

Gen strukturalny – ewolucja pojęcia i dylematy

Mieczysław Chorąży¹

The structural gene – evolution of concept and controversies

W kwartalniku *Nauka* (nr 3/2009, str. 57-108) ukazał się obszerny artykuł M.Ch. pod tytułem „Gen strukturalny – ewolucja pojęcia i dylematy”. Redakcja *Nowotworów* uznała za pożyteczne 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 pierwszą wspomnianego artykułu, za zgodą Redakcji *Nauki*.

Prof. Edward Towpik
Redaktor Naczelny

Pracę tą 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.

W ostatnim półwieczu termin „gen”, powszechnie używany wśród biologów, stał się powszechny także w szerokich kręgach cywilizowanych społeczeństw. „Gen” jest terminem teoretycznym, choć współczesna nauka przyjmuje, że ma on postać materialną. Mimo swej stale powiększającej się złożoności, termin „gen” jest nieodzownym składnikiem komunikacji w różnych kręgach badaczy, wszedł do terminologii medycznej, znalazł miejsce w praktyce ekonomicznej (firmy biotechnologiczne, patenty, agrobiologia), a także w kulturze, socjologii, filozofii i polityce. Termin ten ma wysoki autorytet w dyskusjach i ocenach naukowych, a także jest często używany w potocznych debatach ludzi wykształconych. Stąd wyda-

je się pożytecznym upowszechnienie podstawowej wiedzy o tym, czym jest „gen”, jak był pojmowany półtora wieku temu i jak jest pojmowany obecnie. Dlatego też badania, które doprowadziły do współczesnego poglądu na gen jako na materialną jednostkę dziedziczności, omawiam nieco szczegółowiej.

Od wieków ludzi ciekawiło pytanie, dlaczego potomstwo zwierząt i roślin zawsze jest podobne do swoich rodziców, choć nie jest to podobieństwo absolutne. Rolnicy starych kultur (Babilon i Egipt), prowadząc hodowlę roślin i zwierząt, byli świadomi, że pewne cechy użyteczne mogą być przenoszone z organizmów rodzicielskich na potomstwo. W Biblii znalazło się ostrzeżenie wskazujące, że ludzie tamtych czasów mieli świadomość dziedziczenia cech takich jak hemofilia. Biblia wskazywała, że po śmierci dwóch synów z powodu wykrwawienia, u trzeciego syna nie wolno dokonywać obrzezania. W ubiegłych wiekach istniała świadomość rodzinnego przenoszenia różnych zaburzeń zdrowotnych (dodatkowe palce – polidaktylia, nowotwór siatkówki oka – *retinoblastoma*, samoistne krwawienie – hemofilia).

Gen jako jednostka dziedziczności – wczesne koncepcje

Przegląd wczesnych poglądów na mechanizm dziedziczenia podaję w oparciu o dzieło Michela Morange: *A History of Molecular Biology* (Morange, 1998).

U wszystkich gatunków potomstwo dziedziczy wszystko, co jest szczególne u ich rodziców – ta reguła znana była od czasów Arystotelesa, ale dopiero w XVII i XVIII wieku zaczęto formułować pogląd, że połączenie komórki jajowej, produkowanej przez samice zwierząt z nasieniem (spermą) produkowaną przez osobniki męskie może dać początek życia zarodka. Wiązało się to z odkryciem komórki jajowej, produkowanej przez osobniki żeńskie zwierząt (W. Harvey i A. van Leuwenhoek). W XVII wieku panowała **teoria preformacji** (J. Swammerdam), która głosiła, że plemniki człowieka zawierają w swojej główce uformowanego, pełnego, choć niewidocznego, miniaturowego człowieka. W czasie rozwoju osobniczego (ontogenezy) następowało więc „odkrywanie” formy, a nie progresywne „wypracowywanie” jej.

¹ Prof. em. Centrum Badań Translacyjnych i Biologii Molekularnej Nowotworów
Centrum Onkologii – Instytut im. Marii Skłodowskiej-Curie
Oddział w Gliwicach

Brzemie dziedziczenia spoczywało na ojcu. W procesie rozwojowym miały powstawać nowe formy preadaptowane już do presji środowiskowej, na którą są skazane następne generacje. Odmienne stanowisko głosiła **teoria inkapsulacji**, wskazująca na żeńską komórkę jajową jako miejsce, gdzie miały bytować miniaturowe organizmy potomne, umieszczone jedne w drugich, jak w rosyjskiej laleczce „matrioszka”, w których większa zawiera mniejszą, a ta jeszcze mniejszą, itd. (Moss i Ch. Bonnet). W istocie termin „ewolucja” zawiera w sobie pierwiastek preformistyczny (powstawanie czegoś nowego z czegoś, co już było przedtem).

Inni wyrażali pogląd, że komórki reprodukcyjne nie zawierały preformowanych embryonów, lecz cząsteczki, które miały właściwość formowania zarodka.

Według H. Spencera (1863 r.) w całym organizmie istnieją jednakowe „fizjologiczne jednostki”, odpowiedzialne za dziedziczenie cech. W roku 1868 Charles Darwin zaproponował **teorię pangenezy**, głoszącą, że każda komórka somatyczna organizmu zawiera swobodnie przemieszczające się, niehomogenne, niezależne od siebie, miniaturowe cząsteczki (gemmule), które w narządach rozrodczych tworzą rozwijającego się zarodka. Osobniki rodzicielskie mieszają owe cząsteczki w akcie poczęcia, a jednocześnie, wraz z aktem transmisji cech przenosi się na potomstwo zmienność, która to właściwość stała się kanwą dla rozwinięcia jego teorii selekcji jako siły ewolucyjnej. Zatem „gemmule” miały być odpowiedzialne za dziedziczenie cech gatunkowych i jednocześnie byłyby nośnikami zmienności. Środowisko działające na przypadkową, dziedziczną zmienność miało być sitem dokonującym naturalnej selekcji, czyli doboru osobników najlepiej przystosowanych do środowiska i jednocześnie być siłą napędową ewolucji. Teorię „pangenezy” Darwin rozwinął Hugo de Vries. Według niego komórkowe „pangeny” mogły mutować i podlegać zmienności co do liczby i rodzaju, dzięki czemu powstawały odmiany i gatunki.

W XIX wieku August Weismann zaproponował **teorię plazmy zarodkowej**. Weismann postulował wyróżnienie „plazmy zarodkowej” jako oddzielnego od ciała („soma”) trwałego komponentu organizmu żywego. Plazma zarodkowa bytowała jedynie w komórkach rozrodczych (komórka jajowa i plemnik), była niezależna od „somy” i przekazywana była z pokolenia na pokolenie za pomocą komórek rozrodczych. Plazma zarodkowa zawierała hipotetyczne elementy dziedziczne, zwane przez Weismanna „idantami”. „Idanty” składały się ze zdolnych do samopowieliania się „idów”, złożonych z kolei z „determinantów” i „bioforów”. „Idanty” były odpowiedzialne za organizację (specyfikację) organizmu i przenosiły cechy przodków na potomstwo. W procesie rozwoju „idy” rozdzielały się do różnych części zarodka. Liczba „idów” była redukowana do połowy w czasie formowania się komórek jajowych (oogeneza) i plemników (spermatogeneza). Międzyosobnicza zmienność miała wynikać z różnych kombinacji „idów” w komórkach jajowych.

Doświadczenia G. Mendla w XIX wieku dały impuls do rozwijania badań nad zjawiskami towarzyszącymi procesowi dziedziczenia, a zatem zjawiskami stojącymi

u podstaw życia. Mendel obserwował u krzyżówek groszku ogrodowego przenoszenie na potomstwo prostych cech, takich jak wysokość rośliny, kolor kwiatów lub zabarwienie i wygląd ziaren. Na podstawie swoich badań Mendel sformułował koncepcję, że za przenoszenie cech odpowiedzialne są jednostki dziedziczności, które nazwał „determinantami”, „elementami” lub „zawiązkami”. Pierwsze pokolenie roślin (F1), powstałe z krzyżówki roślin o różnych cechach (np. roślina wysoka x roślina niska, kolor kwiatów czerwony x kolor biały), ma cechę jednolitą, dominującą (wysoki wzrost, czerwona barwa kwiatów), ale już w pokoleniu następnym (F2), powstałym z krzyżówek między osobnikami F1, następuje rozszczepienie cech: powstają rośliny o wysokim i niskim wzroście (lub o czerwonych i białych kwiatach). Cechy takie, jak niski wzrost i biała barwa kwiatów, występują w populacji F2 w 25%. Dokonując następnych krzyżówek roślin z jedną lub dwiema cechami i analizując statystyczny rozkład cech w potomstwie Mendel doszedł do przekonania, że rozmieszczenie cech w potomstwie nie jest zjawiskiem przypadkowym, lecz podlega pewnym regułom. Za każdą cechę jest odpowiedzialna para elementów (alleli). Na podstawie takich statystycznych analiz Mendel sformułował prawa dziedziczenia i rozwoju. W czasie formowania komórek rozrodczych następuje segregacja cech – pojedyncze jednostki (allele) z pary czynników determinujących poszczególne cechy rozchodzą się i przechodzą do komórek płciowych, a po zapłodnieniu allele łączą się ponownie w pary. Komórki płciowe mają po 50% determinantów cech (alleli) każdego z rodziców. Jednostki te (allele) mogą mieć charakter dominujący, recesywny lub mogą występować w postaci mieszanej. Allel Mendla był stanem, warunkiem lub procesem, który wpływał na kształtowanie właściwości komórkowej albo na ukierunkowanie rozwoju w bardzo swoisty, choć niejasny sposób. Badania Mendla pozostały w zapomnieniu do 1900 r., gdy ponownie je „odkryto” (Hugo de Vries, Erich von Tschermak i Carl Correns).

Wprowadzenie do biologii terminu „gen” (z greckiego *genos* – początek) przypisuje się duńskiemu badaczowi Wilhelmowi Johannsen’owi, który w roku 1909 użył tego terminu dla określenia „czegoś” nieokreślonego, neutralnego, co bytuje w komórce (organizmie) i określa pewne aspekty fenotypu. Gen nie miał żadnej formy materialnej i był używany jako termin operacyjny na użytek badań nad dziedzicznością. „Gen” nie został zdefiniowany, lecz zastąpił wszystkie poprzednie nazwy „jednostek”, „elementów”, „czynników”, które miały mieć określoną strukturę lub funkcję w procesie dziedziczenia i rozwoju. Johannsen wskazał też na zasadnicze różnice między genotypem a fenotypem. Termin „genotyp” odnosił się do przenoszonych z pokolenia na pokolenie zbioru par czynników („genów”), z których jeden pochodził od ojca, a drugi od matki. „Fenotyp” określał całkowity zespół cech i ich różnych form występujących u osobników dojrziałych. Fenotyp był czymś realnym, mierzalnym, dostępnym do obserwacji. „Geny” (i genotypy) były czymś niewidocznym, niedostępnym do obserwacji, nie było więc wiedzy o ich wewnętrznych właściwościach.

Do badań nad dziedzicznością Thomas H. Morgan użył muszki owocowej (*Drosophila melanogaster*) ze względu na jej szybki cykl życiowy, możliwość uzyskiwania krzyżówek i łatwe śledzenie cech fenotypowych. Szczególnie interesujące były badania Morgana nad sprzężeniem oraz wymianą fragmentów chromosomów (rekombinacja) w procesie tworzenia komórek płciowych. Geny miały być zorganizowane w formie liniowej. Im bliżej siebie leżały geny w chromosomie, tym rzadziej występowała ich rekombinacja w procesie *crossing-over*. Jednocześnie chromosomy olbrzymie ślinianki muszki owocowej, posiadające morfologiczne znaczniki struktury (jasne i ciemne pasma poprzeczne) i funkcji (puffy), posłużyły Morgonowi do badań zależności między zmianami w genach i relacją tych zmian do znaczników. Badania te pozwoliły na sporządzenie fizycznej mapy chromosomów ślinianki i porównanie jej z mapą uzyskaną z badań nad rekombinacją.

Odkrycie chromosomów (W. Flemming i E. Strasburger) pozwoliło na poznanie zjawiska mitozy i mejozy. Terminem mitozy określa się podział komórki somatycznej na dwie identyczne komórki potomne. Mitozę poprzedza duplikacja chromosomów, a następnie rozmieszczenie ich równoważników do komórek potomnych. Mejoza jest zjawiskiem złożonym, mającym miejsce w czasie formowania komórek płciowych. W czasie mejozy następuje zredukowanie liczby chromosomów do połowy w każdej komórce płciowej obu rodziców, tak jak to postulował Mendel dla swoich „czynników” lub „allel”.

W wyniku badań Mendla i Morgana sformułowano **chromosomalną teorię dziedziczności**. Według tej teorii geny były jednostkami dziedziczności ułożonymi wzdłuż chromosomu, złączonymi ze sobą jak paciorki różańca i mającymi względem siebie stałą pozycję – *locus*. Sporządzone zostały mapy sprzężeń genów, to jest oznaczono względne odległości, w jakich geny rozłożone są względem siebie. Dotychczas nieokreślone pojęcie genu, ujmowane jako stan lub proces, sprowadzono do nadal jeszcze niezdefiniowanego „czynnika”, ale czynnika zlokalizowanego w materialnym chromosomie czyli morfologicznej, dostępnej do obserwacji strukturze komórki. Morgan przewidział, że indywidualny gen może spełniać wiele funkcji i dawać szeroki wachlarz efektów. Dalej domniemywał, że każdy gen może mieć swoiste działanie (efekt) w danym narządzie, ale także swoiste w innym narządzie, a w wyjątkowych przypadkach we wszystkich narządach i w cechach całego organizmu. Te myśli, jak opóźnione echo, wracają po dziesiątkach lat do współczesnej biologii.

Uczeń Morgana, Herman J. Muller (późniejszy laureat Nagrody Nobla) przy pomocy promieni rentgenowskich i związków chemicznych uzyskał różne fenotypy muszki owocowej, przypisując te nowe fenotypy uszkodzeniom (mutacji) genów. Muller uważał geny za „fundament życia” – stabilne struktury, do których są „doczepione” inne elementy organizmu.

Stworzenie podstaw cytogenetyki zawdzięczamy Walterowi Suttonowi. Głosił on, że chromosomy są

nośnikami cech dziedzicznych i w 1903 r. sformułował podstawowe kanony dziedziczenia chromosomalnego: wszystkie organizmy w komórkach somatycznych zawierają podwójny komplet chromosomów; komórki rozrodcze (plemnik i komórka jajowa) zawierają tylko pojedynczy komplet chromosomów – ojca lub matki, a zapłodniona komórka jajowa (zygota) ponownie posiada podwójny komplet chromosomów (ojcowski i matczyne). Z zygoty powstaje organizm potomny o podwójnym komplecie chromosomów, podobnie jak u ojca i matki.

Termin „genetyka” jako nauka o dziedziczności został wprowadzony przez Williama Batesona (początek XIX wieku). Bateson sądził, że istnieją „jednostki właściwości” (niekiedy nazywał je również „genami”), występujące w antagonistycznych parach. Te jednostki nazwał „allelomorfami” (*allelomorph* = inna forma). Zapłodnioną komórkę jajową (zygota), powstającą ze złączenia się w parę podobnych allelomorfów matczynej gamety (komórka jajowa) i ojcowskiej gamety (plemnik), nazywamy homozygotą. Jeśli w gametach rodzicielskich wystąpią allelomorfy niepodobne do siebie, po akcie zapłodnienia powstanie heterozygota. Do dziś pozostał termin „allel” jako alternatywny stan genu strukturalnego.

„Gen” Batesona i „gen” Johannsena nie miały określonej materialnej struktury. Geny były hipotetycznym „czymś” (czynniki, procesy?), co dawało zdefiniowane właściwości (cechy) i co było odpowiedzialne za przekazanie cech rodzicielskich potomstwu.

Na początku XIX wieku nie było dużego zainteresowania fizyczną strukturą genów. Jedyne H. J. Muller podejmował badania nad zmiennością mutacji wywołanych u *D. melanogaster* promieniami Rentgena, sądząc (podobnie jak Max Delbrück), że analiza zmienności mutacji w relacji do energii radiacyjnej pozwoli na zdefiniowanie pewnych cech fizycznych (np. wielkość) genu.

Nowa dyscyplina nauki, genetyka, uważana była za naukę jedynie użyteczną w agronomii, zwłaszcza w hodowli roślin i zwierząt, czego wyrazem było powołanie zwłaszcza w St. Zjednoczonych zespołów i instytucji zajmujących się tą dziedziną. Rozwój genetyki populacyjnej był utrudniony przez fakt, że cechy osobnicze ilościowe mają zwykle charakter wielogenowy (i wieloczynnikowy).

Powiązania genetyki z embriologią były luźne i przez długi czas istniał wyraźny rozdział między tymi dyscyplinami. W dociekaniu podstaw ewolucji zrodziła się **teoria neo-Darwinizmu**, wskazująca na związki procesu ewolucji z genetyką. Teoria ta odrzucała możliwość dziedziczenia cech nabytych (co postulował Lamarck i co dokumentował naiwny eksperyment Weissmana, wykazujący, że mimo obcinania ogonów myszom w kilku pokoleniach nie rodziły się myszy bez ogonów!), podkreślała mutację (genów) jako podstawę istnienia zmienności i wskazywała na eliminację z populacji osobników słabo dopasowanych na rzecz osobników lepiej dostosowanych do środowiska i wykazujących przeto zwiększoną zdolność do reprodukcji. Do sformułowania „syntezy neo-Darwinizmu” przyczyniły się poglądy Theodosiusa Dobzhansky’ego, ale hipoteza miała nadal charakter domniemania, gdyż

natura genu, zjawisko mutacji, mechanizm zmienności i rozwój osobniczy pozostawały nadal w obszarze abstrakcji.

Uderzające są słabe związki genetyki z chemią, która pod koniec XIX wieku poczyniła znaczne postępy w poznawaniu procesów biegnących w żywych organizmach.

Mimo nieokreślonej natury genów, przypisywano im wiele niezwykłych właściwości: geny miały mieć moc działania, być siłą sprawczą (przyczynową), mieć zdolność do samoreprodukcji, co nadawało genom atrybut życia. Geny miały mieć siłę animowania i kontroli rozwoju organizmu, a także miały być białkami lub aperiodycznymi kryształami, upakowanymi w chromosomach, zawierającymi „zapisy kodowe” (Schrödinger, 1998).

Okres poszukiwań i niepewności

Żywiotowy rozwój biochemii na przełomie XVIII wieku i w pierwszej połowie XIX wieku umożliwił również rozwój genetyki – nauki o dziedziczności.

W 1909 roku lekarz londyński Archibald Garrod przedstawił dowód na dziedziczny charakter metabolicznego zaburzenia zwanego alkaptonurią i jednocześnie doszedł do wniosku, że zaburzenie to polegało na niemożności metabolicznego rozerwania pierścienia benzenowego tyrozyny. Dowodził, że najprawdopodobniej u chorych występuje wrodzony brak odpowiedniego enzymu, a ponieważ choroba była dziedziczna, więc musiała być związana z defektem genetycznym.

Inni badacze (Boris Ephrussi, Georg Beadle, Fritz von Wettstein i in.) poczynili obserwacje nad dziedziczeniem koloru oczu u owadów (motyl, muszka owocowa) oraz antocjanin u roślin, wskazując na zależności między powstającym w wyniku reakcji katalizowanych przez enzymy barwnikiem a dziedzicznością. Były to zatem pierwsze przybliżone opisy funkcji genów i ich związków z funkcją enzymów.

Ważnym postępowaniem, wskazującym na ściślejsze relacje między genetyką a biochemią, były badania Georga Beadle i Edwarda Tatum nad pleśnią *Neurospora*. Przy pomocy promieniowania badacze indukowali u zarodników pleśni mutacje, które powodowały utratę zdolności syntezy stosunkowo prostych związków chemicznych (np. witamin, aminokwasów), niezbędnych dla wzrostu. W toku doświadczeń znaleziono takiego mutantu pleśni, który nie był w stanie syntetyzować jednego z aminokwasów, tryptofanu. Krzyżowanie mutantów i hodowle pleśni na odpowiednich prostych pożywkach umożliwiły badaczom precyzyjne określenie szlaku biosyntezy tryptofanu i wykazanie, że każdy etap syntezy był kontrolowany przez oddzielne enzymy i różne, odpowiadające im geny. Tak powstała hipoteza „jeden gen – jeden enzym”. Gen nosił więc atrybuty „planu” dla budowy białka.

Badania metaboliczne wykazały, że geny kontrolują nie tylko ogólne, biologiczne cechy organizmu, lecz także enzymy mające ściśle określone, mierzalne funkcje. Stąd zrodziła się hipoteza, że same geny są niczym innym jak samosyntetyzującymi i samoreplikacyjnymi enzymami.

A ponieważ wiedziano już, że enzymy są białkami hipotezy „jeden gen – jeden enzym” używano również naprzemiennie jako „jeden gen – jedno białko”. Hipoteza ta przez wiele lat była wiodącą koncepcją opisującą naturę genów. Podkreślam ten fakt, aby uświadomić sobie, jakie meandry pokonuje myśl badacza, by na podstawie gromadzonej wiedzy usiłować budować obraz świata.

Okolo 1869 r. lekarz szwajcarski F. Miescher wyizolował z jąder komórek zawartych w ropiejących ranach substancję, której dał nazwę „nukleina”, nazwaną potem kwasem nukleinowym. Podstawową cegiełką kwasu nukleinowego jest nukleotyd. Każdy nukleotyd zbudowany jest z kolei z jednej z czterech rodzajów zasady azotowej o symbolu: A (adenina), T (tymina), G (guanina), C (cytozyna), jednej cząsteczki cukru (deoksyryboza) i reszty fosforanowej (odpowiedzialnej za kwaśny charakter nukleotydu i kwasu nukleinowego). W tekście symbole A, T, G, C używane są wymiennie dla oznaczenia zarówno zasady, jak i nukleotydu.

Duże znaczenie dla rozwoju wiedzy o genie miał fakt rozróżnienia dwóch rodzajów kwasów nukleinowych: DNA zawierającego deoksyrybozę jako składnik cukrowy i RNA, zawierającego rybozę. Stąd pochodzą też akronimy utworzone z ich nazwy w języku angielskim: DNA – *deoxyribo*nucleic acid i RNA – *ribo*nucleic acid. W RNA zamiast tyminy (T) występuje uracyl (U). Metody rozróżniania obu typów kwasów nukleinowych przy pomocy odmiennego sposobu barwienia pozwoliło na ich lokalizację w komórkach zwierzęcych. RNA występował głównie w cytoplazmie (J. Brachet), a DNA w jądrze, w chromosomach (T. Caspersson).

Przez parę dziesiątków lat cząsteczki DNA nie przypisywano specjalnego znaczenia, zwłaszcza, że jeszcze do lat 30. ubiegłego stulecia uważano, że jest to prosta cząsteczka. DNA miał być złożony z czterech nukleotydów, powtarzających się monotonicznie wzdłuż cząsteczki. Choć szwedzki badacz T. Caspersson wykazał, że DNA jest składnikiem chromosomów, sądzono, że stanowi on tylko swego rodzaju ruszt lub podporę dla cząsteczek białkowych, odpowiedzialnych za swoistość genetyczną. Wybór kilkudziesięciu najważniejszych prac opisujących badania nad odkryciem DNA, poznaniem jego struktury i właściwości znajdzie Czytelnik w obszernym tomie opracowanym przez J. H. Taylora (Taylor, 1965).

W 1928 r. F. Griffith opisał niezwykle spostrzeżenie. Niewirulentne (nie dające efektu patologicznego) dwoinki zapalenia płuc u myszy zmieszane z wirulentnymi, ale zabitymi dwoinkami, ulegały przekształceniu (transformacji) i stawały się wirulentnymi. Co więcej, cecha wirulencji pojawiała się w bakteriach niewirulentnych nawet wtedy, gdy do ich kolonii dodawano jedynie ekstrakt z zabitych, wirulentnych bakterii.

W 1944 r. Oswald T. Avery, badacz z prestiżowego Instytutu Rockefellera oraz Colin MacLeod i Maclyn McCarthy, idąc śladem odkrycia Griffitha, dostarczyli dowodu, że czynnikiem transformującym, przenoszącym cechę wirulencji, był DNA. Odkrycie Avery’ego stało się wielkim wyzwaniem dla powszechnie przyjmowanej hipotezy o białkowej naturze genu. Tymczasem doświadczenia

Avery'ego wskazywały jednoznacznie na DNA jako cząsteczkę przenoszącą cechę fenotypową, a więc mającą atrybut genu.

Nieformalna „grupa fagowa” amerykańskich badaczy, działająca w latach 1940-1960, obrała bakteriofagi jako model badawczy zjawisk życia. Trzon grupy stanowili m.in. Max Delbrück, Salvadore Luria, Nikolaj Timofeef-Ressovsky, Alfred Hershey. Delbrück sądził, że bakteriofagi są elementarnymi cząsteczkami biologicznymi, które, przez swą zdolność do infekowania bakterii i gwałtownego namnażania się, stanowią idealny model do badań nad fundamentalną cechą życia – samoreplikacją. Ponadto Joshua Lederberg (1951) wykazał, że bakteriofag może przenosić cechy dziedziczne z jednego szczepu bakterii na inny (zjawisko transdukcji).

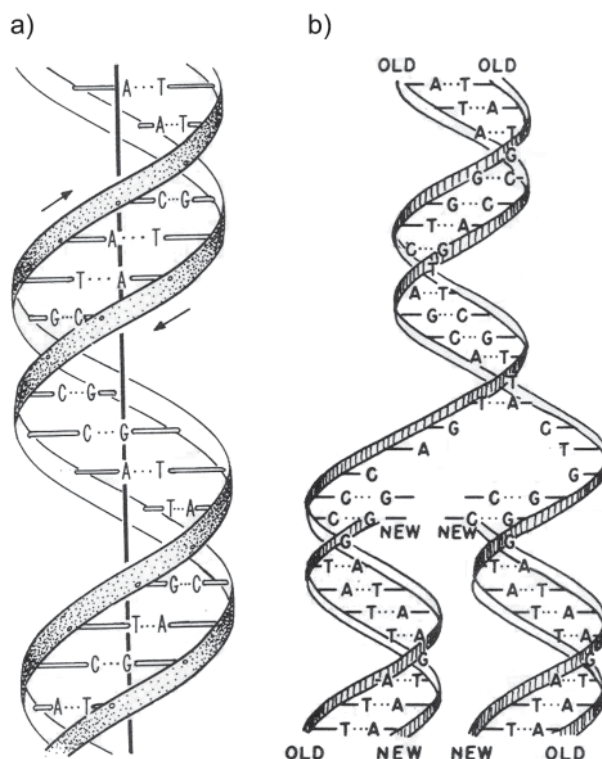
Proces replikacji bakteriofaga ma miejsce w zakażonej bakterii. Ku powszechnemu zdumieniu, Alfred Hershey i Martha Chase (1952) udowodnili jednak, że w procesie zakażenia fag nie wchodzi do komórki bakteryjnej w całości, lecz wstrzykuje jedynie do wnętrza bakterii swój DNA. Białkowa otoczka bakteriofaga pozostaje na powierzchni błony komórkowej bakterii. Konkluzja z doświadczenia Hershey-Chase była prosta: DNA fagowy zawiera całą informację potrzebną dla reprodukcji i uformowania dojrzałego bakteriofaga.

Odkrycie zjawiska transformacji i transdukcji oraz poznanie mechanizmu infekcji bakterii bakteriofagiem były bodźcem dla intensywnych badań nad budową DNA i jego rolę w przenoszeniu cech dziedzicznych. Szybko zbliżał się rozkwit nowej wiedzy o genie jako fizycznej cząsteczce.

Gen materialny – DNA

W omawianym okresie genetycy sądzili, że podstawową i wyjątkową właściwością genu jest zdolność do samoreplikacji. Ta właściwość genu jest właściwością oddzielną od funkcji genu, zatem „samoreplikacja” może być badana niezależnie od funkcji. W konsekwencji badania nad zjawiskiem samoreplikacji DNA nie zajmowały genetyków, lecz głównie fizyków i chemików i sprowadzały się głównie do konstruowania modeli.

Odkrycie struktury DNA (Watson i Crick, 1953), dokonane w Cavendish Laboratory, jest zapewne najbardziej powszechnie znanym odkryciem w biologii. Cytuję tylko pierwszą publikację Watsona i Cricka z *Nature* z dnia 17 kwietnia 1953 r. zaczynającą się od słów: „Chcemy sugerować strukturę (...) kwasu deoksyrybonukleinowego D.N.A.”. Następne serie prac tych autorów znajdzie Czytelnik w zbiorze publikacji (Taylor, 1965). Zarówno Crick, jak i Watson, nie prowadzili badań laboratoryjnych. Ich zasadniczym źródłem wiedzy były nieformalne dyskusje z Maurice Wilkins'em, rzadziej z Rozalindą Franklin – współpracowniczką Wilkina, która również interesowała się badaniem struktury DNA metodą dyfrakcji promieni rentgenowskich. Praca obu badaczy sprowadzała się do konstruowania fizycznego modelu, który uwzględniał fizyczne proporcje, odległości, kąty, wiązania i siły oddziaływania między poszczegól-

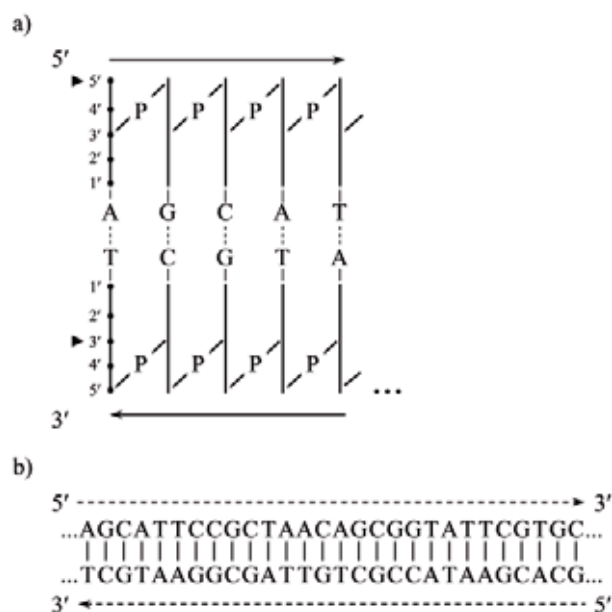


Ryc. 1. Schemat ilustrujący strukturę i replikację DNA. a) struktura DNA – dwa pasma utworzone przez ciąg nukleotydów (nukleotyd to jedna z czterech rodzajów zasad A, C, G, T, reszta cukrowa i reszta fosforanowa), zwinięte wzdłuż osi, tworzą helisę. Pasma utrzymane są między sobą przez wiązania chemiczne między zasadami A-T i G-C. Struktura przypomina „spiralne” schody, w których stopnie stanowią płaszczyzny par zasad, prostopadłe do osi. Odstęp między kolejnymi stopniami wynosi 3,4 Ångströma. b) schemat replikacji DNA. Na każdym „starym” paśmie DNA odtwarza się nowe pasmo. Po ukończeniu replikacji dwie cząsteczki potomne („nowe”) mają sekwencję nukleotydów identyczną, jak w starej cząsteczce. (Zapóżyczono z: J. D. Watson: *The double helix*, Weidenfel and Nicolson, London, 1968)

nymi atomami składników DNA. Nie bez znaczenia dla konstrukcji modelu struktury DNA były wyniki badań Erwina Chargaffa (Chargaff, 1950) nad DNA i zauważona przez niego reguła, że zasady DNA istnieją w ekwiwalentnych ilościach: stosunek A:T = 1 oraz G:C = 1. Reguła ta nasuwała myśl, że w cząsteczce DNA może zachodzić parowanie zasad, czyli chemiczne wiązanie zasad w pary A z T i G z C. Crick uznawał wkład Chargaffa w budowaniu modelu DNA, czemu wielokrotnie dawał wyraz (Crick, 1988).*

W modelu struktury cząsteczki DNA Watsona-Cricka (Ryc. 1) oba pasma DNA biegną anty-równolegle (mają odmienną polarność) i są skrócone wzdłuż swojej osi (helisa). Oba pasma są utrzymane w rejestrze dzięki wiązaniom chemicznym (wodorowym) między parami zasad A-T i G-C, dopasowującym je wzajemnie do siebie. Dowolny ciąg par nukleotydów w cząsteczce DNA można zatem zapisać jako ciąg sekwencji jednego (np. górnego) pasma: AGCATTCCG...

* NB Chargaff odczuwał przez resztę swego życia pewne niedowartościowanie swego odkrycia przez przyszłych noblistów i często powtarzał nieco zgryźliwą sentencję (którą parokrotnie bezpośrednio słyszałem na jego wykładach), że „biolog molekularny to jest taki człowiek, który uprawia biochemię bez posiadania licencji chemika”.



Ryc. 2. Schemat zapisu fragmentu dowolnej sekwencji nukleotydów w dwupasmowej cząsteczce DNA. a) Oba pasma są trzymane w komplementarnym „rejestrze” dzięki wiązaniom zasad A-T i G-C. Każda z zasad związana jest z cukrem (deoksyrybozą) o pięciu atomach węgla (1' do 5'), a ten tworzy wiązanie poprzez atom fosforu z następną cząsteczką cukru, między atomami węgla 3' i 5'. W ten sposób powstaje ciągłość pasm DNA wzdłuż osi cząsteczki. Zasada azotowa związana z deoksyrybozą i resztą fosforanową stanowi nukleotyd. Strzałki pokazują polarność cząsteczki DNA od 5' do 3'. Oznakowanie 5' lub 3' bierze się od wolnego wiązania (na końcu cząsteczki DNA) przy 5' lub 3' pozycji atomu węgla deoksyrybozy. b) Uproszczony zapis sekwencji. Literami (A, T, G, C) oznaczone są jedynie zasady azotowe tworzące „kod”. Reszty cukrowe i fosforanowe powtarzają się, zatem nie ma potrzeby uwzględniać ich w zapisie

Każda para nukleotydów A-T lub G-C zajmuje w osi cząsteczki pewną przestrzeń (ok. 3,4 Ångströma) i przeto służy jako miara długości (wielkości) cząsteczki DNA lub jego fragmentu. Mówimy więc, że np. gen ma 1500 par zasad lub 1,5 kilo par zasad (para zasad w języku angielskim – *base pair*; stąd zwykle pochodzą skrót *bp*), tu: 1500 bp lub 1,5 kb).

Dowolną sekwencję nukleotydów możemy zapisać jako ciąg symboli literowych (Ryc. 2b), bez zaznaczania cząsteczek cukrowych i reszt fosforanowych, które tworzą jedynie zewnętrzne rusztowanie, utrzymujące ciągłość cząsteczki DNA (Ryc. 2a).

Oś cząsteczki („wnętrze” cząsteczki) stanowią pary A-T i G-C, ułożone jedna nad drugą i powiązane wiązaniami wodorowymi, natomiast reszty fosforanowe i deoksyrybozy są „na zewnątrz”. Reszty fosforanowe nadają charakter kwaśny cząsteczce DNA, stąd w nazwie znajduje się słowo „kwas” (kwas deoksyrybonukleinowy), co niespecjalistom, np. w znanej anegdocie z Łysenką, nasuwa popołite skojarzenia**.

** Prof. Ilja B. Zbarsky opowiadał mi, że po referacie Prof. Bielozierskiego o budowie i znaczeniu kwasu deoksyrybonukleinowego, sceptyczny Łysenko poprosił o próbkę roztworu DNA. Ku zdumieniu asystenta Bielozierskiego, który następnego dnia dostarczył próbkę, Łysenko wylał na dłoń parę kropli roztworu, posmakował je językiem i oznajmił: „Eto na wierno nie kislota” („To na pewno nie kwas”).

Żadne z dotychczasowych odkryć w biologii nie miało tak wielkiego wpływu na rozwój nauki, jak odkrycie struktury cząsteczki DNA, dokonane przez Watsona i Cricka.

Model Watsona-Cricka sugerował koncepcję semikonserwatywnej replikacji DNA czyli odtwarzania (syntezy) identycznych względem każdego z komplementarnych pasm kopii (czyli takiego samego ciągu sekwencji zasad), w wyniku czego powstają dwie cząsteczki potomne, identyczne z cząsteczką macierzystą (Ryc. 1b). Semikonserwatywny model replikacji DNA został potwierdzony w r. 1957 przez Matthew Meselsona i Franklina Stahla (zob. Taylor, 1965). Watson i Crick sądzili, że model cząsteczki DNA zapewnia jedną z podstawowych cech postulowanych dla czynnika dziedziczności: możliwość powielania cząsteczki z zachowaniem budowy i układu sekwencji zasad (nukleotydów) cząsteczki macierzystej.

Struktura cząsteczki DNA, złożona z dwóch komplementarnych w stosunku do siebie pasm i semikonserwatywna replikacja sugestywnie, aczkolwiek myląco, ujmowana jest często jako inherentna zdolność DNA „do samoreplikacji”. W istocie utworzenie dwóch cząsteczek potomnych (replikacja) z cząsteczki DNA, od początkowej fazy oddzielania się pasm do ukończenia procesu prowadzona jest przez złożone systemy enzymatyczne.

W późniejszych latach odkryto, że pasmo DNA służy również jako matryca do syntezy komplementarnej nici RNA. Cząsteczka DNA jest uważana za strukturę stabilną mimo, że DNA ma (wprawdzie ograniczone) zdolności autokatalityczne (DNAzy) (por. Adamala i Pikuła, 2004). Również niektóre szczególne sekwencje DNA mają zdolność do przemieszczania się (por. niżej), a w komórkach somatycznych zachodzi zjawisko rekombinacji (np. formowanie i rearanżacja genów w procesie odpowiedzi immunologicznej, rearanżacja sekwencji kodujących receptor powierzchniowy limfocytów typu T).

Odkrycie Watsona i Cricka zostało jednoznacznie przyjęte przez biologów i genetyków jako odkrycie materialnej struktury genu. W euforii towarzyszącej odkryciu badacze ci wypowiedzieli wiele deklaracji cytowanych do dzisiejszego dnia, nadających genowi (DNA) moc sprawczą i plasujących gen (DNA) na centralnej pozycji w zjawisku życia: „Odkryliśmy istotę życia”, „Geny są istotą życia”.

Nawet po upływie dziesiątków lat przekonanie, że geny/DNA są jakąś samoistną siłą sprawczą procesów rozwojowych i samego życia jest przewodnią myślą wypowiedzianą przez wielu badaczy. Współczesne definicje genu opierają się na uznaniu go jako realnej, fizycznej jednostki genomowej i formowane są w oparciu o sekwencje DNA. „Genes are made of DNA” stwierdza czołowe dzieło „The Molecular Biology of the Cell” (Alberts i in., 1994).

Kod genetyczny

Identyfikacja DNA jako materialnego nośnika cech dziedzicznych i odkrycie struktury DNA otwarły następny

problem: na czym polega „kod genetyczny” i rola DNA (identyfikowanego już z genem) w syntezie białka, które jest ostatecznym wyrazem funkcji genu. W latach 50. było już ugruntowane przekonanie, że synteza białka w cytoplazmie polega na uformowaniu łańcucha aminokwasów (łańcuch polipeptydowy), który następnie spontanicznie przyjmuje konformację przestrzenną (fałduje się) i staje się cząsteczką funkcjonalną. Podstawowych aminokwasów jest 20. Ustalony też był fakt, że DNA znajduje się w jądrze komórkowym, a synteza białka ma miejsce w cytoplazmie. Pierwsze pytanie było: jak przekłada się instrukcja zawarta w DNA na ciąg aminokwasów w białku?

Już w pierwszej publikacji w *Nature* (kwiecień, 1953), poświęconej wyłącznie strukturze cząsteczki DNA, Crick i Watson przedstawili krótki komentarz, ekscytując świat badaczy: „Nie uszedł naszej uwadze fakt, że swoiste parowanie [zasad], które postulujemy, natychmiast nasuwa sugestię co do możliwego kopiowania materiału genetycznego”. To wyważone sformułowanie było podyktowane, jak przyznaje Crick, zapewnieniem sobie pierwszeństwa w kwestii dalszych prac nad kodem genetycznym (Crick, 1988).

George Gamow sugerował, że istnieje bezpośrednia relacja między sekwencjami nukleotydów w DNA a sekwencją aminokwasów białka. Gamow rozważał nawet strukturę „kodu”, która miałaby obejmować trzy zasady (nukleotydy), czyli triplet tworzący kodon, który miał być kodem nakładającym się. Pogląd Gamowa, choć krytykowany przez Cricka, był jednak pewną inspiracją dla niego i nasunął myśl o bezpośrednim związku między nukleotydami DNA i aminokwasami oraz kolinearnością między DNA i łańcuchem białkowym. Jednak molekularny mechanizm rozpoznający taki kod był nadal trudny do wyobrażenia.

Na zasadzie podobnej odległości między nukleotydami w paśmie DNA i wiązaniami peptydowymi aminokwasów w białku Alexander Dounce w 1953 r. sugerował, że DNA może służyć jako matryca dla syntezy RNA, a ten ostatni może być wtórną matrycą dla syntezy białka. Podobną sugestię wysunęli już wcześniej francuscy badacze André Boivin i Roger Vendrely. Oba te fakty są mało znane, cytując je za Morange (Morange, 1998).

Tymczasem badania nad syntezą białka w układach bezkomórkowych wskazywały, że dla syntezy niezbędne są m.in. mikrosomy, ekstrakty komórkowe zawierające enzymy i RNA oraz ATP jako źródło energii. Paul Zamecnik i in. wykazali, że w takim systemie znakowany aminokwas był „ładowany” na małą cząsteczkę RNA, a następnie wchodził w łańcuch polipeptydowy. Już w 1954 r. Crick przewidział, że w systemie syntezy białka musi istnieć drobna cząsteczką RNA („adaptor”), która wchodzi w interakcję z tripletem (kodonem) matrycy (DNA?, RNA?), a drugi region adaptora wiąże aminokwas i dostarcza go do powstającego polipeptydu. Adaptor miałby mieć rolę „reduktora”, tłumaczącego trzy sąsiadujące nukleotydy pasma DNA na jeden aminokwas w łańcuchu białka. Musi istnieć, dowodził Crick, przynajmniej 20 takich rodzajów małych cząsteczek RNA – jeden dla każdego aminokwasu. Istotnie, kilka lat póź-

niej odkryto klasę cząsteczek transportującego (transportującego aminokwas) RNA (tRNA). Okazało się jednak, że rozpoznają one triplety nie DNA, lecz wówczas jeszcze nie znanego informacyjnego RNA – cząsteczki pośredniczącej między matrycą DNA a ośrodkami syntezy białka.

W 1957 r. Crick i Leslie Orgel zaproponowali koncepcję „ramki odczytu” (reading frame) jako teoretycznego rozwiązania kodu. Trzy sąsiadujące w pasmie DNA nukleotydy czyli triplet kodują dany aminokwas, ale gdy ramka odczytu przesunie się o jeden nukleotyd, to wszystkie następujące triplety tracą swój sens kodujący. Crick na posiedzeniu British Society of Experimental Biology przedstawił pogląd, że końcowa konformacja przestrzenna cząsteczki białkowej jest zdeterminowana sekwencją aminokwasów i zachodzi spontanicznie, a także sugerował, że swoistość kwasów nukleinowych polega na sekwencji zasad (nukleotydów) i ogłosił wstępną wersję swego słynnego „dogmatu głównego” (central dogma), wskazując jedynie, że część „RNA musi być pod kontrolą DNA”. Ostateczna wersja głównego dogmatu, że przepływ informacji biegnie od DNA przez RNA do białka nie mogła być jeszcze sprecyzowana, bo w tym czasie nie została jeszcze odkryta klasa informacyjnego RNA (messenger RNA – mRNA).

„Jest teraz – czytamy w pracy Cricka i in. (Crick i in., 1961) – dużo pośrednich dowodów, które sugerują, że sekwencja aminokwasów wzdłuż łańcucha polipeptydu białka jest określona przez sekwencję zasad wzdłuż pewnej szczególnej części kwasu nukleinowego w materiale genetycznym”.

Informacyjny RNA (mRNA) został odkryty w 1960 r. przez Artura Pardee, François Jacoba i Jaques Monod, którzy na podstawie badań własnych nad indukowalnym enzymem beta-galaktozydazą i jej genem u bakterii postulowali istnienie krótko żyjącego RNA. Taka labilna cząsteczką RNA miałaby pośredniczyć w przenoszeniu informacji między genami (DNA) a mikrosomami jako ośrodkami syntezy białka (mikrosomy określono później jako rybosomy). Do zbliżonych wniosków prowadziły też badania Heinrich Matthaei i Marshalla Nirenberga w National Institutes of Health (NIH), jak też wyniki badań François Grosa.

Tuż przed końcem 1961 r., w *Nature* Crick i in. (zob. Taylor, 1965) podali ogólne cechy uniwersalnego kodu:

- a) grupa trzech zasad [nukleotydów] koduje jeden aminokwas
- b) kod nie ma charakteru kodu „nakładającego”
- c) sekwencja zasad jest prawdopodobnie odczytywana od określonego punktu startowego. Nie ma specjalnych „przecinków” wskazujących jak wybierać prawidłowy triplet
- d) kod jest prawdopodobnie zdegenerowany; to znaczy, że jeden szczególnie aminokwas może być kodowany przez kilka tripletów.

W tym czasie ukształtowały się również elementy centralnego dogmatu Cricka, oparte na następujących stwierdzeniach: 1) cała genetyczna (dziedziczna) informacja jest przechowywana w kwasach nukleinowych;

2) struktura podwójnej helisy DNA wyjaśnia, jak informacja jest przechowywana i kopiowana; 3) informacja jest przechowywana jako kod digitalny (literowy); 4) informacja przepływa nieodwracalnie od kwasów nukleinowych do białek. Ostatnie stwierdzenie zostało spopularyzowane jako powiedzenie, które wyszło z pracowni Cricka i Watsona: „DNA robi RNA, RNA robi białko, białko robi nas”.

Mimo, że w omawianym okresie zebrano dużo dowodów, że „kod genetyczny” był zapisany w DNA, a cząsteczką pośredniczącą w syntezie białka był „messenger” RNA (mRNA), molekularny mechanizm przeniesienia „kodu” z DNA na białko i natura kodu nie były do końca jasne. Wyniki dziesiątków prac wskazywały na powiązanie DNA z syntezą białka i udział w tym procesie RNA. Oprócz mRNA i tRNA zidentyfikowano także rybosomalny RNA (rRNA), stwierdzono, że wszystkie klasy RNA powstają w jądrze oraz, że są one komplementarne do sekwencji DNA. Dowodem, że synteza RNA zależy od DNA były m.in. obserwacje nad hamowaniem syntezy RNA w różnych komórkach, po zablokowaniu matrycy DNA aktynomycyną D. Opracowanie modeli syntezy białka w systemach bezkomórkowych pozwoliło na rozwój koncepcji dotyczących funkcji w tym procesie różnych klas RNA.

Stwierdzono, że w bakteriiach zainfekowanych fagiem ma miejsce szybka synteza RNA i że ten RNA bierze następnie udział w syntezie fagowych białek oraz tworzy hybrydową cząsteczkę – fagowy DNA i „nowy” RNA. Wykrycie cząsteczek hybrydowych DNA:RNA, a także oddzielenie dwóch pasm DNA (denaturacja) i ponowne odtworzenie struktury dwupasmowej (renaturacja) wskazywały na ścisłe powiązania strukturalne między tymi cząsteczkami. Jak wspominałem, przegląd tych prac znajdzie Czytelnik w zbiorze pod red. J. H. Taylora (zob. Taylor, 1965).

W 1961 r., na 5. Światowym Kongresie Biochemicznym w Moskwie, pracujący w National Institutes of Health (NIH), niemiecki biolog Heinrich Matthaei i Marshall Nirenberg przedstawili doniesienie o „złamaniu kodu genetycznego” (Nirenberg i Matthaei, 1961). System syntezy białka, opracowany przez tych badaczy, zawierał mieszaninę podstawowych aminokwasów (wśród nich, w kolejnych doświadczeniach, jeden aminokwas był znakowany izotopem radioaktywnym), ATP, drobnocząsteczkowy RNA, ekstrakt bakteryjny i syntetyczną matrycę, ale typu RNA – polimer złożony z samych uracyli, U-U-U-U-U- ... czyli poliU. Inkubacja takiej mieszaniny, a następnie analiza zsyntetyzowanego białka wykazała, że matryca poliU kodowała monotonną cząsteczkę polipeptydu („białka”), złożoną wyłącznie z jednego rodzaju aminokwasu – fenyloalaniny (akronim: Phe). Łańcuch polipeptydowy miał zatem budowę: Phe-Phe-Phe-...

Wkrótce uzyskano syntetyczne polinukleotydy (syntetyczne RNA) o znanej i zaprogramowanej sekwencji nukleotydów (G. Khorana, M. Grunberg-Manago i S. Ochoa), które użyte w systemie Matthaei i Nirenberga pozwoliły w ciągu kilku lat poznać wszystkie kodujące triplety, czyli „złamać” kod genetyczny. Dla przykładu:

triplet UUU koduje fenyloalaninę, GCU – alaninę, AGC – serynę, CUG – leucynę, itd. Kod genetyczny, opisujący triplety (kodony) RNA i odpowiadające im aminokwasy oraz kodony sygnałne „stop”, znajdzie Czytelnik w każdym podręczniku biologii.

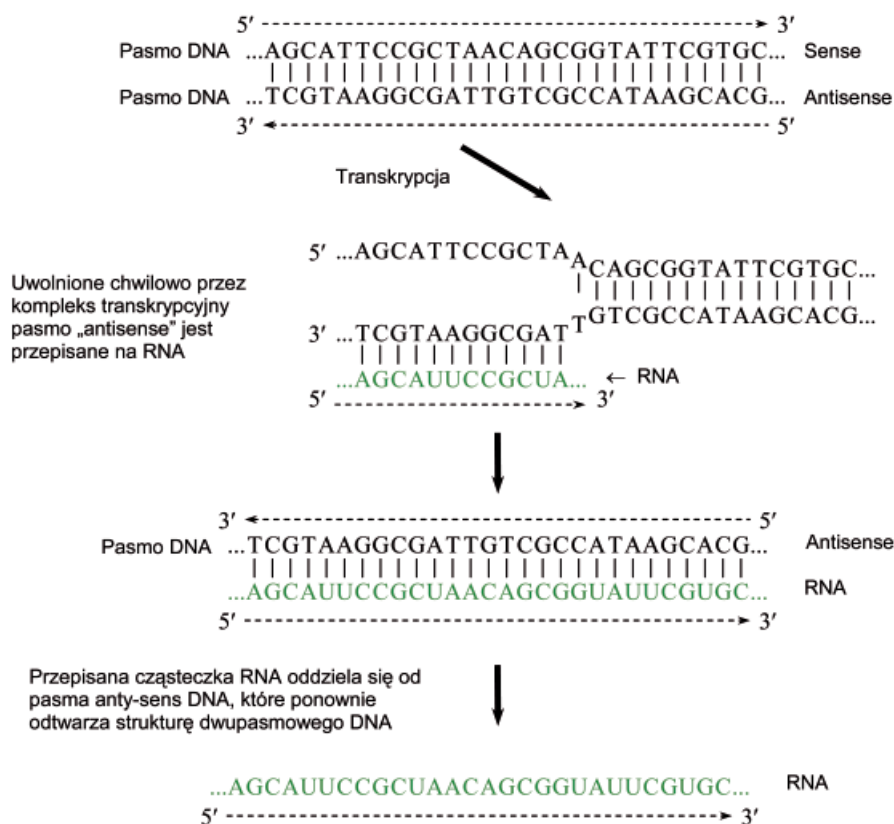
Gen klasyczny – elementarz

Ówczesna, oparta na hipotezie Cricka, definicja genu głosiła, że gen jest „odcinkiem DNA” kodującym białko (polipeptyd).

Dwa główne procesy są niezbędne dla odczytania informacji zakodowanej w genie (DNA) i wyrażenia jej w postaci syntezy swoistego białka: transkrypcja i translacja.

Proces transkrypcji czyli przepisania kodujących sekwencji nukleotydowych DNA na liniową, prekursorową cząsteczkę informacyjnego RNA (mRNA, messenger RNA) zaczyna się w określonym punkcie sekwencji genu zwanej promotorem. Promotor zwykle poprzedza początek genu (jest „na lewo” od początku, albo inaczej: mieści się „w górę” od początku genu). Na sekwencjach promotora usadawia się kompleks transkrypcyjny, przesuwany się od początku do końca genu i rozrywający chwilowo wiązania w helisie DNA. W wyniku transkrypcji powstaje pojedyncza nić mRNA, która po ukończeniu transkrypcji na końcu genu oddziela się od pasma DNA. Oba pasma DNA ponownie tworzą dwupasmową helisę. Każdy gen może dać początek tysiącom cząsteczek komplementarnego mRNA. „Informacja genetyczna” w postaci sekwencji tripletów nukleotydów (kodonów) zawarta jest na tzw. paśmie sensowym („sense”) DNA. Informacyjny RNA (mRNA) kopiowany jest z pasma DNA, komplementarnego do pasma „sense” czyli pasma antysensowego („antisense”). W wyniku transkrypcji pasma anty-sensowego, w cząsteczce mRNA odtworzona zostaje sekwencja pasma kodującego czyli pasma sensowego (por. Ryc. 3), ze znanym nam już wyjątkiem: w miejscu, gdzie w paśmie kodującym DNA znajduje się T (tymina), w mRNA pojawia się U (uracyl). Transkrypcja i pojawianie się cząsteczek mRNA w komórce jest wyrazem „aktywacji” genu lub jego „ekspresji”. Przyjęto, że sekwencje kodujące białko będą przedstawiane jako zapis na mRNA, począwszy od lewej strony (koniec 5’ cząsteczki mRNA), od kodonu AUG. Kodon AUG jest pierwszym tripletem, od którego zaczyna się synteza białka.

Cząsteczka mRNA, oprócz sekwencji tripletów (kodonów) kodujących poszczególne aminowasy w polipeptydzie, niesie także u obu końców cząsteczki dodatkowe sekwencje (nie ulegające translacji) oraz inne specyficzne sekwencje, jak zlokalizowany na początku cząsteczki (5’) sygnał (kodon) AUG dla startu odczytywania (translacji) i sygnały „stop” (kodony UAA, UAG i UGA) dla zakończenia odczytu (u końca 3’) i inne. Sekwencje te służą jako swoiste miejsca dla lokalizacji białek i ich kompleksów, niezbędnych dla procesu odczytu, a także jako sekwencje biorące udział w procesach kontrolnych, związanych np. z trwałością cząsteczki mRNA. Odczyt,



Ryc. 3. Schemat transkrypcji. Przed rozpoczęciem transkrypcji pasma DNA zostają chwilowo odsunięte od siebie przez kompleks transkrypcyjny. „Przepisanie” sekwencji z DNA na RNA zachodzi na paśmie antysensowym, komplementarnym do pasma sensowego. Zauważmy, że sekwencja w pre-mRNA oraz polarność są identyczne, jak w paśmie sensowym DNA (czyli „genie”), z tym, że tyminę (T) zastąpił uracyl (U). Zatem sekwencje genowe możemy odczytywać na paśmie sensowym DNA, jak i na paśmie RNA, począwszy od 5’

czyli translacja informacji zawartej w mRNA na łańcuch białkowy, czyli synteza białka, zachodzi na rybosomach w cytoplazmie komórki z dostępnych tam wolnych aminokwasów i przy użyciu cząsteczek tRNA – „adaptorów”, postulowanych przez Crick’a i bardzo złożonych systemów „pomocniczych”.

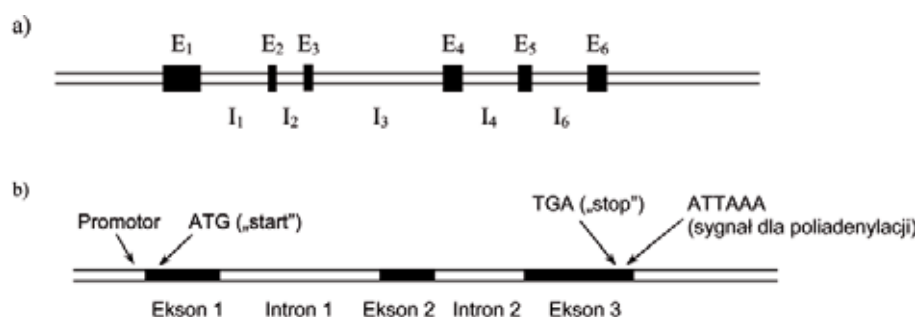
Każda cząsteczka mRNA może służyć do syntezy tysięcy cząsteczek białka, które koduje. Jednak żywotność cząsteczki mRNA jest regulowana. Powyższy opis ilustruje „ekspresję” genu i dotyczy w zasadzie genu bakteryjnego, który występuje jako ciągła sekwencja DNA.

Sądzone, że w komórkach jądrazstych (eukariotycznych) struktura genu jest podobna do struktury w komórkach bakteryjnych. Ale w końcu lat 70. zaczęły pojawiać się zdumiewające doniesienia, że u niektórych wirusów i w komórkach eukariotycznych organizmów wyższych geny kodujące białko z reguły mają sekwencje kodujące nie w formie ciągłej, lecz są one poprzerywane sekwencjami niekodującymi. Walter Gilbert zaproponował, aby fragmenty kodujące (znajdujące się następnie w „dojrzałym” mRNA) nazywać **eksonami (E)**, zaś przedzielające je sekwencje niekodujące **intronami (I)**. Geny o strukturze *intron-ekson-intron-ekson...* Gilbert nazwał genami **mozaikowymi**. Nazywa się je także genami „podzielonymi” – *split genes* (Pierre Chambon i wsp.). Graficznie strukturę takich genów przedstawiamy jako podwójną linię, a kodujące sekwencje genowe (eksony,

E) jako pionowe krótkie linie, o różnej grubości, co ma ilustrować ich względną długość (Ryc. 4a) i względne proporcje między eksonami (E) i intronami (I). Innym, wygodnym przedstawieniem graficznym jest pokazanie eksonów (E) jako odcinków wypełnionych na podwójnej linii DNA (Ryc. 4b).

Zatem w komórkach eukariotycznych (jądrzastych) procesy związane z ekspresją genu wyglądają inaczej. Skopiowana z genu wielka cząsteczka zwana pre-mRNA zawiera zarówno kopie eksonów, jak i intronów i podlega procesowi „składania” (*splicing*). Z cząsteczki pre-mRNA zostają usunięte introny, a eksony zostają połączone ze sobą w ciągłą sekwencję, stanowiącą teraz ostateczną, dojrzałą (znacznie mniejszą niż pre-mRNA) cząsteczkę informacyjnego RNA (mRNA). Cząsteczka mRNA jest teraz gotowa podjąć funkcję kodowania białka w cytoplazmie komórki.

Liczba eksonów/intronów jest różna (od kilku do kilkudziesięciu) w różnych genach i u różnych gatunków. Największa znana mi z literatury liczba eksonów występuje u muszki owocowej *D. melanogaster* w genie o symbolu *Dsam* i wynosi 115. Eksony te są rozrzucone na wielkim sięgającym milionów par zasad fragmencie DNA. Ludzki homolog tego genu posiada ponad 100 eksonów, podobnie rozsianych na dużym obszarze w znacznych odległościach od siebie.



Ryc. 4. Powszechnie używane schematy ilustrujące budowę genu mozaikowego: a) gen o 6 eksonach; pokazano względne proporcje długości eksonów (E) i intronów (I); b) gen o trzech eksonach; zaznaczone są podstawowe sekwencje sygnałowe. Opis w tekście. Sygnały „start” i „stop” pokazują miejsca początku i końca przepisywania „informacji” na sekwencję aminokwasów białka

Introny we wszystkich genach są z reguły znacznie dłuższe od eksonów i ich łączne sekwencje zajmują często ponad 95% sekwencji pierwotnego transkryptu (por. Ryc. 4a i 5a). Jeszcze kilkanaście lat temu introny były uważane za „śmieciowego” DNA (junk DNA, por. niżej). Tymczasem okazało się, że w samych intronach mogą się znajdować sekwencje kodujące, należące do innego genu, a także sekwencje regulatorowe, np. tzw. sekwencje wzmacniające transkrypcję (*enhancery*), często leżące w znacznej odległości „w dół” (czyli „na prawo”) od początku genu. Sądzono, że po wycięciu z pre-mRNA introny jako transkrypt „śmieciowy” ulegają kompletnej degradacji jako sekwencje niepotrzebne. Obecnie wiemy, że introny jako główny składnik w pierwotnym transkrypcie (pre-mRNA) i, wraz z transkryptami intergenowych sekwencji unikatowych niekodujących białka, stanowią ważny materiał, użyteczny w regulacji wielu procesów biologicznych. Ta porcja sekwencji, która znajduje się w intronach, najlepiej koreluje ze złożonością rozwoju ewolucyjnego (Mattick i Gagen, 2001, także por. niżej).

Następująca definicja dobrze opisuje gen „klasyczny”, taki, jaki znaleźliśmy w końcu ubiegłego stulecia: „Gen jest segmentem DNA wciągniętym w produkowanie łańcucha polipeptydowego; gen obejmuje regiony poprzedzające i następujące po regionie kodującym, jak również sekwencje wtrącone (introny), leżące między indywidualnymi segmentami kodującymi (eksonami)” (zob. Lewin, 1994). Gen „klasyczny” miał swoją ściśle zdefiniowaną liniową strukturę, swój początek i koniec. Geny miały być ułożone jeden za drugim jako oddzielne jednostki funkcjonalne, jak paciorki różańca. Obecnie wiemy, że geny mogą nakładać się („overlapping”) i wówczas z pasma „sense” odczyt zachodzi z przesunięciem ramki odczytu, w wyniku czego na jednym paśmie „sense” mogą być zakodowane dwa polipeptydy. Także pasmo „antisense”,

do niedawna uważane za nie kodujące, może służyć do kodowania i jest wykorzystane do transkrypcji.

Prof. dr hab. n. med. Mieczysław Chorąży

Centrum Badań Translacyjnych

i Biologii Molekularnej Nowotworów

Centrum Onkologii – Instytut im. Marii Skłodowskiej-Curie

Oddział w Gliwicach

e-mail: chorazy@io.gliwice.pl

Piśmiennictwo

- Adamala KI, Pikula S. Hipotetyczna rola autokatalitycznych właściwości kwasów nukleinowych w procesie biogenezy. *Kosmos* 2004; 53: 123-31.
- Alberts A, Bray D, Lewis J i wsp. *Molecular biology of the cell*. New York & London: Garland Publishing Inc., 1994.
- Chargaff E. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* 1950; VI/6, 201-209.
- Crick F HC, Barnett L, Brenner S i wsp. General nature of the genetic code for proteins. *Nature* 1961; 192: 1227-32.
- Crick F. *What Mad Pursuit – A Personal View of Scientific Discovery*. Wyd. Basis Books, 1988.
- Lewin B. *Genes*. Oxford New York Tokyo: Oxford University Press, 1994.
- Mattick JS, Gagen MJ. The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. *Mol Biol Evol* 2001; 18: 1611-30.
- Morange M. *A History of molecular biology*. Cambridge and London: Harvard University Press, 1998.
- Nirenberg MW, Matthaei JH. The dependence if cell-free protein synthesis in *E. Coli* upon naturally occurring or synthetic poliribonucleotides. *Proc Natl Acad Sci* 1961; 47, 1588-1602.
- Schrödinger E. *Czym jest życie?* (tłum. z ang.). Warszawa: Pruszyński i Sk-a; 1998.
- Taylor JH (red.). *Selected papers on molecular genetics*. New York and London: Academic Press; 1965 (replika 55 wybranych oryginalnych prac dotyczących, genetyki, poszukiwań mat. genetycznego, struktury i replikacji DNA, funkcji materiału genetycznego itp.).
- Watson JD, Crick FHC. Molecular structure of nucleic acids: a structure of deoxyribose nucleic acid. *Nature* 1953; 171: 737-38.
- Watson J. D. *The double helix*. London: Weidenfeld and Nicholson; 1968.

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