



## WSTĘP

W 1869 roku Friedrich Miescher odkrył DNA, który nazwał „nukleinę”. W 1918 roku Phoebus Levene zidentyfikował cztery zasady: adeninę (A), cytozynę (C), guaninę (G) i tyminę (T), elementy budulcowe DNA (Levene, 1919). Trzydzieści lat później Erwin Chargaff odkrył, że udział zasad w DNA różni się między gatunkami, ale ustalił, że w obrębie gatunku zasady w DNA są zawsze obecne w stałych proporcjach: w takiej samej liczbie A jak T oraz takiej samej C i G (Chargaff, 1950). Wielka kariera DNA rozpoczęła się jednak w 1944 roku wraz z eksperymentem Avery’ego, MacLeoda i McCarty’ego (Avery i in., 1944), udowadniającym, że to nie białka czy RNA, jak wcześniej podejrzewano, ale DNA jest substancją odpowiedzialną za transformację bakterii. Proces transformacji bakterii został po raz pierwszy opisany w 1928 roku przez Griffitha, który odkrył, że ekstrakt ze zjadliwego szczepu *Pneumococcus pneumoniae* może indukować konwersję szczepu nie-zjadliwego w zjadliwy (Griffith, 1928). Kolejnym kamieniem milowym w badaniach DNA był model struktury i replikacji DNA zaproponowany przez Watsona i Cricka (1953). Kilka lat później Crick wygłosił wykład podczas sympozjum „Biological Replication of Macromolecules” [Biologiczna replikacja makrocząsteczek], które odbyło się w University College London. Wystąpienie to, opublikowane w 1958 roku pod tytułem „On protein synthesis” [O syntezie białek] (Crick, 1958), uważane jest za niezwykle ważne, formułujące tzw. centralny dogmat biologii molekularnej: DNA tworzy RNA, a RNA tworzy białko. Od tego czasu genów nie uważano już za tajemnicze „jednostki dziedziczenia”, jak po odkryciach Grzegorza Mendla, ale za fragmenty łańcucha DNA, w których zapisywane są informacje o strukturze RNA i białek. Główny wysiłek skierowano odtąd na odszyfrowanie sposobu działania genów.

W tym kontekście nie sposób nie wspomnieć o pracy François Jacoba i Jacquesa Monoda z Instytutu Pasteura w Paryżu. Badali oni geny zaangażowane w metabolizm laktozy u bakterii *Escherichia coli*. Odkryli, że geny te, *lacZ*, *lacY* i *lacA*, kodujące odpowiednio  $\beta$ -galaktozydazę, permeazę i transacetylazę, znajdują się blisko siebie na chromosomie *E. coli* i podlegają transkrypcji na pojedynczej cząsteczce matrycowego RNA (mRNA). Transkrypcja rozpoczyna się od elementu zwanego promotorem, rozpoznawanego przez polimerazę RNA. Proces ten zależy od dwóch innych genów, operatora i regulatora. Operator znajduje się blisko pierwszego genu strukturalnego (*lacZ*), a regulator jest genem kodującym represor – białko, które wiąże się z operatorem i zapobiega rozpoczęciu transkrypcji przy braku laktozy. Promotor, operator i geny strukturalne tworzą tzw. operon. Transkrypcja operonu laktozy zachodzi tylko przy braku glukozy, preferowanego źródła węgla, oraz w obecności laktozy będącej inhibitorem represora. François Jacob i Jacques Monod wraz z André Lwoffem zostali wspólnie uhonorowani Nagrodą Nobla w dziedzinie fizjologii lub medycyny w 1965 roku „za odkrycia dotyczące genetycznej kontroli syntezy enzymów i wirusów”.

Kolejne lata przyniosły dalsze informacje na temat struktury i funkcji genomów. Stwierdzono, że u eukariontów większość genów nie jest ciągła. Składają się z eksonów i intronów. Są one przepisywane na pre-mRNA, który podlega procesowi zwanemu splicingiem, przeprowadzanemu przez małe jądrowe rybonukleoproteiny (snRNP), które wiążą się z końcami 5' i 3' intronu i powodują, że intron tworzy pętlę. Intron jest następnie usuwany z sekwencji, a dwa pozostałe eksony są ze sobą łączone.

Na początku lat siedemdziesiątych XX wieku rozwinęła się nowa gałąź genetyki nazwana inżynierią genetyczną. Możliwe stało się wprowadzanie precyzyjnych zmian w materiale genetycznym i przenoszenie genów z jednego organizmu do drugiego. Klasycznym przykładem użytecznej manipulacji genetycznej jest konstrukcja szczepu bakteryjnego zawierającego ludzki gen kodujący insulinę. Insulina to lek niezastąpiony w leczeniu cukrzycy. Wcześniej pozyskiwany był z trzustek świń i krów; teraz jest łatwo dostępny dzięki możliwości jego ekstrakcji z bakterii rosnących w ogromnych fermentorach. Rozwój inżynierii genetycznej wzbudził pewne obawy związane z bezpieczeństwem eksperymentów. Teoretycznie możliwe było wykorzystanie tych technik do stworzenia niebezpiecznych mikroorganizmów do użytku wojskowego. W 1975 roku Paul Berg, jeden z twórców inżynierii

genetycznej, zorganizował w Asilomar Conference Center (Kalifornia, USA) konferencję poświęconą wszelkim możliwym zastosowaniom nowych technik. Około 150 genetyków z całego świata dyskutowało o zagrożeniach związanych z inżynierią genetyczną. Uczestniczyłem w tej konferencji na zaproszenie Davida Baltimore'a, naukowca z Cancer Center of the Massachusetts Institute of Technology, który wraz z Renato Dulbecco i Howardem Martinem Teminem otrzymał Nagrodę Nobla „za odkrycia dotyczące interakcji między wirusami nowotworowymi a materiałem genetycznym komórek”. Drugim polskim uczestnikiem Konferencji Asilomar był prof. Wacław Gajewski, kierownik Zakładu Genetyki Uniwersytetu Warszawskiego.

Ostatecznym wnioskiem płynącym z Konferencji Asilomar było kontynuowanie rozwoju inżynierii genetycznej przy jednoczesnym dbaniu o prowadzenie eksperymentów w warunkach laboratoryjnych adekwatnych do potencjalnie najbardziej niebezpiecznego elementu danego eksperymentu. Na przykład konstruowanie szczepu *Escherichia coli* wytwarzającego insulinę można przeprowadzać praktycznie w każdym laboratorium biologicznym, gdyż ani ten gatunek bakterii, ani insulina nie są niebezpieczne dla człowieka. Natomiast eksperymenty na genetycznie zmodyfikowanych szczepach wykorzystywanych do opracowywania leków na choroby zakaźne powinny być prowadzone tylko w specjalnie wyposażonych laboratoriach, dających pewność, że groźne dla człowieka bakterie nie wydostaną się do środowiska. James Watson, laureat Nagrody Nobla za pracę nad strukturą DNA, był zdecydowanym zwolennikiem intensywnego rozwoju inżynierii genetycznej. W czasie, gdy trwała Konferencja Asilomar, był doradcą rządu Stanów Zjednoczonych do spraw broni biologicznej. Twierdził, że wie o naturalnie występujących, niemodyfikowanych inżynieryjnie szczepach bakterii przechowywanych w magazynach wojskowych USA mogących zabić każdego mieszkańca naszej planety na osiem różnych sposobów i że w porównaniu z tym wytwory manipulacji genetycznych to drobnostka.

Dzięki postępowi w technikach sekwencjonowania DNA uzyskano ogromną liczbę nowych informacji na temat struktury genomu. W połowie lat 80. XX wieku zsekwencjonowanie pełnego ludzkiego genomu wydawało się niemożliwe, głównie ze względu na koszt takiego projektu – wyższy niż roczny budżet Instytutu Maxa Plancka. Pomimo wielu obiekcji Human Genome Project [Projekt poznania genomu ludzkiego] został uruchomiony w 1990 roku. Prowadzony był przez International Human Genome Sequencing

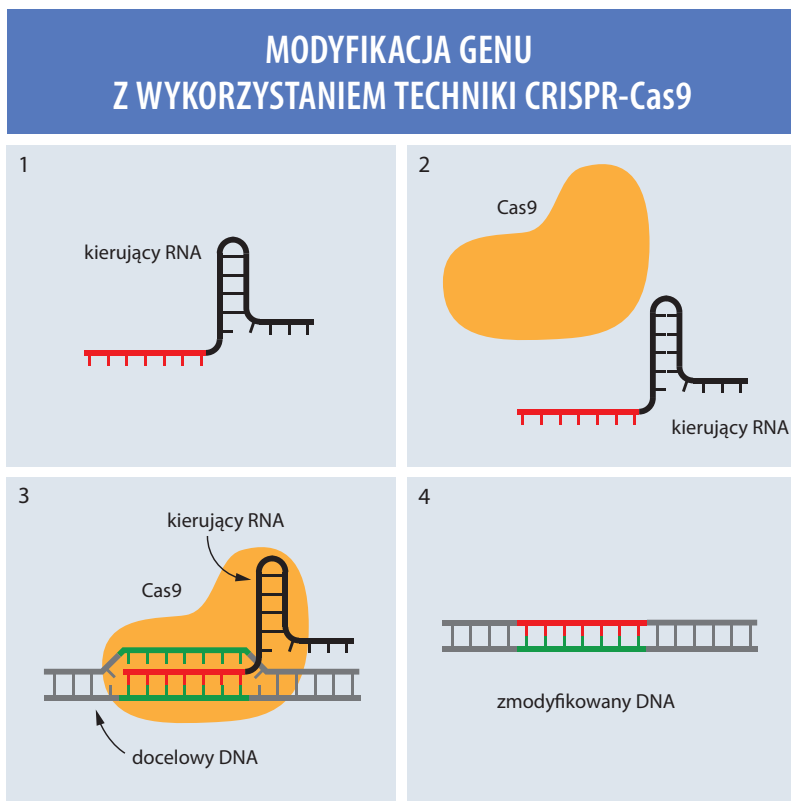
Consortium (IHGSC) złożone z naukowców z 18 krajów. Zgodnie z jedną z podstawowych zasad obowiązujących członków konsorcjum wszystkie informacje zbierane w ramach tego projektu dotyczące sekwencji genomu ludzkiego były publicznie dostępne w ciągu 24 godzin od ich uzyskania. W 1998 roku do wyścigu w sekwencjonowaniu ludzkiego genomu przystąpiła prywatna firma biotechnologiczna Celera Genomics prowadzona przez prof. Craiga Ventera. Celera zapowiedziała zsekwencjonowanie całego ludzkiego genomu w ciągu trzech lat. W lutym 2001 roku obie grupy w dwóch oddzielnych artykułach (Lander i in., 2001; Venter i in., 2001) jednocześnie opublikowały szkice sekwencji ludzkiego genomu. Dzięki stałemu doskonaleniu metod sekwencjonowania DNA i dobrej współpracy między obiema grupami dotarły one na metę jednocześnie, a oba projekty zostały ukończone przed terminem. IHGSC realizował drugą końcową fazę projektu (IHGSC, 2004), mającą na celu uzupełnienie luk i ustalenie sekwencji DNA w niejednoznacznych obszarach, które pozostały nieoznaczone. W końcowej fazie zsekwencjonowano 99% ludzkiego genomu. Jego ostateczna forma zawierała 2,85 miliarda nukleotydów z przewidywanym wskaźnikiem błędu 1 zdarzenia na 100 000 zsekwencjonowanych zasad. Ponadto IHGSC zmniejszył liczbę luk 400-krotnie: z 147 821 luk pozostało jedynie 341. Pozostałe były związane z trudnymi technicznie regionami chromosomów. Chociaż wcześniejsze wstępne publikacje przewidywały aż 40 000 genów kodujących białka, w fazie końcowej szacunki te zmniejszyły się do przedziału między 20 000 a 25 000.

Dziś bez trudu można dostrzec spektakularny sukces projektu. Jego zakończenie zapoczątkowało nową erę w medycynie. Teraz możemy się dowiedzieć, które z naszych chorób mają podłoże genetyczne i zidentyfikować odpowiedzialne za nie geny. Human Genome Project przyniósł również znaczny postęp w technologiach wykorzystywanych do sekwencjonowania DNA.

Szczególnie uderzającym odkryciem Human Genome Project jest stwierdzenie, że sekwencja nukleotydów jest prawie identyczna (99,9%) u dwóch dowolnych osób. Jednak nawet pojedyncza zmiana nukleotydu w jednym genie może być odpowiedzialna za chorobę. Dlatego poznanie sekwencji genomu ludzkiego przyczyniło się również w znacznym stopniu do wyjaśnienia mechanizmów molekularnych leżących u podstaw wielu ludzkich chorób. Co więcej, połączenie podejścia cytogenetycznego ze znajomością

cią sekwencji genomu przenosi naszą wiedzę na temat chorób człowieka na zupełnie nowy poziom. I chociaż na początku do Human Genome Project odnoszono się sceptycznie, z pewnością zostanie uznany za jedno z najważniejszych przedsięwzięć naukowych naszych czasów.

Całkiem niedawno opracowano nową, potężną metodę inżynierii genetycznej, noszącą nazwę CRISPR-Cas9 (ryc. 1). Jest ona skutecznym i niezawodnym sposobem dokonywania precyzyjnych, ukierunkowanych zmian



**Rycina 1.** Poglądowy schemat metody CRISPR-Cas9 (metoda edycji genów CRISPR): 1) zaprojektowany i zsyntetyzowany kierujący RNA (ang. *guide RNA*) z fragmentem pasującym do docelowej sekwencji DNA (czerwony); 2) kierujący RNA jest dodawany do komórki docelowej wraz z białkiem Cas9; 3) kierujący RNA łączy się z pasującym docelowym fragmentem DNA gospodarza (zielony), po czym białko Cas9 tworzy dwuniciowe pęknięcie w precyzyjnie określonym miejscu docelowym; 4) pożądane modyfikacje sekwencji DNA można wprowadzić dokładnie w wybranym miejscu w genomie.

w genomie żywych komórek. Geny CRISPR (są to zgrupowane, regularnie rozmieszczone, krótkie powtórzenia palindromiczne; ang. *clustered regularly interspaced short palindromic repeats*) i gen Cas kodujący nukleazę (ang. *CRISPR-associated protein*) mają kluczowe znaczenie dla adaptacyjnej odporności bakterii i archeonów, umożliwiając tym organizmom reakcję na inwazyjny materiał genetyczny i jego eliminację.

Powtórzenia palindromiczne odkryto w latach 80. u *E. coli* (Ishino i in., 1987), ale ich funkcja nie została potwierdzona do roku 2007, kiedy to Barrangou i współpracownicy wykazali, że *S. thermophilus* może nabyć oporność na bakteriofaga w wyniku integracji fragmentu genomu infekującego wirusa do swojego locus CRISPR (Barrangou i in., 2007). Zidentyfikowano trzy typy mechanizmów CRISPR, z których typ II jest najlepiej zbadany. W tym przypadku DNA atakujących wirusów lub plazmidów zostaje pocięty na małe fragmenty i włączony do locus CRISPR w otoczenie serii krótkich powtórzeń – około 20 par zasad (pz). Loci są transkrybowane, a transkrypty następnie przetwarzane w celu wygenerowania małych RNA (crRNA – CRISPR RNA), wykorzystywanych do kierowania endonukleaz do komplementarnych sekwencji w atakującym DNA.

Stosując technikę CRISPR-Cas9, można precyzyjnie zniszczyć gen docelowy lub zmienić gen zmutowany w jego wariant prawidłowy. Oczekuje się, że technika ta zostanie wykorzystana do „naprawy” genów odpowiedzialnych za choroby uwarunkowane genetycznie; zdaniem wielu naukowców nie jest ona jednak jeszcze całkowicie bezpieczna i może pociągać za sobą niepożądane skutki. Mimo to grupa chińskich badaczy podjęła pierwsze próby i wstrzyknęła człowiekowi komórki zawierające edytowane geny. Uważali oni, że edycja genów może poprawić zdolność komórek odpornościowych do zwalczania nowotworu. Zespół kierowany przez onkologa Lu You z Uniwersytetu Syczuańskiego w Chengdu dostarczył zmodyfikowane komórki pacjentowi z agresywnym nowotworem płuc w Szpitalu Zachodnio-chińskim (Liang i in., 2015). Naukowcy usunęli komórki odpornościowe z krwi biorcy, a następnie zakłócili kodowanie genu dla białka PD-1, które normalnie hamuje odpowiedź immunologiczną komórki: nowotwory wykorzystują tę funkcję do proliferacji. Edytowane komórki hodowano w celu zwiększenia ich liczby i wstrzykiwano z powrotem pacjentowi. Żywiono nadzieję, że bez PD-1 edytowane komórki zaatakują komórki nowotworowe i pokonają chorobę. Planuje się zastosowanie tej samej techniki zarówno

w Chinach, jak i USA do leczenia pacjentów z innymi rodzajami nowotworów (Cyranoski, 2016).

Nowe techniki analizy DNA – poza ich zastosowaniami w medycynie – są bardzo ważnym narzędziem do badania ewolucji żywych organizmów. Odkrycie, że DNA może przetrwać w antycznych szczątkach roślinnych lub zwierzęcych i może być amplifikowany za pomocą reakcji łańcuchowej polimerazy, a następnie zsekwencjonowany, miało duży wpływ na badania nad ewolucją. Pierwsze opublikowane artykuły sugerowały, że dzięki badaniu kopalnego DNA można cofnąć się nawet o setki milionów lat. Jednak doniesienia (Cano i in., 1993) o sekwencjach DNA uzyskanych z owadów zatopionych w bursztynie okazały się nieprawdziwe. Takie owady są datowane na ponad sto milionów lat, podczas gdy wiadomo, że DNA prawdopodobnie nie może przetrwać dłużej niż około miliona lat. Nawet jednak badając okres tak krótki geologicznie, można uzyskać odpowiedzi na ważne pytania dotyczące ewolucji, systematyki, paleoekologii, genezy chorób i procesów ewolucyjnych na poziomie populacji.

## Antyczny DNA

Każda próbka DNA uzyskana z pochodzących z wykopalisk szczątków ludzkich, roślinnych lub zwierzęcych nazywana jest antycznym DNA (aDNA). Analizy aDNA przyciągnęły uwagę nie tylko biologów molekularnych, ale także opinii publicznej. Było to wynikiem kilku spektakularnych odkryć, zwłaszcza tych dotyczących historii naszego gatunku. Badania antycznego DNA ujawniły złożoność historii współczesnego człowieka. Udowodniły zarazem, że nasi przodkowie krzyżowali się w okresie środkowego oraz wczesnego górnego paleolitu z neandertalczykami – każdy współczesny nie-Afrykanin ma w swoim genomie domieszkę (od 1 do 4%) DNA neandertalczyka. Doszło do licznych niezależnych krzyżowań, które obejmowały nie tylko neandertalczyków, ale także denisowian oraz innych, niezidentyfikowanych, homininów. Badania aDNA ujawniły m.in. skandynawskie pochodzenie dynastii Rurykowiczów oraz pewne przesłanki o takim też pochodzeniu dynastii Piastów. Dzięki analizom DNA możliwe było prześledzenie ścieżek migracji ludzi oraz relacji między populacjami zamieszkującymi obecnie Ziemię.

Niemal równie interesujące są wyniki badań historii gatunków zwierząt i roślin, także tych, które wyginęły tysiące lat temu lub też całkiem niedawno. Istnieje wiele gatunków zwierząt, które znamy tylko z ich szczątków zachowanych w wiecznej zmarzlinie lub głęboko w jaskiniach; niektóre słynne przykłady to mamut, niedźwiedź jaskiniowy, nosorożec włochaty czy ptak dodo z Mauritiusa. Badając ich DNA, możemy ustalić relacje między nimi a ich wciąż żyjącymi kuzynami. Istnieją nawet projekty bliskie powieściom science-fiction, których celem jest ożywienie niektórych z tych gatunków w procesie zwanym de-ekstynkcją; na przykład poprzez wprowadzenie DNA wyizolowanego ze szczątków mamuta do komórek jajowych słonia.

Postęp w dziedzinie inżynierii genetycznej umożliwia transfer genów od jednego osobnika do drugiego przy wykorzystaniu DNA wyizolowanego z żywych gatunków lub z pozostałości wymarłych. Obecnie powszechną praktyką jest ulepszanie gatunków roślin poprzez wyposażanie ich w gen bakteryjny, warunkujący uodparnianie ich na owadzie szkodniki.

W kolejnych rozdziałach przedstawiono metody analizy aDNA oraz ich zastosowanie w badaniach ewolucji człowieka oraz wybranych gatunków roślin i zwierząt.