniono też znaczenia wartości %; być może jest to % sekwencji kodującej genu w retrogenie i retrogenu w genie; nie wiadomo jednak, czy porównanie dotyczy długości czy sekwencji; aminokwasów czy nukleotydów?

Analiza z użyciem bibliotek ludzkiego cDNA z różnych tkanek pozwoliła na określenie wzoru ekspresji ludzkich osieroconych retrogenów. Analiza występowania tych retrogenów i ortologów ich genów rodzicielskich w siedmiu innych gatunkach pozwoliła na zaproponowanie filogenetycznego drzewa opisującego ewolucję tych sekwencji. Stosunek Ka/Ks sugeruje, że większość osieroconych genów pozostaje pod wpływem silnej selekcji oczyszczającej.

Szeroki zakres ekspresji osieroconych retrogenów (w wielu tkankach i organach, w opozycji do wybiórczej ekspresji większości retrogenów, ograniczonej do jąder) wydaje się wspierać sugerowaną przez Autorów teorię o przejęciu przez osiercone retrogeny funkcji genów rodzicielskich. Z drugiej strony, analiza otoczenia sekwencji retrogenu *CHMP1B*, który u człowieka jest retrogenem osieroconym, natomiast u myszy współlistnieje z wieloeksonowym funkcjonalnym genem rodzicielskim, wykazała, że retrogeny w obu gatunkach posiadają w swoim bezpośrednim sąsiedztwie wspólne sekwencje regulatorowe, odmienne od tych, które znajdują się w regionach 5' i 3' genu rodzicielskiego u myszy. Obserwacja, że retrogeny w obu gatunkach mogą być regulowane w sposób podobny, lecz inny niż gen rodzicielski, wydaje się świadczyć o subfunkcjonalizacji tych retrogenów. W sumie, opisane wyniki nie dostarczają niezbitego dowodu na to, że istnienie osieroconych retrogenów wskazuje na zastąpienie istniejącej funkcji, a nie na subfunkcjonalizację połączoną z zanikiem funkcji „rodzicielskiej”. Autorzy zdają sobie z tego sprawę (vide rozdział o asocjacji z chorobami oraz ostatnia strona dyskusji), ale w związku z tym powinni nieco ostrożniej formułować stwierdzenia o zastępowaniu genów rodzicielskich przez retrogeny. W tym kontekście ciekawe byłoby na przykład zbadanie fenotypowego efektu wyciszenia genu rodzicielskiego w gatunkach, gdzie geny rodzicielskie są nadal obecne; czy Autorzy rozważali możliwość badania wyniku knockoutu genu *CHMP1* u myszy?

W rycinie 4 nastąpiła omyłkowa zamiana podpisów (A < B).

Historia ewolucyjna osieroconych retrogenów wskazuje, że największa ich ilość powstała na stosunkowo wczesnych etapach ewolucji zwierząt. Według Autorów jest to zaskakujące, gdyż wiadomo, że retropozycja była szczególnie intensywna u ssaków. Tymczasem, powstawanie osieroconych retrogenów jest wynikiem nie tyle intensywności retropozycji, ile procesu zanikania funkcjonalnych genów rodzicielskich; dla lepszego zrozumienia tego procesu ważne byłoby określenie roli retrogenów w osłabieniu selekcji genów rodzicielskich.

Inne powiązane z tym pytanie dotyczy czasu od powstania retrokopii do zaniku genu rodzicielskiego. Drzewo filogenetyczne przedstawione w publikacji sugeruje względną czasową bliskość tych dwóch procesów. Ta bliskość może być jednak jedynie pozorna, wynikająca z wyboru przedstawionych na tym drzewie taksonów. Analiza innych gatunków, zwłaszcza wśród bezkręgowców, pozwoliłaby zapewne na znalezienie taksonów posiadających zarówno gen wieloeksonowy jak i osierocony u człowieka retrogen (tak jak dla młodego retrogenu *CHMPI* u myszy) i dokładniejsze określenie, kiedy w ewolucji nastąpił zanik funkcji poszczególnych genów rodzicielskich i jak silna była konserwacja sekwencji kodowanych białek.

Inna interesująca obserwacja przedstawiona w opublikowanej pracy dotyczy związku pomiędzy wiekiem retropozycji i intensywnością/zakresem ekspresji; odwrotnie do oczekiwań, najmłodsze z osieroconych retrogenów (retropozycja u przodka zwierząt ciepłokrwistych) wykazują najsilniejszą i najszerszą ekspresję. Część dyskusji poświęcona jest omówieniu znaczenia badań funkcjonalnych retrogenów w genomie człowieka dla medycyny. Autorzy słusznie zauważają, że istnienie osieroconych retrogenów oraz niezależne powstawanie retrogenów w różnych liniach ewolucyjnych wymaga dużej ostrożności w kwestii przełożenia wyników z organizmów modelowych na organizm ludzki. W sumie, publikacja prezentuje bardzo ciekawe informacje dotyczące udziału retrokopii w ewolucji genomów, mogące stanowić podstawę dalszych badań naukowych.

Dołączony do trzech publikacji tekst obejmuje 6-stronicowe streszczenia w języku polskim i angielskim, opisujące pokrótce stan wiedzy w zakresie tematu pracy oraz opublikowane wyniki i plany badawcze. W streszczeniu tym, podobnie zresztą jak w publikacjach, daje się zauważyć brak konsekwencji w nomenklaturze dotyczącej tematu. Jest to po części usprawiedliwione faktem licznych biologicznych odstępstw od ogólnego schematu, jednak konsekwentne zdefiniowanie i rozłączne stosowanie (przynajmniej w obrębie dysertacji) określeń: „retrokopia”, „retrogen (funkcjonalna retrokopia)”, „retropseudogen”, ułatwiłoby zrozumienie o czym w danej chwili jest mowa. Najśluszniesze wydaje się nazywanie **retrokopiami** wszystkich powielonych w procesie retropozycji kopii genów rodzicielskich, **retropseudogenami** – kopii nieaktywnych, a **retrogenami** – retrokopii ulegających ekspresji i funkcjonalnych (ekspresja i funkcja też nie są równoznacznymi pojęciami); określenie „funkcjonalny retrogen” zawiera cechy redundancji i lepsze byłoby sformułowanie „funkcjonalna retrokopia”. W założeniu taka właśnie nomenklatura była przez Doktorantkę przyjęta (vide *Glossary* w przeglądówce), jednak w wielu miejscach terminy „retrogen” i „retrokopia” używane są zamiennie w sposób osłabiający jasność przekazu, np.: str.1, \$2, l. 1 i 4; \$3, l.8; str. 284, \$2, l. 9; publikacja 1, str. 285, \$2, l. 2; str. 290, \$2; str 292, \$ 3, l. 4, 10 – tam powinna być mowa o retrokopiach, nie retrogenach; publikacja 1, str. 286, \$2, l.1; str. 287, \$2, l. 2 – tam powinna być mowa o retrogenach lub funkcjonalnych retrokopiach.

Podobny brak jednolitej nomenklatury w literaturze przedmiotu może być – poza wymienionymi przez Doktorantkę różnicami metod identyfikacji i kryteriów definiowania kopii jako funkcjonalnej – jedną z przyczyn, dla których określenie liczby retrogenów (funkcjonalnych retrokopii) w genomach nie jest jednoznaczne.

Mimo, że streszczenie dysertacji jest zredagowane bardzo starannie, Autorka nie ustrzegła się kilku błędów literowych/słownych, które wymieniam poniżej:

Str. 2, \$2, l.13: Rps21rg1 > Rps23rg1

Str. 3, \$2, l. 7: komórki spermatog~~enetyczne~~ > spermatog~~eniczne~~

Str. 4, \$3, l. 2 – „retrogeny są także miejscem intensywnego powstawania nowych intronów u ssaków” – czy jest tego tak dużo, by usprawiedliwić określenie „intensywny”?

Na zakończenie pragnę podkreślić, że wyniki uzyskane przez mgr Ciomborowską i współautorów stanowią ważki wkład w badania nad ewolucją i rolą funkcjonalnych retrogenów w ludzkim genomie. Przedstawione przeze mnie uwagi krytyczne nie umniejszają wysokiej wartości ocenianej dysertacji i stanowiących jej podstawę publikacji; większość tych uwag można potraktować jako „ziarna dyskusji” w zakresie niezwykle interesującego tematu funkcjonalnych retrogenów w genomie człowieka. Z zadowoleniem konstatuje, że Doktorantka uczestniczy w dalszych badaniach, które stanowią kontynuację przedstawionej pracy; projekty te to wielkoskalowa identyfikacja i analiza retrogenów i ich genów rodzicielskich w ponad 600 genomach zwierzęcych; poszukiwanie i charakterystyka kolejnych przykładów intronizacji w retrogenach; badanie funkcjonalnych, uznawanych dotąd za pseudogeny, retrogenów specyficznych dla człowieka i istotnych w kształtowaniu różnic międzygatunkowych.

Podsumowując stwierdzam, że w mojej ocenie rozprawa doktorska mgr Joanny Ciomborowskiej spełnia wszystkie wymagania, stawiane tego typu rozprawom. Zwracam się więc do Wysokiej Rady Wydziału Biologii Uniwersytetu Adama Mickiewicza w Poznaniu z wnioskiem o dopuszczenie pani mgr Ciomborowskiej do dalszych etapów przewodu doktorskiego i o wyróżnienie przedłożonej do oceny pracy doktorskiej.

Ewa Ziolkiewicz, prof. dr hab.

Instytut Genetyki Człowieka PAN w Poznaniu

Poznań, 24 kwietnia 2014 r.