

Do trzech razy...

Już dwa razy ogłoszono, że osiągnięto pełne poznanie ludzkiego genomu. Po raz pierwszy, w roku 2000, wówczas do pełnej sekwencji brakowało 10%. Po raz drugi – w marcu 2022 roku, ponad 100 naukowców ze wszystkich stron świata (kto tu zasłużył na Nobla?) zasugerowało zakończenie prac. Zostało jeszcze 0,4% nieoznaczonych sekwencji, to szansa na trzecie oświadczenie. Genom (ludzki) to pełny zasób cząsteczek DNA przypadających na komórkę, wspólny dla wszystkich osobników gatunku (*Homo sapiens*).

Podstawowe składniki DNA, nukleotydy, połączone są w łańcuchy. DNA zbudowany jest z 4 różnych rodzajów nukleotydów. Kolejność ich ułożenia w DNA (sekwencja) stanowi zapis informacji o różnorodnych funkcjach komórki i organizmu jako całości. Genom człowieka zbudowany jest z ponad 3 miliardów nukleotydów. Są uporządkowane w 23 łańcuchach, każdy wchodzi w skład odrębnego „tworu” w jądrze komórkowym – chromosomu. Dodatkowo płeć odróżniają chromosomy Y (XY – mężczyźni) i X (XX – kobiety).

W roku 2015 oceniano, że różnica między konkretnym genomem a referencyjnym wynosi 0,6%, dlatego jesteśmy różni. Ustalenie sekwencji nukleotydów nazywa się sekwencjonowaniem.

Chemiczną naturę materiału genetycznego poznaliśmy 70 lat temu, sekwencjonowanie zaczęto w 1977 roku, po wynalezieniu odpowiednich metod (Nagroda Nobla w 1980 r.). Noblowskie metody były eleganckie, ale zbyt wolne i zbyt kosztowne. Pierwszą opublikowano sekwencję faga $\Theta X 174$ (5,4 tys. nukleotydów). Dalszy postęp wymagał automatyzacji procedur, a następnie ich ulepszania lub opracowania nowych. Kolejne lata przyniosły dziesiątki nowych propozycji przyspieszenia metod obliczeniowych do analizy szybko uzyskiwanych wyników sekwencjonowania. Oczywiście oznaczało to także obniżanie kosztów, wyrażane w koszcie ustalenia jednej pozycji nukleotydu w łańcuchu. Kiedy świat naukowy i finansowy uznał, że sekwencjonowanie DNA jest ważne i **konieczne** (połowa lat 80.), jedna pozycja nukleotydu w sekwencji kosztowała około 1 dolara, projekt „ludzki genom” wyceniano na 3 miliardy dolarów. Od tego czasu cenę oznaczenia takiej sekwencji obniżono do 1 tysiąca, obecnie jest dostępna dla indywidualnych zleceniodawców.

Gdy poznawano kolejne genomy innych organizmów, starano się przypisywać im funkcje. Kod genetyczny poznano w latach sześćdziesiątych. Odróżniano takie fragmenty DNA, których sekwencja kodowała syntezę drugiego typu kwasu nukleinowego – RNA, którego sekwencja z kolei kierowała syntezą białka. Tak rozumiano wówczas, czym jest gen: jest fragmentem DNA kodującym białko, te cząsteczki sprawują w komórce wszystkie funkcje życiowe. W pierwszych zsekwencjonowanych organizmach (wirusy, bakterie) cały DNA kodował białka. Sprawa skomplikowała się, gdy zabrano się do sekwencjonowania genomów organizmów bardziej złożonych, także człowieka. W 2000 roku okazało się, że białka kodowane są przez niecałe 30% genomu człowieka. A pozostałe 70%...? Rozumowanie trzeba było odwrócić: szukać funkcji dla poznawanych sekwencji. To zajęło badaczom, wyposażonym w wyszukany sprzęt i technologie, kolejnych 20 lat. W sekwencjach z 2022 roku niewiadomą jest już tylko 0,4%.

W naszym genomie większość stanowią sekwencje powtarzające się. Długie fragmenty łańcuchów, które prawie identycznie powtarzają się obok lub w innych łańcuchach DNA. Niektóre z nich to pozostałość po sekwencjach bardzo dawnych wirusowych zakażeń ludzi i ich przodków. W wyniku pomyłek w procesach replikacji (podwajania) DNA fragmenty łańcuchów mogą przenosić się z miejsca na miejsce. Nie znamy nadal wszystkich procesów towarzyszących podziałowi komórki (a więc także replikacji DNA). Są fragmenty łańcucha zabezpieczające charakterystyczne „końce” chromosomów (telomery). Wreszcie odkryto dużo rodzajów RNA kodowanych przez DNA, bardzo różnej długości (nawet tylko kilkanaście nukleotydów), które nie kodują białek – nazywamy je obecnie rDNA. Wiele cząsteczek takiego RNA odgrywa role związane z regulacją przebiegu i intensywnością różnorodnych procesów.

W 2022 roku nie umiemy zsekwencjonować DNA znajdującego się w regionie centromeru – środkowej części chromosomu, zaangażowanej w procesy podziału komórki i rozdzielania nowo utworzonych potomnych chromosomów. Ta lokalizacja sugeruje funkcję takich sekwencji. To już tylko 0,4% całego genomu.

Większość tych wniosków wysnuto dzięki opracowaniu nowej metody sekwencjonowania bardzo długich odcinków DNA (pierwotne metody pozwalały na uzyskanie sekwencji ok. 500 nukleotydów).

Dla kolejnych badaczy pozostaje zrozumienie i wykorzystanie wiedzy o genomie z zakresu: medycyny, historii naszego gatunku i jego ewolucji, odczytywania źródeł zmienności w obrębie gatunku, genomiki (porównywanie genomów w całym zakresie gatunków), dynamiki rozwoju osobniczego. Od nowa odżyją dyskusje o modyfikacjach genomów w ogóle, ludzkiego w szczególności. Jest o czym myśleć doświadczalnikom i teoretykom wszelkich rodzajów, filozofom, pisarzom, malarzom – nam wszystkim. Zapraszam!

Magdalena FIKUS (magda.fikus@gmail.com)