

Magdalena FIKUS

Instytut Biochemii i Biofizyki PAN
WarszawaINFORMACJA GENETYCZNA: WYROK CZY
MOŻLIWOŚĆ*

„Mówimy o terapii genowej, jak gdyby mogła zmienić czyjś los, ale można także go zmienić, jeśli spłaci się czyjeś zadłużenie na karcie kredytowej”.

(James Watson)

Zaczęło się od serii przypadków. Wybrano 21 dawców, którzy sami określili swoją przynależność etniczną. Od każdego pobrano 130 ml krwi, od mężczyzn także nasienie, pięciokrotnie. Zakażając część limfocytów dawców wirusem Epsteina-Barra utworzono nieśmiertelne linie komórek. Z jądrowego DNA wykonano banki genomowe. Do ostatecznych badań wybrano DNA pięciu osobników, kierując się wieloma okolicznościami, m.in. jakością banku genów. DNA wybranych dawców posłużyło w firmie Celera do oznaczenia sekwencji nukleotydów ludzkiego genomu. Cała procedura była zgodna z Deklaracją Helsińską o prawach człowieka.

Weszliśmy w wiek XXI z pokaznym posagiem genetycznym. Przez pierwsze 25 lat poprzedniego wieku trwała identyfikacja komórkowych struktur nośników materiału genetycznego, które nazwano chromosomami. W połowie wieku dwu badaczy — James Watson, który nie-

*UWAGA: Tekst został zrekonstruowany przy pomocy środków automatycznych; możliwe są więc pewne błędy, których sygnalizacja jest mile widziana (obi@opoka.org). Tekst elektroniczny posiada odrębną numerację stron.

¹Referat wygłoszony na VII Krakowskiej Konferencji Metodologicznej „Konieczność i przypadek”, Kraków 8-9 maja 2001.

dawno przekroczył dwudziestkę, Francis Crick, trzydziestkę — opisując budowę chemiczną i strukturę przestrzenną kwasu deoksyrybonukleinowego, DNA, dokonało największego przyrodniczego odkrycia stulecia. Trzeba było upływu kolejnych 25 lat i badań wieńczonych licznymi nagrodami Nobla, aby opisane zostały w głównych zarysach zasady regulacji ekspresji genów, czyli regulacji intensywności reakcji prowadzących od DNA do dwu grup cząsteczek: kwasu rybonukleinowego (RNA) oraz białek. Od początku lat 70. rozpoczęła się era manipulacji genetycznych, czyli procedur wiodących do kontrolowanych zmian genetycznych, zwanych potocznie inżynierią genetyczną. Jednym z ważniejszych wyników osiągniętych w tym okresie jest opracowanie metody ustalania kolejności czterech podjednostek tworzących DNA (nukleotydów A, G, C, T), zwanej sekwencjonowaniem DNA. Sekwencja nukleotydów to zapis informacji genetycznej w języku czteroliterowym. Do końca XX wieku oznaczono sekwencje 205 naturalnych plazmidów, 185 organelli, genomów 600 wirusów, 31 eubakterii, 7 archebakterii, 1 grzyba, 2 zwierząt bezkręgowych i 1 rośliny. W lutym 2001 roku ogłoszono triumfalnie zakończenie prac dwu niezależnie pracujących zespołów naukowych, które podały sekwencje ponad 90% genomu ludzkiego. Ciekawe, wykorzystano różne sposoby przygotowania DNA, różne algorytmy obliczeniowe i... różne źródła finansowania. W każdej grupie uczestniczyło po kilkaset osób. Jedna składała się z pracowników instytucji publicznych USA (ponad 70% wyników), Wielkiej Brytanii, Japonii, Chin, Francji, Niemiec, druga finansowana była przez prywatne instytuty o amerykańskich siedzibach, TIGR i Celera. Oba programy wyszły z fundamentalnego założenia, że budowa wszystkich białek i RNA zakodowana jest w DNA. Milcząco zakładano, że po zidentyfikowaniu wszystkich cząsteczek występujących w komórce i organizmie uzyska się szczegółową wiedzę o tym, jak one działają i jak ich działanie jest regulowane. J. Collins, szef grupy publicznej, przewiduje, że ok. 2030 roku uzyskanie własnego, indywidualnego paszportu genetycznego będzie kosztować mniej niż \$1000. Sądzi również, że technologia genowa w zastosowaniach medycznych będzie udziałem jedynie mieszkańców krajów o zaawanso-

wanej technologii, co przyczyni się do dalszego rozwarstwiania świata na bogatych i biednych. Wydawać by się mogło, że odczytanie sekwencji genomu osobnika jest jednocześnie sentencją wyroku. Skoro jednak liczba i ilość kodowanych przez DNA białek podlega regulacji, a często procesy te rozpoczynają się, kończą, lub ich prędkość ulega zmianom, w wyniku oddziaływania organizmu z otoczeniem, to do realizacji genetycznego wyroku wkracza przypadek i nieoznaczoność. Geny i genom ludzki porównywano na wiele sposobów z uprzednio poznanymi w innych gatunkach. Nikogo nie zdziwiło, że znaleziono wśród nich pewną grupę genów i białek uniwersalnych, obecnych jako podstawowe genetyczne wyposażenie wszystkich organizmów. Ich istnienie wskazuje na utrwalone i zoptymalizowane w toku ewolucji procesy metabolizmu komórek. Tym bardziej zainteresowanie badaczy zwraca się nie ku informacjom o genach wspólnych z innymi stworzeniami, a ku poznaniu genów swoistych człowiekowi. Niestety człowiek położony jest daleko w sensie ewolucyjnym od dwojga poznanych uprzednio zwierząt: nicienia *Caenorabditis elegans* i muszki *Drosophila melanogaster*. Z niecierpliwością oczekiwane są zatem dane o sekwencji genomowej gryzoni (mysz, szczur), ryb i wreszcie naczelnych.² Ustalone jako charakterystycznie ludzkie cechy genomu być może dzielimy z kręgowcami, może tylko ze ssakami. A za takie uznano:

1. Tylko 1,2 — 1,4% ludzkiego DNA to obszary eksonowe, podczas gdy transkrybowany RNA zajmuje 28% genomu. Introny ludzkich genów są bardzo długie, a geny rozrzucone po genomie rzadko. Na 1 milion nukleotydów w drożdżach przypada średnio 483 geny, u poznanych bezkręgowców ok. 200, u człowieka 15. Identyfikacja genów w sekwencji ludzkiego DNA jest zatem bardzo trudna, a wyniki przeczesywania poznanej sekwencji programami komputerowymi nie zawsze wiarygodne.

2. Ponad 50% DNA człowieka stanowi kilkanaście typów sekwencji nie kodujących, powtarzających się. Ich rola fizjologiczna nie została ustalona, część z nich porównuje się do pasożytniczych częste-

²W maju 2001 w zespołach Celera oznaczono sekwencję genomową myszy, jednak ze względów komercyjnych nie podano jej do publicznej wiadomości

czek, część do symbiotycznych. Są nierównomiernie rozrzucone po genomie, niektóre mogą dać początek nowym genom i prawdopodobnie dawały w przeszłości. Dla porównania: powtarzające się sekwencje stanowią 3% genomu muszki, 6,5% genomu nicienia, i 10,5% rośliny.

3. Z wyróżnionych 1278 rodzin genów wspólnych dla poznanych już organizmów eukariotycznych tylko 94 uprzednio nie opisano. To tak, jak by obserwowany dziś kres działań procesów ewolucji nie wiele dodał do repertuaru genetycznego. Jak należało się spodziewać dla całego wielotkankowego organizmu ssaka, genom człowieka bogaty jest w geny i białka układu immunologicznego, krwionośnego, neurologicznego; a z punktu widzenia jego komórek — w białka regulujące procesy transkrypcji (syntezy RNA na matrycy DNA).

4. Wzdłuż danego chromosomu i w różnych chromosomach różna jest częstotliwość rekombinacji (wymiany fragmentów DNA), u mężczyzn rekombinacja była szczególnie aktywna w chromosomie 12, u kobiet w 16. Drastyczne różnice między płciami na poziomie genetycznym pojawiają się także w trakcie dojrzewania płciowych komórek; mutacje są 5 razy częstsze w komórkach męskich.

5. W każdym okresie życia osobniczego, choć częściej w życiu płodowym i dzieciństwie, mogą objawić się choroby warunkowane genetycznie. Ponad 1000 z nich zostało molekularnie i klinicznie zróżnicowanych. Cechy związane z zachowaniami są zawsze wielogenowe, dlatego mówi się, że geny nadają osobowości człowieka określone predyspozycje. Trudności w ścisłym zdefiniowaniu tych predyspozycji zwiększa fakt, iż nie ma dobrych sposobów na ilościowe oceny takich np. cech jak agresywność, inteligencja, podatność na stresy itd. Geny związane z zachowaniem zazwyczaj ulegają ekspresji w mózgu, najślabiej poznanym organie człowieka, nawet na poziomie fizjologicznym, a co dopiero molekularnym.

Genom jest zatem u danej osoby w zasadzie niezmiennym zapisem informacji. Ale tylko zapisem. Jej realizacja zależy od dwu ciągów bardzo złożonych procesów: transkrypcji i translacji. W wyniku transkrypcji powstają cząsteczki RNA, jedno niciowego kwasu nukleinowego, komplementarne do jednej z nici kodującego go DNA. O trans-

kryptach RNA wiemy mniej niż o DNA. Sam proces regulacji intensywności ich syntezy jest złożony i podlega m.in. wpływom warunków panujących wokół organizmu i wokół i wewnątrz danej komórki. To znaczy, że gen może, ale nie musi w danym momencie być transkrybowany, a RNA może być z różną szybkością degradowany, jeszcze zanim włączy się do procesu translacji. A kilka lat temu odkryto zadziwiające zjawisko: z tego samego pierwotnego produktu transkrypcji jednego genu, (pre-mRNA), może w komórkach eukariotycznych, dzięki różnemu wycinaniu intronów, powstawać więcej niż jedna dojrzała cząsteczka informacyjnego RNA, (m-RNA). Jeżeli tak, jeżeli eksony łączą się ze sobą w różnej kolejności, to ostatecznie uformowane różne informacyjne RNA kodują różne białka, ponieważ kolejność trójek nukleotydów stanowiących kodony określonych aminokwasów jest różna w różnych wersjach m-RNA. Ten sam gen może kodować więcej niż jedno białko, a opisany powyżej proces dojrzewania m-RNA, zwany „*alternative splicing*” (*ang.*), jest wyjątkowo intensywny w komórkach ludzkich. Liczba białek decydujących o naszej różnorodności i odrębności gatunkowej znacznie przewyższa liczbę genów.

Jeżeli chodzi o m-RNA nie koniec na tym. Kilkanaście lat temu odkryto także, że do już dojrzałej cząsteczki mRNA pewne enzymy mogą w części kodującej dodawać pojedyncze i kilka naraz nukleotydów, co podobnie do alternatywnego dojrzewania zmienia sens zapisu genetycznego, co więcej, nie ma swojego odpowiednika w sekwencji genu. Odnotowywano także dla danego mRNA przesunięcia co do miejsca startu i końca transkrypcji, a więc raz jeszcze, różne odkodowywanie informacji genetycznej zapisanej w tej samej sekwencji DNA.

Wreszcie w roli jaką odegra jako matryca informacji genetycznej dany mRNA niebagatelne znaczenie ma trwałość w komórce danej cząsteczki — jej „czas życia” może się wahać od kilku sekund do godzin. W rozważaniach o grze konieczności i przypadku w procesach ekspresji informacji genetycznej pozostał jeszcze niezwykle bogaty obszar proteomu, czyli zespołu syntetyzowanych białek, w danym gatunku, w danych tkance i komórce, w danym momencie ich

życia. Skalę problemu rozszerza jeszcze fakt, że jakkolwiek coraz bardziej wzbogaca się nasza wiedza o przestrzennych strukturach białek, to w dalszym ciągu nie można przewidzieć funkcji tych cząsteczek w oparciu o samo poznanie struktury. Ta ostatnia niechętnie odkrywa swoje tajemnice, metody badawcze są w tej dziedzinie trudne w wykonaniu, kosztowne i cały problem wymaga zastosowania kilku metod jednocześnie. Nadal przewidywanie struktury przestrzennej białka ze znajomości sekwencji aminokwasów, bez danych doświadczalnych, jest grą samą w sobie. Wiadomo, że to samo białko, w zależności od otaczających warunków, może przyjąć różne struktury przestrzenne i spełniać różne funkcje. Dramatycznym przykładem takiej sytuacji są złoży amyloidowe obserwowane w przebiegu różnych chorób neurodegeneracyjnych. W budowie wielu białek eukariotycznych wyróżnić można obszary przyjmujące pewien kształt, takie obszary nazywamy domenami. Wiele domen białek wyższych eukariotów (w tym człowieka), w niższych organizmach występuje jako odrębne białka o innej funkcji, niż wtedy gdy są częścią większej cząsteczki. Domeny mogą także występować w różnych kombinacjach, co znakomicie zwiększa liczbę rozwiązań alternatywnych. I tak np. pewna domena proteazy serynowej, enzymu trawiącego inne białka i podobnego do trzustkowej tripsyny, oddziałuje w komórkach ludzkich z 18. różnymi domenami, kompleksy te mają różne funkcje. Tę samą domenę w komórkach muszki odnajduje się w 8 kompleksach, u nicienia w 5, a w drożdżach tylko w jednym.

Toczy się spór wokół doniesień grupy Celera o obecności w ludzkim genomie 223 genów pochodzących wprost od bakterii. Są badacze, którzy twierdzą, że jest to błąd w oznaczeniach i artefakt. Po dokładnym zbadaniu tej grupy genów zostało się tylko ok. 40 genów. Wyjaśnienie tego sporu nie jest błahym problemem, choć liczba genów niewielka — wiąże się on z wielką zagadką biologii molekularnej, z pytaniem czy możliwe było w trakcie ewolucji, a więc czy możliwe jest obecnie, tak zwane horyzontalne przenoszenie genów między gatunkami. Wiadomo, iż zdarza się to między mikroorganizmami, jednak brak bezpośrednich dowodów na takie wędrówki genów mię-

dzy prokariotami i eukariotami. Gdyby były one możliwe, wzrosła by liczba argumentów przeciw manipulacjom z transgenicznymi roślinami i zwierzętami. Czy jednak jesteśmy nosicielami genów pochodzącej w prostej linii od bakterii pokażą najbliższe lata, w których weryfikowane będą doświadczenia grupy Celera. Horyzontalne przenoszenie genów mogłoby być kolejnym przykładem roli przypadku w rozwoju życia.

„Przetłumaczone” z mRNA białko podlega jeszcze modyfikacjom potranslacyjnym. Może zostać udekorowane przez wyspecjalizowane w takiej funkcji enzymy różnymi dodatkowymi grupami funkcyjnymi: acetylową, metylową, fosforanową i innymi. Te dekorujące grupy bardzo często określają dalszą aktywność i czas życia cząsteczki białka, stanowią także gatunkowy wyróżnik, zauważany przez systemy immunologiczne. Przyjmowanie określonych struktur przestrzennych, o czym już była mowa, także może łączyć się z modyfikacjami chemicznymi, z których najważniejszą jest jednoczesna redukcja bocznych odgałęzień aminokwasów, grup SH, z utworzeniem utlenionego mostka S-S, spinającego często oddalone odcinki łańcucha białkowego.

U człowieka, choć znaleziono tylko 30-40 tysięcy genów, liczbę różnorodnych białek, które może wytwarzać nasz organizm ocenia się na bliską 100 tysiącom. W związku zaś z różnicami w funkcjach różnych tkanek, również w przybliżeniu sądzi się, że w danym typie komórek powstaje w danym momencie ok. 10 tysięcy białek. No i nie budzi żadnej wątpliwości intuicyjnie przewidywalna sytuacja, że te 10 tysięcy białek wchodzi w wielorakie oddziaływania między sobą i innymi cząsteczkami obecnymi w komórce, oddziaływania dyktowane potrzebą chwili, łatwo powstające i łatwo się rozpadające. O ich istnieniu bardzo często decyduje także charakterystyczny transport substancji wewnątrz komórki i jej podzielenie na przedziały wysoko wyspecjalizowanymi błonami. Nad całością procesów realizacji informacji genetycznej czuwają swoiste czynniki sygnałowe (białka, produkty genów regulacyjnych). Nie ulega zatem wątpliwości, że procesy ekspresji genów podlegają różnorodnym modyfikacjom, przyczyn tych

zjawisk nie w pełni pojmujemy, więc też i wynik końcowy często nie daje się przewidzieć. Także metody badań doświadczalnych zespołów genów, ich transkryptów i ich końcowych produktów białkowych wymagają rozwinięcia, lub wręcz stworzenia. Przykładem może być technika tzw. „czipów” — DNA, pozwalająca na jednoczesne badanie intensywności wielu tysięcy genów w próbkach pobranych z komórek w określonym stanie fizjologicznym, od chorego, od osobników w różnym wieku itd. Wydaje się, że w biologii molekularnej jesteśmy świadkami przekraczania niezwykle ważnej granicy, dzielącej zbieranie szczegółowych faktów i tworzenie cząstkowych hipotez, od teorii uogólniających, które pozwolą spojrzeć na komórkę i organizm, kiedyś także na biosferę, jako na całości, funkcjonujące zgodnie z dającymi się zrozumieć regułami. Kończy się epoka pojedynczych badaczy, którzy całe swoje życie poświęcili badaniu jednego genu lub jednego białka.

Ubóstwo metod badawczych w stosunku do cech wielogenowych zmusza do analiz statystycznych. Przez wiele lat w sporze o cechy ludzi — warunkowane czy to przez geny, czy przez środowisko — najważniejsze dane uzyskiwano z badań jednojajowych (monozygotycznych) bliźniąt, wychowywanych wspólnie lub osobno.³ Z badań tych jasno wynika, że cechy złożone są tak różnie warunkowane (w tym obecnością różnych wersji kilkunastu, często kilkudziesięciu genów), że i ich podłoże trudno przy dzisiejszej wiedzy jednoznacznie opisać. Zaangażowane w grę geny stwarzają predyspozycję organizmu do określonych cech, nie zaś pewność, iż cechy takie się pojawią. Cały problem jeszcze bardziej komplikuje fakt (także statystyczny), że w miarę gromadzenia przez jednostkę doświadczeń życiowych wzrasta rola genów w przejawianiu się cech, a maleje wpływ środowiska. Dziedziczność nie jest synonimem niezmienności.

³Korelacja współczynnika inteligencji, IQ, z genetycznym pokrewieństwem, wyrażona w procentach wynosi dla tej samej osoby testowanej dwa razy 87, dla jednojajowych bliźniąt razem wychowywanych 86, osobno — 76, dwujajowych wychowywanych razem 55, rodzeństwa 47, rodziców i dzieci żyjących razem 40, osobno — 31, adoptowanych dzieci żyjących razem — 0. 100 oznaczało by doskonałą zbieżność, 0 — losową różnicę. Cyt. za M. Ridley, „Genom”, Rebis, 2001.

Geny są tylko rusztowaniem, na którym kształtujemy nasze życie. Genom przytrafiają się przypadkowe przygody, prowadzące do zmian, które genetycy nazwali mutacjami. Utrwalone w kolejnych pokoleniach mutacje stają się koniecznością. Ale też, jak w cytowanej już książce napisał Matt Ridley: "...cała ludzka natura jest elastycznie zaprogramowana z góry w naszych chromosomach i wyjątkowa dla każdego z nas. Każdy ma niepowtarzalną i odmienną, endogenną naturę. Swoje **ja**".

Genom ludzki opisany został w *Nature*, numerze specjalnym 6822, z dn. 15 lutego 2001 (Konsorcjum Projektu Genomu Człowieka, zespoły akademickie finansowane przez budżety wielu krajów) i w *Science*, numerze 5507 z dn. 13 lutego 2001 (prywatna firma Celera).