

Gen strukturalny – ewolucja pojęcia i dylematy

Mieczysław Chorąży*

The structural gene – evolution of concept and controversies

Pracę tę dedykuję tym wszystkim byłym i obecnym pracownikom Zakładu Biologii Nowotworów, Zakładu Biologii Molekularnej i Zakładu Radiobiologii Doświadczalnej i Klinicznej Centrum Onkologii – Instytutu im. Marii Skłodowskiej-Curie, Oddział w Gliwicach, których rzetelność naukowa, żarliwość i oddanie nauce pozwalały zawsze przetrwać trudne lata.

Genom – ile DNA, ile genów?

DNA komórek prokariotycznych (beźjadrzastych, np. bakterie) jest zorganizowany jako jeden chromosom, składający się z jednej cząsteczki DNA „rozpuszczonej” w cytoplazmie. Cała taka cząsteczka DNA obejmuje geny ułożone blisko jeden za drugim, każdy z nich ma swój wymiar przestrzenny, dość wyraźne „granice” (początek i koniec) i każdy z nich ma nieprzerwany ciąg sekwencji kodujących.

Natomiast DNA organizmów eukariotycznych (posiadających jądro), zarówno jedno- jak i wielokomórkowych, upakowany jest w chromosomach i zamknięty w jądrze komórkowym. Jeden chromosom, zawiera jedną

cząsteczkę DNA. Człowiek posiada 22 chromosomy autosomalne, oznaczone według malejącej długości od 1 do 22 i dwa różne chromosomy płciowe X i Y (u kobiet XX, u mężczyzn XY). Uporządkowany według wielkości zestaw chromosomów nazywamy **kariotypem**. W jądrze komórek człowieka poszczególne chromosomy zajmują ściśle określone terytoria. Nawet w okresie cyklu życiowego komórki zwanym interfazą, gdy ma miejsce duża aktywność genów, gdy chromosomy wydają się rozproszone, a DNA chromosomalny jest pozornie rozpuszczony i wymieszany, szkielet każdego indywidualnego chromosomu jest utrzymany, a jego DNA „nie miesza” się z DNA innego chromosomu. DNA każdego chromosomu zajmuje określone terytorium jądra.

W organizmach diploidalnych takich jak człowiek każdy chromosom występuje w dwóch kopiach – jedna pochodzi od matki i jedna od ojca. **Genomem** nazywamy całkowitą „informację genetyczną” czyli DNA (choć nie są to terminy równoważne), zawartą w komórce i charakterystyczną dla danego organizmu. Wielkość genomu najczęściej podajemy jako długość wyrażoną liczbą par zasad (nukleotydów) i odnosimy zwykle do haploidalnego zestawu chromosomów. „Typowy” genom człowieka, czyli sumaryczna długość cząsteczek DNA wszystkich chromosomów w komórce haploidalnej (zawierającej zestaw jednej kopii chromosomów danego gatunku), ma wielkość około 3×10^9 par zasad/nukleotydów (3×10^9 bp) czyli 3 miliardy par zasad. W przeliczeniu na miarę metryczną DNA każdej komórki człowieka ma długość około 1,5 metra! Ta wielkość daje nam przybliżone pojęcie o tym, jak silnie skondensowany i gęsto upakowany musi być DNA w indywidualnym chromosomie i w jądrze komórkowym, które – jeśli przyjmiemy jego kulisty kształt – ma średnicę zaledwie od kilku do kilkunastu μm .

Wielkość genomów różnych gatunków mieści się w szerokich granicach, od około 10^6 bp u mykoplazmy i bakterii aż do 10^{11} bp u kwitnących roślin i zimnokrwistych kręgowców klasy *Amphibia*. Czyli w haploidalnych genomach różnych gatunków ilość DNA waha się do

* Em. Prof. dr hab. n. med. Mieczysław Chorąży, Centrum Badań Translacyjnych i Biologii Molekularnej Nowotworów, Centrum Onkologii – Instytut, Oddział w Gliwicach

W kwartalniku *Nauka* (nr 3/2009, str. 57-108) ukazał się obszerny artykuł Mieczysława Chorążego pod tytułem „Gen strukturalny – ewolucja pojęcia i dylematy”. Redakcja *Nowotworów* uznała za pozytywne dokonać przedruku tej publikacji. Praca podaje podstawowe informacje z zakresu biologii i genetyki o tym, jak kształtowało się pojęcie „genu” i jak współcześnie to pojęcie postrzegamy. Redakcja posiada informację, że autor tej pracy przygotowuje cykl kilku publikacji, wprowadzających Czytelnika w zagadnienia współczesnej biologii komórki, które będą także wywierać wpływ na kierunki badań nad rakiem. W bieżącym numerze *Nowotworów* publikujemy część drugą wspomnianego artykułu, za zgodą Redakcji *Nauki*.

100 000 razy. Wielkość genomu bakteriofagów jest rzędu 10^5 bp. Jeszcze około 20 lat temu spekulowano, że złożoność organizmu będzie ściśle korelować z ilością materiału genetycznego (ilością DNA), a zatem również z liczbą genów i liczbą chromosomów. W miarę gromadzenia danych doświadczalnych i rozwijania metod badawczych okazało się, że sprawa jest bardziej złożona. Nie ma prostej korelacji między biologiczną złożonością gatunków, a ilością DNA, ani liczbą chromosomów, ani ilością kodującego lub niekodującego DNA.

U bakterii, sekwencje DNA kodujące białko (czyli geny) zajmują cały lub prawie cały genom. Ale w organizmach wyższych (ssaki) sekwencje kodujące białko stanowią tylko małą część genomu, np. u człowieka najwyżej około 1,5% genomu. Jeśli przyjąłmy, że liniowa długość DNA genomu człowieka wynosi około 1,5 m, to łączna długość sekwencji kodujących (tj. genów) zajmuje zaledwie około 2 cm. U niektórych kręgowców klasy *Amphibia*, mających genom wielkich rozmiarów, ta proporcja jest jeszcze mniejsza. W genomie człowieka i innych gatunków zwierząt oraz roślin jest znacznie więcej DNA, niż byśmy mogli oczekiwać z oczywistych różnic w złożoności tych organizmów. Przez wiele lat uznawano niesłusznie (por. niżej), że ten „nadmiarowy” DNA u człowieka i innych organizmów wyższych nie ma znaczenia funkcjonalnego i jest wynikiem procesów ewolucyjnych, a nawet ukuto nazwę dla „nadmiarowego” DNA – ewolucyjne „śmiecie” (*junk DNA*). Organizmy prokariotyczne takich „nadmiarowych” sekwencji DNA nie posiadają.

W latach 90. badaczy intrygowało głównie zagadnienie, jak dużo jest genów u każdego gatunku. Sądzono, że im gatunek jest na wyższym stopniu ewolucyjnego rozwoju, tym więcej będzie posiadał DNA i genów. Jeszcze kilkanaście lat temu szacowano, że człowiek powinien posiadać przynajmniej 100 000 genów, pojmowanych nadal jako kodujące białko sekwencje DNA, mające swój materialny początek i koniec. Szacunki bardzo różniły się w zależności od użytej metody. Analiza sekwencji DNA organizmów posiadających mały genom i poznanie sekwencji aminokwasowej różnych białek pozwalało „uśrednić” wielkość genu i uściślać liczbę genów. Geny kodujące rybosomalny RNA (rRNA), transportujący RNA (tRNA), a także geny kodujące histony występują zwykle w wielu kopiach. Często nazywa się je „genami powtarzalnymi”. Należy je odróżnić od „sekwencji powtarzalnych” lub „powtarzających się” (sekwencje repetytywne) (por. niżej). Liczne kopie jednego genu występują czasami w komórkach niektórych nowotworów.

W latach 90. w Stanach Zjednoczonych opracowano i rozpoczęto projekt badawczy zmierzający do ustalenia sekwencji (czyli kolejności występowania w DNA) 3 miliardów par zasad/nukleotydów DNA całego genomu człowieka (Human Genome Project, HGP). Ustanowienie projektu HGP przez USA (Narodowe Instytuty Zdrowia, NIH oraz Departament Energii, DOE) i międzynarodową organizację Human Genome Organization (HUGO) było poprzedzone usilnymi zabiegami naukowców, takich jak Walter Gilbert, James Watson, Charles Cantor, Leroy Hood (cyt. za Lewontin, 1991).

Wierzono, że HGP pozwoli ustalić liczbę genów i zrozumieć, na czym ostatecznie polega zjawisko życia, ustalić cały, liniowo zakodowany, „plan organizmu” z wszelkimi anatomicznymi i fizjologicznymi cechami i zachowaniem. W. Gilbert obrazowo ujął swój entuzjazm wobec HGP: „Sekwencję trzech miliardów par zasad DNA człowieka będzie można zapisać na dysku CD, wyciągnąć taki dysk CD z kieszeni i oznajmić: „Oto jest istota ludzka: to JA!”.

Projekt wzbudzał wielkie nadzieje. Sądzono, że poznanie sekwencji pozwoli na ustalenie liczby genów oraz genów i sekwencji unikatowych dla człowieka, czyli na ustalenie czym człowiek różni się ewolucyjnie od zwierząt wyższych i najbliższych mu naczelnymi (np. szympansa). Oczekiwano dalej, że zostaną poznane defekty i uszkodzenia genów determinujące powstawanie chorób (łącznie z rakiem), co z kolei pozwoli na rozwinięcie metod diagnostycznych i opracowanie nowych skutecznych technik leczniczych, jak np. „terapia genowa” (Zalewski, 2001). Każdy człowiek będzie mógł mieć swój „paszport genetyczny” dla identyfikacji, prosty i niezbędny np. w sprawach ubezpieczeniowych. Powstały też złudne iluzje, że znając zestaw genów i sekwencji DNA człowieka, będzie można ten gatunek uszlachetniać przez zabiegi inżynierii genetycznej, a nawet zwalczyć negatywne ludzkie zachowania.

W 2001 r. dwa ośrodki badawcze w USA, rządowy (NIH) pod kierunkiem F. Collinsa i prywatny (Celera) J. C. Ventera, ogłosiły równocześnie w tygodnikach *Nature* (Lander i in., 2001) oraz tygodniku *Science* (Venter i in., 2001) sekwencję ludzkiego DNA i pokazały „mapę genetyczną” człowieka. Był to wielki sukces badawczy, który jednocześnie okazał się być wielkim zaskoczeniem dla nauki. Wbrew wcześniejszym oczekiwaniom, liczbę „genów” człowieka oszacowano w przedziale 25 000 – 30 000. Liczba ta obejmowała sekwencje posiadające atrybuty genu strukturalnego, kodującego białko, jak również geny kodujące te klasy RNA, które były ostatecznymi produktami (rRNA, tRNA). Okazało się, że trudno jest zdefiniować i rozpoznać indywidualne geny. Rozrzut w szacunkowej liczbie genów zależy od metod przyjętych do identyfikacji, opartych na swoistych sekwencjach (sekwencje sygnałne charakterystyczne dla genu, jak np. promotor, sekwencje stop, otwarte ramki odczytu).

Inne grupy badawcze zaczęły wkrótce donosić o ustaleniu sekwencji i szacunkowych liczbach genów u kilku gatunków zwierząt i roślin. Porównanie liczby genów (sekwencji kodujących białko) między gatunkami wyraźnie wskazało, że nie ma także oczekiwanej korelacji między liczbą genów, a złożonością organizmów. Obecnie liczbę genów kodujących białko w genomie człowieka szacuje się na około 22 700, a nawet zaledwie 20 000 (cyt. wg Pheasant i Mattick, 2007). Człowiek, który umieścił siebie na szczycie drzewa ewolucyjnego, ma porównywalną liczbę genów do myszy (22 500), jeżowca (ok. 23 000) i nicienia *Caenorhabditis elegans* (ok. 19 000 – 20 000) i tylko o jedną trzecią więcej niż muszka owocowa *Drosophila melanogaster* (ok. 14 000), a znacznie niższą niż jednokomórkowy protista, orzęsek *Tetrahymena thermophila* – 27 000 genów (cyt. wg Pheasant i Mattick, 2007). Te

szacunkowe dane wskazują, że w procesie rozwoju osobniczego geny kodujące białko *per se* nie są odpowiedzialne za „programy rozwojowe”. Program (plan) rozwoju jest przypisywany genomowym mechanizmom regulacyjnym lub innym nieznanym jeszcze procesom. Drożdże posiadają około 6 000 genów, a bakterie między 5 000 a 10 000 genów. Sekwencje genomu człowieka i szympansa różnią się zaledwie niewiele ponad 1,0%. Czyżby – jak chcą niektórzy badacze – człowiek i człowieczeństwo, cechy osobowości człowieka, jego wielki ładunek intelektualny tkwił w tej drobnie DNA, stanowiącej różnicę z najbliższym krewniakiem ewolucyjnym?

Pozagenowe sekwencje DNA, występują u wielu gatunków i różne opisane niżej sekwencje posiadają często zdumiewające podobieństwo, a poprzez ich porównywanie można prześledzić ich koleje ewolucyjne (Jurka i in., 2007).

Co zawiera ludzki, nadmiarowy DNA, uważany do niedawna za ewolucyjne „śmiecie”? Jak wspomnieliśmy, mniej niż 2% ludzkiego DNA zawiera sekwencje kodujące, zatem nadmiarowa część DNA zajmuje aż około 98% genomu. Na tę część składają się sekwencje powtarzające się (*repetitive sequences, repetitive DNA*), występujące w licznych kopiach oraz fragmenty sekwencji unikatowych o nieznannej funkcji, nie posiadające swoich homologów w genomie. Sekwencje powtarzające się o prostej budowie, tandemowe, złożone z kilku (2-6) nukleotydów, powtarzających się zwykle kilkadziesiąt razy, np. dinukleotydy – [AT]_n dające trakty ATATATA..., trinukleotydy np. [AAT]_n lub [CGG]_n, tworzące odpowiednio ciągi: AATAATAAT... i CGGCGGCGGCGG..., nazywane są **mikrosatelitami**. Bardziej złożone sekwencje powtarzające się, mające motyw podstawowy od kilkunastu do kilkudziesięciu nukleotydów, definiowane są jako **minisatellity**. Zajmują one odcinki DNA o długości nawet kilkunastu tysięcy pz. Niekiedy „inwazja” krótkich (kilka nukleotydów) traktów w rejony kodujące ma miejsce w chorobach degeneracyjnych, głównie typu neurodegeneracyjnego. Do grupy powtarzających się sekwencji rozproszonych zaliczamy także sekwencje typu SINE (*short interspersed nucleotide elements*), z których najlepiej poznana sekwencja *Alu* o długości około 300 nukleotydów występuje w około 500 tys. kopii w genomie człowieka. Grupa sekwencji LINE (*long interspersed nucleotide elements*) o długości kilku tysięcy nukleotydów występuje w setkach tysięcy kopii. Znanych jest wiele innych rzadziej występujących elementów.

Wiele sekwencji powtarzających się ma zdolność przemieszczania (transpozycji) w genomie. Grupy rodzin takich sekwencji nazywamy ogólnie sekwencjami/elementami transposonalnymi (TE – *transposable elements*) lub wprost **transpozonomi**. Istnienie elementów transposonalnych, zwanych również elementami mobilnymi, postulowała w latach 30. Barbara McClintock na podstawie badań nad genetyką kukurydzy. Elementy transpozonalne są to sekwencje o dużej złożoności co do struktury, pochodzenia i różnorodności. Ich funkcja w procesach ewolucji, specjacji (powstawaniu gatunków), w rozwoju osobniczym i rola w genomie współczesnym nie jest do

końca poznana. Transpozony odegrały prawdopodobnie wiodącą rolę w procesie ewolucji (Jurka i in., 2007; Kapitono v i in., 2004). W warunkach stresowych transpozony mogą się uaktywnić, i dzięki swoim zdolnościom autokatalitycznym (aktywność własnego enzymu, transpozazy, umożliwiającego wycinanie i insercję) mogą zmieniać miejsce w genomie tego samego osobnika lub gatunku lub też ulegać transpozycji horyzontalnej do genomu innego gatunku. Masowy, horyzontalny współczesny transfer transpozonów opisano u wrotki (Gladyshev i in., 2008). Znane są przykłady chorób powstałych wskutek transpozycji aktywnych TE. Z kodujących sekwencji transpozonów (transpozaza, gen *gag* u retrotranspozonów) ewoluowały niektóre geny kodujące białko u ssaków. Transpozony ingerowały także w powstanie i funkcję sekwencji kodujących microRNA. W „pozagenowym” DNA znajdują się liczne powtarzające się kopie genów nieczynnych (pseudogeny, por. niżej).

Między genem a białkiem

Obraz prostego przekazywania informacji z genu (DNA) przez mRNA do białka i prostej liniowej relacji gen (DNA), struktura białka i jego funkcja bardzo komplikują odkrycia mechanizmów zmieniających na poziomie mRNA treść informacji pierwotnej zawartej w DNA (genie). W wyniku działania tych mechanizmów sekwencja aminokwasów w cząsteczce białka (a zatem jej struktura i funkcja) będą różnić się od tej „zapisanej” w genie.

Jedna z hipotez wyjaśniających ewolucję głosiła, że wprowadzenie do procesu ewolucyjnego „nowości” miałyby polegać na duplikacji genu. Jedna kopia genu miała służyć do kodowania „starego” białka, nowa kopia byłaby uwolniona dla ewolucyjnego „eksperymentowania” i ostatecznie posłużyłaby do wykształcenia nowych funkcji kodujących. Ta hipoteza nie wyjaśnia jednak faktu, że w miarę powstawania organizmów o coraz większej złożoności nie wzrasta proporcjonalnie liczba genów kodujących białko.

Dogmat jeden gen – jedno białko – jedna funkcja uległ obaleniu, gdy okazało się, że liczba rodzajów cząsteczek białek jest przynajmniej o jeden rząd większa niż liczba genów kodujących białko. W dodatku uległ również obaleniu pogląd, że przestrzenna trójwymiarowa konformacja cząsteczki białkowej jest determinowana wyłącznie sekwencją aminokwasów w polipeptydzie, co pierwotnie miał zdeterminować gen kodujący. Przestrzenna forma białka zmienia się w zależności od mikrośrodowiska, np. oddziaływań danej cząsteczki białkowej z innym białkiem, różnymi drobnymi cząsteczkami nie-białkowymi, jonami, itd. Przestrzenna forma zależy też od chemicznych potranslacyjnych modyfikacji białek. Takie modyfikacje obejmują fosforylację, metylację, acetylację, sumoylizację, ubikwitynację, glutamylację, itp. Modyfikacje potranslacyjne zmieniają funkcję białka. Funkcjonalna różnorodność białek określana terminem „pleiotropia” zwiększa złożoność biologiczną białek.

W ostatnich dziesięcioleciach poznano mechanizmy (zapewne nie wszystkie), które wyjaśniają, jak można

osiągnąć różnorodność strukturalną i funkcjonalną białek kodowanych przez jeden gen. Najważniejsze z nich zostaną przytoczone niżej.

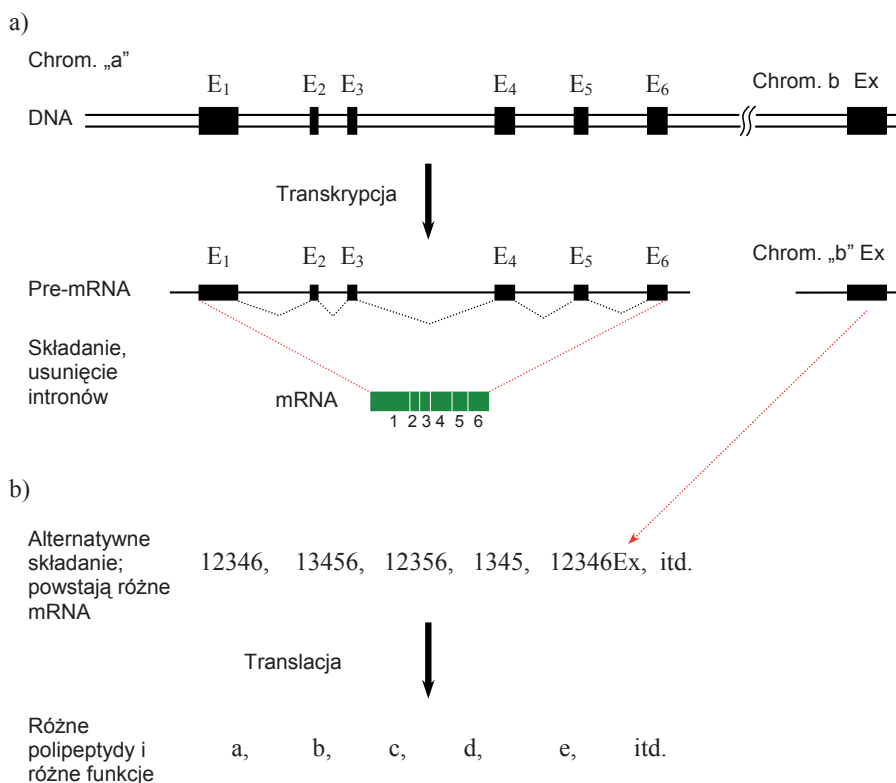
Alternatywne składanie mRNA

Oprócz możliwości odtwarzania identycznych kopii genu (DNA) głównym zadaniem genu jest programowanie syntezy białka. Istnieje wielka różnorodność białek oraz konieczność ich wybiórczo kontrolowanej syntezy w czasie rozwoju embrionalnego, a potem prowadzenia syntezy w różnych komórkach, tkankach i narządach oraz w odpowiedzi na zmienne bodźce środowiskowe. Jednym z mechanizmów zapewniających generowanie z jednego genu (DNA) różnych polipeptydów o różnych funkcjach jest alternatywne składanie pre-mRNA.

Wielka cząsteczka pre-mRNA (zwana też jednostką transkrypcyjną) podlega wspomnianemu wyżej procesowi składania (*splicing*). W procesie tym zostają usunięte długie niekodujące sekwencje intronowe, a eksony zostają złączone w jeden ciąg nukleotydów. W wielu przypadkach składanie przebiega w sposób alternatywny (Ryc. 5), generując z sekwencji jednego genu (jednego pre-mRNA) wiele różnych cząsteczek mRNA, kodujących różne polipeptydy (białka). Dla przykładu (Ryc.5a),

jeśli w pre-mRNA mamy trakt sześciu eksonów (E) oznaczonych kolejno 1 2 3 4 5 6, to w przypadku alternatywnego składania mogą powstać cząsteczki mRNA o ciągu E, jak w pre-mRNA, lecz także, oprócz intronów, mogą zostać usunięte niektóre eksony i utworzone zostaną cząsteczki mRNA o ciągach E (Ryc. 5b): 1 2 3 4 6 albo 1 3 4 5 6, albo 1 2 3 5 6, albo 1 3 4 5 itd. Takie cząsteczki RNA będą kodowały różne izoformy polipeptydu o różnych funkcjach. Zupełnie nieoczekiwanym było odkrycie, że w procesie składania transkryptu (pre-mRNA) genu „a” może być pobrany ekson z transkryptu genu „b”, leżącego w dużej odległości na tym samym chromosomie lub nawet na innym chromosomie. Wówczas może powstać mRNA o ciągu eksonów 1 2 3 4 6, Ex (Ryc. 5b). Po translacji aktywnie złożonych mRNA powstają różne izoformy białek. To odkrycie burzy pojęcie genu jako ściśle określonej jednostki strukturalnej, mającej swoje zdefiniowane granice.

Skrajnym przykładem zdumiewających możliwości alternatywnego składania jest transkrypt genu *Dscam*. U człowieka gen *DSCAM* (*Down syndrome cell adhesion molecule*) koduje duże polipeptydy, biorące udział w rozwoju układu nerwowego. Neurony w trakcie rozwoju osobniczego dostają etykietę w postaci swego spektrum polipeptydów, co pozwala im zająć odpowied-



Ryc. 5. Składanie mRNA. a) Zarówno eksony (E), jak i introny genu (DNA) na chromosomie „a” i niekiedy nawet część genu na chromosomie „b” są przepisane (transkrypcja) na długi transkrypt pierwotny (pre-mRNA). W procesie składania z pre-mRNA zostają wycięte sekwencje intronowe, a sekwencje eksonowe zostają spojone w jedną, „dojrzałą” cząsteczkę mRNA. b) Cyfry pokazują kilka różnych alternatywnych sekwencji eksonów w cząsteczce mRNA po alternatywnym składaniu, gdy oprócz intronów zostają wycięte również niektóre eksony. W pierwszej cząsteczce o ciągu 12346 został wycięty ekson E5, w drugim ciągu 13456 został wycięty ekson E2, itd. W ostatnim przykładzie o ciągu 12346x został wycięty ekson E5, ale dodatkowo został „doklepany” ekson „Ex” z transkryptu odległego chromosomu „b”. Różne cząsteczki mRNA, powstałe w wyniku alternatywnego składania, zostają przetłumaczone (translacja) na różne polipeptydy, które mogą mieć różne funkcje

nie miejsce w mózgowiu i połączyć się z odpowiednimi komórkami nerwowymi. Polipeptydy DSCAM biorą udział w złożonych kompleksach, umożliwiających sterowany ruch eksonów komórki nerwowej w rozwoju osobniczym.

Dobrze poznany u muszki owocowej homologiczny gen *Dscam* posiada 115 eksonów. Potencjalnie alternatywne składanie pre-mRNA *Dscam* może generować 38 016 różnych cząsteczek mRNA, a zatem tyleż różnych izoform polipeptydów (Schmuckler i in., 2000). Dotychczas zdefiniowano kilka tysięcy takich izoform. Geny *Dscam* należą do rodziny genów immunoglobulinowych, które również wykazują wielką zdolność rekombinowania. U muszki owocowej polipeptydy *Dscam* stanowią także składową systemu odporności. W komórkach systemu nerwowego i gdzie indziej repertuar alternatywnych polipeptydów *Dscam* zmienia się w czasie i jest elementem pozwalającym odróżnić sąsiadujące komórki (np. w różnych typach fotoreceptora). Transkrypty genów kodujące inne rodziny białek wciągniętych w proces neurogenezy podlegają również alternatywnemu składaniu (Neves i in., 2004). Składanie alternatywne przejawia specyficzną narządową (Xu, Modrek i Lee, 2002), zmienia się w stadiach rozwojowych, a także posiada pewne cechy swoistości gatunkowej. Szacuje się, że proces alternatywnego składania obejmuje 35-59% genów człowieka. Proces regulowanego składania zachodzi w trakcie tak podstawowego zjawiska, jak determinacja płci u muszki owocowej (zob. Alberts i in., 1994). W wyniku alternatywnego składania „informacja pierwotna” genu zostaje przetworzona na „informację wtórną”. Alternatywne składanie pre-mRNA może obejmować także początek i koniec cząsteczki, zawierające sekwencje nie ulegające translacji, ale wpływające na wydajność translacji, stabilność cząsteczki mRNA lub jej komórkową lokalizację.

Przypuszcza się, że w trakcie ewolucji zachodził proces generowania lub utraty eksonów. Proces ewolucyjny „eksperymentował” i poddawał próbie nowe funkcje alternatywnie złożonych sekwencji, utrzymując podstawowy, stary zapis. Zjawisko alternatywnego składania jest w odniesieniu do genów człowieka częste i komórka ludzka może generować wiele alternatywnych form białka o potencjalnie różnej strukturze i funkcji. Jednak mamy mało dowodów na to, że alternatywne składanie daje wiele nowych funkcji enzymatycznych i strukturalnych (Tress i in., 2007).

Niekiedy transkrypcja obejmuje dwa sąsiadujące geny i wówczas powstaje białko chimeryczne oraz jego liczne izoformy (Akiva i in., 2006).

Redagowanie mRNA

Proces redagowania RNA (*RNA editing*) jest złożony ze skomplikowanych mechanizmów biochemicznych, zmieniających na poziomie mRNA treść pierwotnej informacji odczytanej z sekwencji nukleotydowych w DNA (Smith, Gott i Hanson, 1997). Proces ten poznano u bakterii, wirusów, pierwotniaków, grzybów, drożdży, roślin,

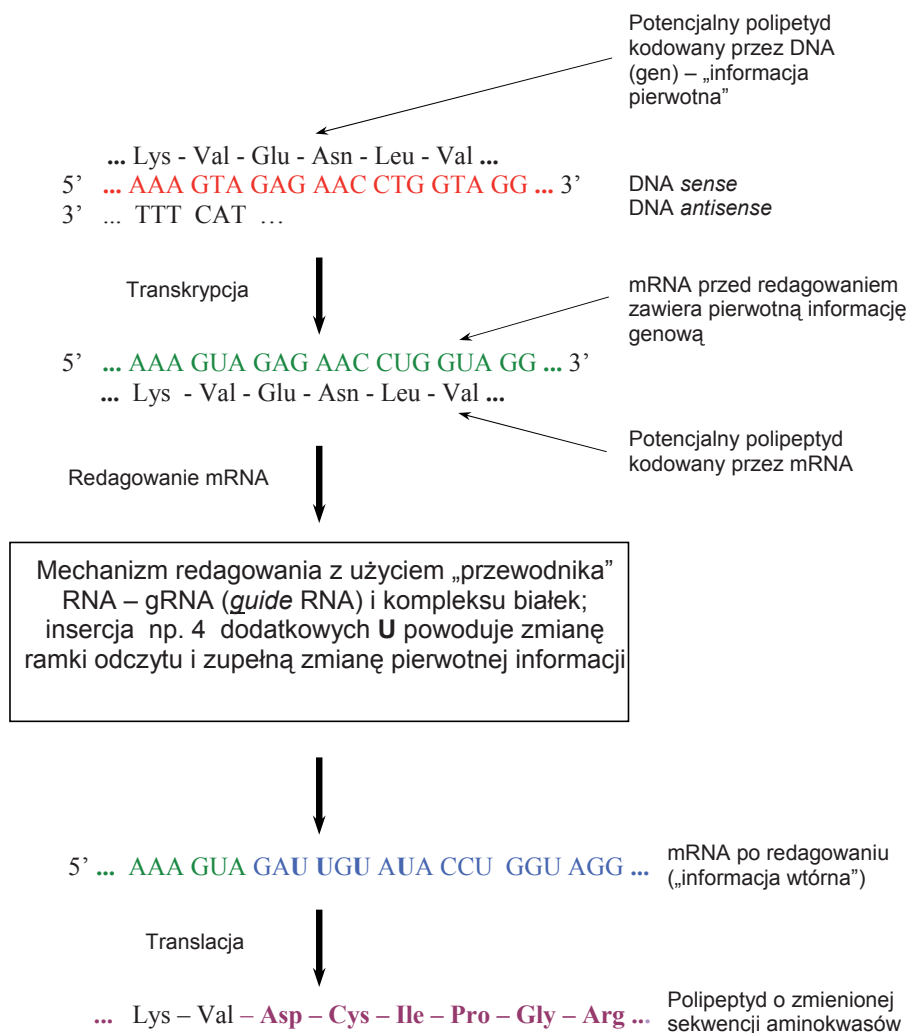
insektów i ssaków. Proces dotyczy też samego DNA, ale jest tu raczej wyjątkiem (w limfocytach B). Redagowanie RNA obejmuje złożone mechanizmy insercji i/lub usuwania (delecji) cząsteczek urydyny (U) w sekwencji mRNA. Prosty mechanizm redagowania polega na chemicznym przekształceniu w cząsteczce mRNA w określonych pozycjach kodonów C w U i A w I (I = inozyna, rzadki nukleotyd, gdy znajdzie się w kodonie odczytywanym jest jako G).

Redagowanie mRNA intensywnie zachodzi w mitochondriach *Trypanosoma* i w chloroplastach roślin. Zachodzi również u zwierząt, co umożliwia generowanie narządowych izoform białek. Redagowanie występuje też w syntezie niektórych białek, związanych z funkcją kanałów wapniowych oraz neuroprzekazników w mózgu człowieka. W tych ostatnich przypadkach często zachodzi podstawienie aminokwasu przez redagowanie mRNA z udziałem deaminaz adenozyne, zmieniających A w I. Utrata lub deregulacja funkcji deaminaz adenozyne idzie często w parze z zaburzeniami neurologicznymi. Proces zmiany A w I spostrzegano także na granicy intron/ekson, co w efekcie skutkowało tworzeniem nowej formy mRNA.

Konsekwencje redagowania mRNA są głębokie. Dla przykładu, w komórkach wątroby człowieka gen *APOB* daje transkrypt mRNA, kodujący białko apolipoproteiny B o długości 4 563 aminokwasów. Taki sam mRNA (jako transkrypt genu *APOB*) powstaje w komórkach jelita, ale koduje małą cząsteczkę apolipoproteiny, zawierającą jedynie 2 152 aminokwasów. Jest to wynik redagowania mRNA w komórkach jelita, polegający na podstawieniu U w miejsce C w środkowej części cząsteczki mRNA. Podstawienie zachodzi w kodonie CAA (kodującym aminokwas glicynę), co przekształca CAA w kodon UAA, który jest sygnałem „stop” dla procesu translacji, w wyniku czego syntetyzuje się jedynie połowa cząsteczki białka. Proces ten zmienił pierwotną treść informacji w genie, co w konsekwencji dało dwie izoformy cząsteczki białkowej, różniące się wielkością i funkcją – apolipoproteina B wątrobową jest niezbędna dla transportu cholesterolu, podczas gdy apolipoproteina jelitowa bierze udział w absorpcji lipidów.

W mitochondrialnych mRNA u pierwotniaków *Trypanosoma* podczas redagowania zachodzi intensywna insercja i delecja U przy użyciu cząsteczek „RNA-przewodnika” (*guide RNA*, *gRNA*) (Stuard i Panigrahi, 2002). Jest to wysoce złożony mechanizm, którego zasadę ilustruje Rycina 6. Nie jest jasnym dlaczego, jak się wydaje, intensywne redagowanie zachodzi głównie w mitochondriach.

Redagowanie zachodzi również w rRNA i tRNA. W przypadkach redagowania mRNA, proces ten wywiera głęboki efekt na pierwotną strukturę białka, wprowadzając sygnał „stop” kończący translację wcześniej niż to zakładała informacja pierwotna w genie (DNA), lub wprowadzając zmianę ramki odczytu. Podobnie głęboki wpływ na pierwotną strukturę białka mają zmiany w antykodonie tRNA.



Ryc. 6. Uproszczony schemat redagowania mRNA. W wyniku redagowania powstaje wtórna informacja w mRNA, dająca po translacji polipeptyd odmienny od zakodowanego w genie. Insercję nowych U w czasie redagowania zaznaczono wytłuszczoną czcionką

Uniwersalność zjawiska redagowania pre-mRNA zapewnia możliwość generowania wielu nowych izoform białka i przypuszczalnie ma jeszcze wiele innych funkcji biologicznych. Dla przykładu, dzięki mechanizmom redagowania, elementy powtarzające się *Alu* (por. wyżej), występujące w wielkich liczbach w genomie człowieka, podlegają głębokiej przebudowie. W wyniku takich procesów mogą powstawać nowe sekwencje typu eksonów, które po „wyjściu” z sekwencji *Alu* (eksonizacja) zasilają repertuar eksonów w genomie.

Redagowanie ma również miejsce w pri-microRNA przez mechanizm zamiany A na I i jest tkankowo swoiste. Takie redagowanie przenosi swoistość tłumienia ekspresji na inny zestaw genów (Kawakara i in., 2007).

Nie wiemy, co kieruje wysoce swoistym procesem redagowania RNA, ani jak ten proces jest regulowany i koordynowany, jak zachodzi w procesie rozwoju i jaka jest jego swoistość. Zjawisko redagowania RNA, podobnie jak składanie mRNA, narusza nasze postrzeganie genu jako autonomicznego odcinka DNA, dziedzicznego i niezmiennego, jako jedyne źródła bezpośredniej

instrukcji dla syntezy struktury pierwszorzędowej polipeptydu.

Gene sharing

Termin *gene sharing* (Piatigorsky, 2007) oznacza sytuację, gdy jeden i ten sam gen (DNA) jest używany wspólnie dla syntezy jednego polipeptydu, który w zależności od kontekstu (zapotrzebowania, stężenia polipeptydu w tkance) spełnia zupełnie różne funkcje molekularne. Termin *gene sharing* trudno jest przetłumaczyć na zwięzły termin w języku polskim. Może termin „dzielenie się genem” najbardziej dokładnie oddaje jego sens. W teorii ewolucji panowało przekonanie, że oprócz „korzystnych” mutacji, które wytrzymywały presję selekcyjną, podstawowym mechanizmem wprowadzającym „nowości” ewolucyjne (tj. nowe białka o nowych funkcjach) była duplikacja genów. Zrobienie drugiej kopii genu umożliwiłoby zatrzymanie funkcji „starego” genu, podczas gdy nowa kopia może służyć dla wprowadzenia innowacji w strukturze i funkcji nowego białka. Odkrycie zjawiska „dzielenia się genem” wskazuje na inny możliwy mechanizm

innowacyjny: gen może nabywać drugą funkcję bez konieczności duplikacji genu i bez utraty jego funkcji pierwotnej. Co więcej, „dzielenie się genem” i multifunkcjonalność białek istniały prawdopodobnie wcześniej niż zjawisko duplikacji genów. Kilka przykładów zjawiska *gene sharing* przytaczam za Piatigorsky’em (Piatigorsky, 2007).

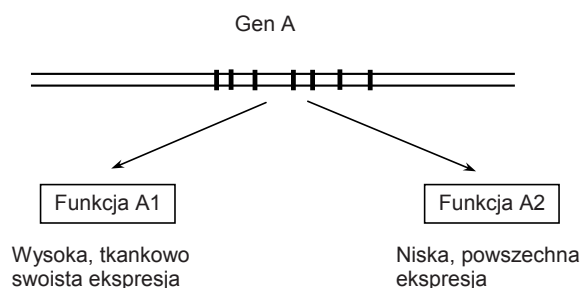
Pod koniec lat 80. odkryto, że u kaczki, kury i żółwia krystaliny – białka soczewki oka spełniające funkcje optyczne (refrakcja), pełnią w innych narządach funkcje enzymatyczne. Wkrótce okazało się, że delta2 krystalina soczewki oka kury jest białkiem identycznym z liazą/liazą argininosukcynylową, a eta-krystalina soczewki kaczki pełni także funkcję enzymu dehydrogenazy mleczanowej B. U ptaków i gadów są duże różnice międzygatunkowe w budowie krystalin. Krystaliny wywodzą się od drobno-cząsteczkowych białek szoku cieplnego i u wielu gatunków zatrzymują tę funkcję nawet w samej soczewce (jako białka chaperonowe). U kałamarnicy krystalina ma jednocześnie funkcję transferazy-S-glutationowej, u innych mięczaków – funkcję dehydrogenazy aldehydowej.

Białko gelsolina, w wysokich stężeniach działa jako regulator szkieletu aktynowego rogówki u ryby danio, ale w niskich stężeniach działa jako czynnik modulujący formowanie się wzoru brzuszno-grzbietowego w czasie rozwoju embrionalnego. Rodopsyna u muszki owocowej posiada funkcję fototransmittera i jednocześnie organizatora cytoszkieletu aktynowego dla morfogenezy fotoreceptora. Białko ksantynooksydoreduktaza, enzym katabolizmu purynowego, ma również inną funkcję. Używane jest do budowy otoczki wydzielającej kropelki mleka w gruczołach mlecznych. Syntaza cytrynowa ma funkcje metaboliczne, a także funkcje strukturalne w cytoszkielecie.

Te przykładowe białka enzymatyczne mają funkcję białek o właściwościach refrakcyjnych i innych strukturalnych wówczas, gdy ich stężenie jest odpowiednio wysokie, czyli dwie różne właściwości białka wynikają z różnic w ekspresji genu i jego lokalizacji tkankowej/narządowej. Co więcej ten sam polipeptyd może znajdować się w różnych przedziałach komórki, gdzie przejawia różne funkcje w zależności od interakcji z innymi cząsteczkami białkowymi i lokalnymi kompartmentu.

Zjawiska te są przykładem zmian struktury i funkcji polipeptydu w zależności od jego stężenia, mikrośrodowiska w obrębie komórki i miejsca syntezy w tkance/narządzie. Należy ostrożnie podchodzić do dogmatu, że w każdym miejscu i w każdym czasie białko będzie zawsze miało tę samą funkcję, jaką wykazuje w danym momencie i danym miejscu, gdyż „funkcja genów i białek zależy od kontekstu” (Piatigorsky, 2007) (Ryc. 7). Zjawisko „dzielenia się genem” może być zjawiskiem trwałym, wiecznym, lub też może być zdarzeniem przejściowym w procesie ewolucji („testowanie” nowych funkcji białka).

Piatigorsky dyskutuje na temat dość subtelnych różnic między zjawiskiem *gene sharing* a pleiotropią. Pierwsze ma miejsce, gdy gen syntetyzuje białko o dwóch oddzielnych funkcjach, drugie zaś oznacza raczej sytuację, gdy gen „kontroluje” różne cechy fenotypowe. Tak



Ryc. 7. Dzielenie się genem (*gene sharing*). Jeden i ten sam gen koduje jeden polipeptyd, który w zależności od stężenia, miejsca i kontekstu pełni odmienne funkcje molekularne

dla przykładu liaz argininosukcynylowa jest enzymem biorącym udział w syntezie mocznika u ssaków, argininy u ptaków, ale jednocześnie pełni funkcję białka refrakcyjnego w soczewce oka. Natomiast enzym, syntaza tlenu azotu, choć jest również syntetyzowany „na podstawie” informacji jednego genu, bierze udział w wielu różnych zjawiskach fenotypowych, jak np. procesy zapalne, neurotransmisja, oddychanie, apoptoza, itd.

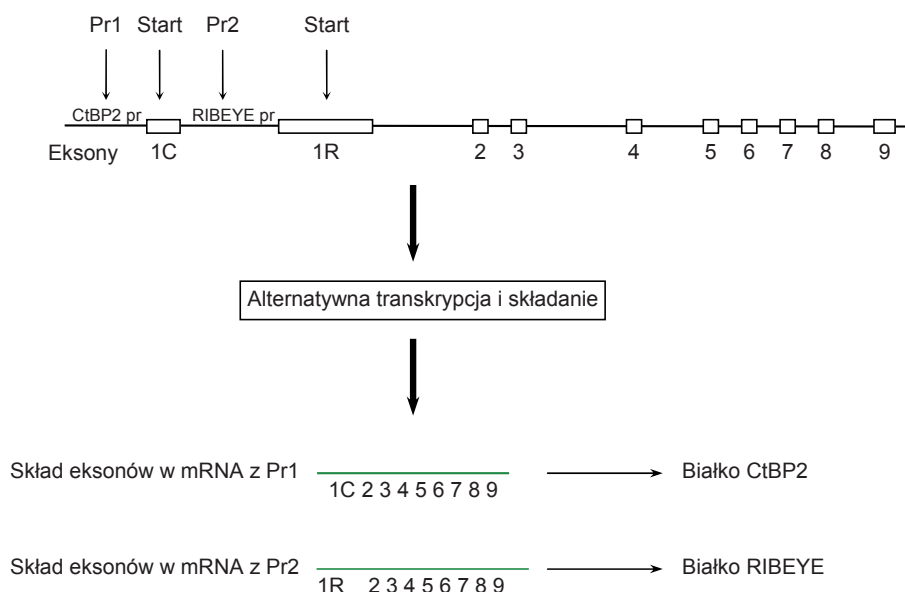
W żywych systemach polipeptydy (białka) zwykle wykonują liczne funkcje. Enzym GAPDH (dehydrogenaza glicerolaldehydo-3-fosforanowa) jest przykładem białka o wielu funkcjach – jest enzymem glikolitycznym, białkiem powierzchniowym wiążącym się z innymi białkami, posiada rolę w transdukcji sygnału, w soczewce występuje jako krystalina, jest koaktywatorem innych, białek, itd. Podobnie wielofunkcyjna jest enolaza – inny enzym działający jako białko szoku cieplnego, mające rolę w organizacji mikrotubul, które wiąże aktynę i plasminogen, a w dodatku jest również krystaliną.

Zasadnicza nauka, jaka wypływa ze zjawiska „dzielenia się genem”, to przypomnienie, że ani gen, ani białko nie są elementami autonomicznymi, działającymi w sposób niezależny, izolowany, ale są elementami dynamicznego systemu (sieci).

Promotory alternatywne

Promotory są swoistymi sekwencjami DNA, które umiejscawiano zwykle przed początkiem genu strukturalnego. Stanowią one elementy niezbędne dla inicjacji transkrypcji i kontroli ekspresji genów. Okazało się jednak, że ich położenie w stosunku do genu w organizmach eukariotycznych jest bardzo różne, a ponadto komórka używa ich alternatywnie, przez co zwiększa się plastyczność ekspresji genów w organizmie wielokomórkowym. Transkrypcja inicjowana z alternatywnych promotorów różni się poziomem, a transkrypty (a zatem i kodowane przez nie białka) wykazują różne izoformy. Alternatywne promotory wykazują swoistość tkankową i różnią się odpowiedzią na system sygnałów (Ayoubi i Van de Ven, 1996).

Użycie w procesie transkrypcji alternatywnych promotorów jest jeszcze innym mechanizmem generującym różnorodność mRNA i białka. Transkrypcja DNA (genu) zaczyna się od sekwencji zwanej promotorem, mieszczącej się zwykle na początku genu. Okazuje się, że gen może posiadać kilka miejsc promotorowych.



Ryc. 8. Alternatywne użycie promotorów Pr1 lub Pr2 oraz alternatywne składanie daje dwa różne mRNA i dwa białka CtBP2 i RIBEYE o różnej lokalizacji i różnej funkcji. Rycina jest dalece uproszczonym schematem wzorowanym na Ryc. 3.1 w monografii Piatigorsky'ego (Piatigorsky, 2007)

Przez użycie alternatywnych promotorów uzyskuje się nowe, zróżnicowane i elastyczne narzędzie kontroli ekspresji genów: możliwość generowania białek różniących się u początku cząsteczki, kontrolę poziomu transkrypcji, swoistość tkankową, swoistość związaną ze stadium rozwojowym, możliwość kontroli wydolności translacji i czasu trwania cząsteczek mRNA. Regulacja genów przez mechanizm alternatywnych promotorów jest szeroko rozpowszechniona u organizmów żywych.

U kręgowców wiele genów posiada alternatywne promotory używane w zależności od narządu, tkanki, stanu fizjologicznego komórki, stadium rozwoju, zdolności komórki do odpowiedzi na szczególne stany mikrośrodowiska i odpowiedzi na warunki metaboliczne. Rozmieszczenie cząsteczek białkowych w komórce, poziom białka i jego funkcja jest także uwarunkowana użyciem odpowiedniego promotora. Wprowadzenie do zagadnienia użycia alternatywnych promotorów znajduje się w przeglądowej pracy (Ayoubi i Van de Ven, 1996).

Użycie alternatywnych promotorów skutkuje głębokimi zmianami fenotypowymi. „Alternatywne miejsca inicjacji transkrypcji mogą spowodować transformację intronów w eksony kodujące białko i wytwarzać białka o różnej strukturze i funkcji” (Piatigorsky, 2007). Przytaczam ten cytat ze świetnej monografii Piatigorsky'ego, gdzie podany jest przykład genu kodującego dwa odrębne białka w zależności od użycia alternatywnych promotorów (Ryc. 8). Użycie pierwszego promotora genu *CtBP2/RIBEYE* daje białko CtBP2 – posiadające funkcję powszechnego represora transkrypcji. Użycie alternatywnego, drugiego promotora oraz alternatywnego składania daje białko RIBEYE występującego w synapsach siatkówki, w ślimaku ucha i przysadce. W wyniku użycia alternatywnych promotorów z „informacji” zapisanej w jednym genie powstają dwa różne rodzaje białka o różnej lokalizacji i różnych funkcjach.

Poprzez użycie alternatywnych promotorów uzyskuje się nowe, zróżnicowane i elastyczne narzędzie kontroli ekspresji genów, możliwość generowania białek różniących się na początku cząsteczki, kontrolę poziomu transkrypcji, swoistość tkankową, swoistość związaną ze stadium rozwojowym, możliwość kontroli wydolności translacji i czasu trwania cząsteczek mRNA. Regulacja genów przez mechanizm alternatywnych promotorów jest szeroko rozpowszechniona u organizmów żywych.

U kręgowców wiele genów posiada alternatywne promotory używane w zależności od narządu, tkanki, stanu fizjologicznego komórki, stadium rozwoju, zdolności komórki do odpowiedzi na szczególne stany mikrośrodowiska i odpowiedzi na warunki metaboliczne. Rozmieszczenie cząsteczek białkowych w komórce, poziom białka i jego funkcja są także uwarunkowane użyciem odpowiedniego promotora.

Nonsense-mediated mRNA decay

W prawidłowych komórkach translacja mRNA w łańcuch białkowy biegnie od sygnału inicjującego „start” (kodon AUG) do sygnału kończącego „stop” (kodony UAA, UAG, UGA, zwane także kodonami *nonsense*). Wskutek mutacji punktowej lub mutacji z przesunięciem ramki odczytu, w mRNA mogą być generowane kodony *nonsense*, będące przedwczesnymi sygnałami terminacji translacji (PTC – *premature termination codons*). Ich obecność uniemożliwia dokończenie syntezy cząsteczki białka, w wyniku czego powstają krótsze („przycięte”) cząsteczki białka. Szacuje się, że tak uszkodzone białka mają związek z jedną trzecią chorób genetycznych, w tym również nowotworowych.

Komórki eukariotyczne posiadają zdolność rozpoznawania i degradowania takich transkryptów mRNA, w których pojawiają się przedwczesne sygnały terminacji

translacji (PTC). System degradujący cząsteczki mRNA posiadające sygnały PTC nazywa się systemem NMD (*nonsense-mediated mRNA decay*). System NMD radykalnie redukuje poziom mRNA z przedczesnymi sygnałami PTC, dzięki czemu synteza uszkodzonych (skróconych) cząsteczek białka spada.

Prawdopodobnie system NMD ma funkcje bardziej rozległe niż tylko rodzaj „kontroli jakości”, możliwe, że reguluje „szum genomowy” i nadzoruje aktywność szerokiego spektrum genów (kinazy białkowe, fosfatazy, czynniki transkrypcyjne, geny biorące udział w metabolizmie aminokwasów), a także bierze udział w mechanizmach składania mRNA. System NMD nadzoruje także niefunkcjonalne transkrypty z retrowirusów i transpozonów, trzymając je w uśpieniu i jest aktywnym składnikiem regulacyjnym aktywności genowej (Mendel i in., 2004; Alonso, 2005). System NMD jest ewolucyjnie konserwowany i występuje powszechnie u drożdży, robaków, roślin, owadów i u ssaków (Maquat, 2005). Jego rola w rozwoju osobniczym i ewolucji oraz znaczenie w klinice nie jest w pełni poznane (Kuzmiak i Maquat, 2006; Madghalchi i in., 2001).

Inaktywacja genów supresorowych jest związana z dziedziczną formą nowotworów. Większość sporadycznych nowotworów nie wykazuje inaktywacji tych genów. Głęboka mutacja lub utrata jednej kopii genu supresorowego, np. genu *RB* w komórce germinacyjnej, ma charakter recesywny, a zatem nie wywołuje wyraźnego efektu fenotypowego, czyli nie wiąże się z powstawaniem groźnego nowotworu oka – siatkówczaka, ponieważ pozostaje druga czynna kopia (drugi allel). Utrata heterozygotyczności czyli zniesienie funkcji drugiej kopii (drugiego allelu) genu supresorowego (tu genu *RB*) prowadzi do powstania dziedzicznej formy siatkówczaka. To zjawisko znane jest w raku dziedzicznym i odpowiada hipotezie „dwóch trafień” Knudsona. Najczęstszym sposobem unieczynniania allelu w genie *RB* jest nabycie mutacji nonsensowych, które generują kodony przedczesnej terminacji translacji – PTC (Holbrook i in., 2004). W genie *BRCAl* germinalne mutacje typu PTC (w liczbie około 30) zidentyfikowano w 22 kodonach. System NMD degraduje większość mRNA posiadających PTC, ale jednocześnie obserwuje się obecność białek „przyciętych”. Przypuszczalnie pochodzą one z PTC zlokalizowanych blisko kodonu inicjującego, gdyż ta lokalizacja PTC (podobnie jak lokalizacja blisko kodonu terminującego) wymyka się spod nadzoru systemu NMD (Perrin-Vidoz i in., 2002).

NMD jest skomplikowanym systemem i funkcją kompleksu białek kodowanych przez geny z rodziny *Rent1* (*Upf1*), które mają zdolność rozpoznawania mRNA posiadającego kodon PTC („stop”), z wyjątkiem lokalizacji w eksonie końcowym, lub w pobliżu początku genu.

System NMD nie narusza jednak redagowanych cząsteczek mRNA, w których celowo powstał nowy sygnał „stop”, skracający cząsteczkę białka, jak w podanym wyżej przykładzie jelitowej apolipoproteiny B.

Obecność systemu NMD nasuwa konieczność zachowania ostrożności w interpretacji „ekspresji genów” przy użyciu mikromacierzy DNA.

Prof. dr hab. n. med. Mieczysław Chorąży

Centrum Badań Tranlacyjnych

i Biologii Molekularnej Nowotworów

Centrum Onkologii – Instytut im. Marii Skłodowskiej-Curie

Oddział w Gliwicach

e-mail: chorazy@io.gliwice.pl

Piśmiennictwo

- Akiva P, Toporik A, Edelheit S i wsp. Transcription-mediated gene fusion in the human genome. *Genome Res* 2006; 16: 30-36.
- Alberts A, Bray D, Lewis J i wsp. *Molecular Biology of the Cell*. Garland Publishing Inc., New York & London, 1994.
- Alonso C. Nonsense-mediated RNA decay: a molecular system micro-managing individual gene activities and suppressing genomic noise. *BioEssays* 2005; 27: 463-466.
- Ayoubi TAY, Van de Ven WJM. Regulation of gene expression by alternative promoters. *FASEB J* 1996; 10: 453-460.
- Gladyshev E, Meselson M, Arkipova IR. Massive horizontal gene transfer in *Bdelloid Rotifers*. *Science* 2008; 320: 1210-1213.
- Holbrook JA, Neu-Yilik G, Hentze MW i wsp. Nonsense-mediated decay approaches the clinic. *Nature Genetics* 2004; 36: 801-808.
- Jurka J, Kapitonov VV, Kohany O i wsp. Repetitive Sequences in Complex Genomes: Structure and Evolution. *Ann Rev Genomics Hum Genet* 2007; 8: 241-259.
- Kawakura Y, Zinshteyn B, Sethupathy P i wsp. Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science* 2007; 315: 1137-1140.
- Kapitonov VV, Pavlicek A, Jurka J. Anthology of human repetitive DNA. W: Meyers, RA (red.), *Encyclopedia of Molecular Cell Biology and Molecular Medicine* (Vol. 1, s. 251-305), Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA, 2004.
- Kuzmiak HA, Maquat LE. Applying nonsense-mediated mRNA decay research to the clinic: progress and challenges. *Trends Molec Med* 2006; 12, 306-316.
- Lander ES i wsp. Initial sequencing and analysis of the human genome. *Nature* 2001; 409: 860-921.
- Lewontin RC. *Biology as ideology – the doctrine of DNA*, HarperPerennial, 1991.
- Madghalchi SM, Frischmeyer PA, Mendell JT i wsp. *Rent1*, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Hum Mol Genet* 2001; 10: 99-105.
- Mendell JT, Sharifi NA, Meyers JL i wsp. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nature Genetics* 2004; 36: 1073-1078.
- Maquat L. Nonsense-mediated mRNA decay in mammals. *J Cell Science* 2005; 118: 1773-1776.
- Neves G, Zucker J, Daly M i wsp. Stochastic yet biased expression of multiple *Dscam* splice variants by individual cells. *Nature Genetics* 2004; 36: 240-246.
- Perrin-Vidoz L, Sinilnikova OM, Stoppa-Lyonnet D i wsp. The nonsense-mediated mRNA decay pathway triggers degradation of most *BRCAl* mRNAs bearing premature termination codons. *Hum Mol Genet* 2002; 11: 2805-2814.
- Pheasant M, Mattick JS. Raising the estimate of functional human sequences. *Genome Res* 2007; 17: 1245-1253.
- Piatigorsky J. *Gene sharing and evolution – The diversity of protein function*. Cambridge – Massachusetts, London – England: Harvard University Press, 2007.
- Schmuckler D, Clemes JC, Shu H i wsp. *Drosophila Dscam* is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* 2000; 101: 671-684.
- Smith HC, Gott JM, Hanson MR. A guide to RNA editing. *RNA* 1997; 3: 1106-1123.
- Stuard K, Panigrahi AK. *Molecular Microbiology* 2002; 45: 591-696.

- Tress ML, Martelli PL, Frankish A i wsp. The implications of alternative splicing in the ENCODE protein complement. *PNAS* 2007; 104: 5495-5500.
- Venter JC. The sequence of the human genome. *Science* 2001; 291: 1304-1351.
- Xu Q, Modrek B, Lee C. Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic Acids Res* 2002; 30: 3754-3766.

Zalewski Z. Mapa ludzkiego genomu – złudne (?) nadzieje i realne (?) zagrożenia. *Sztuka leczenia* 2001; 7: 65-75.

Przedruk II części artykułu „Gen strukturalny – ewolucja pojęcia i dylematy” z kwartalnika *Nauka* 2009; nr 3: 57-108.