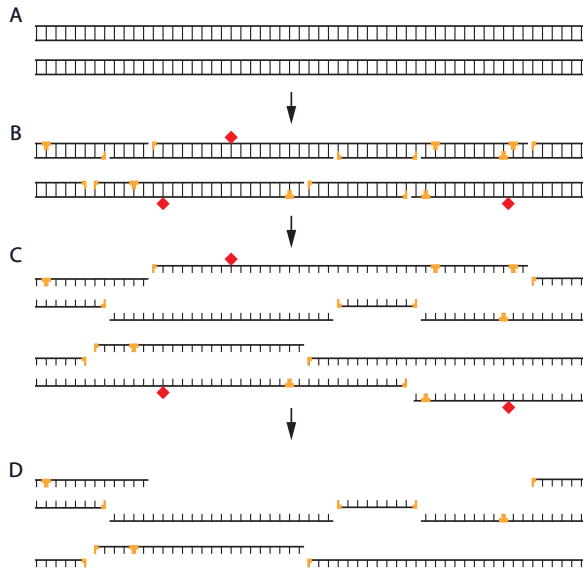


METODY ANALIZY aDNA

Metody analizy antycznego DNA są zazwyczaj zasadniczo takie same jak w przypadku współczesnego DNA. Trzeba tylko pamiętać, że aDNA jest zwykle mocno zdegradowany. Jego degradacja zależy od wieku i warunków (m.in. temperatura, wilgotność) dominujących w miejscu, w którym znaleziono materiał kopalny. Najczęściej cząsteczki aDNA są mocno rozdrobnione i noszą ślady modyfikacji chemicznych ze względu na wiele procesów pośmiertnych, takich jak aktywność enzymatyczna lub spontaniczny rozpad makromolekuł (ryc. 2). Sygnał z tak zdegradowanego materiału genetycznego może być łatwo zagłuszony przez zaledwie kilka zanieczyszczających cząsteczek długołańcuchowego nienaruszonego współczesnego DNA. Dlatego badacze zajmujący się aDNA powinni przedsięwziąć liczne środki ostrożności przeciwko zanieczyszczeniu próbki. Zazwyczaj próbki poddawane analizom aDNA mają znacznie niższą zawartość rzeczywistego aDNA niż zanieczyszczającego współczesnego DNA ze środowiska czy pochodzącego od osób zajmujących się pobraną próbką. Dlatego ważne jest, aby ekstrakcję i wszystkie inne manipulacje w przypadku aDNA przeprowadzać w specjalnych, odpowiednio przygotowanych do pracy z aDNA pomieszczeniach. Laboratorium aDNA to zazwyczaj przestrzeń fizycznie oddzielona od laboratoriów, w których wykonywane są prace nad współczesnym DNA. Powinna być często naświetlana promieniowaniem UV, które niszczy cząsteczki DNA znajdujące się na powierzchni i unoszące się w powietrzu, minimalizując w ten sposób ryzyko zanieczyszczenia. Dostęp do takich laboratoriów musi być ściśle ograniczony do osób przeprowadzających eksperyment. Osoby te powinny nosić sterylne fartuchy laboratoryjne, maseczki i rękawiczki (ryc. 3). Wejście do laboratorium powinno być możliwe



Rycina 2. Schematyczne przedstawienie procesu degradacji DNA. A) Długie nienaruszone cząsteczki DNA po śmierci komórki gospodarza ulegają zaawansowanemu rozkładowi. B) Cząsteczki, które nie zostały strawione przez enzymy lityczne uwalniane podczas śmierci komórkowej lub przez mikroorganizmy żerujące na rozkładającym się ciele, z czasem akumulują pęknięcia w niciach (przedstawione na niciach cząsteczek DNA jako przerwy), pojawiają się także spontaniczne uszkodzenia chemiczne (czerwone romby i żółte znaczniki). C) Fragmenty DNA dostępne do tworzenia bibliotek DNA są pofragmentowane i zawierają modyfikacje chemiczne, z których część uniemożliwia dalszą analizę poprzez np. blokowanie enzymów wykorzystywanych podczas przygotowywania biblioteki DNA lub odczytu sekwencji DNA (takie zmiany są tu reprezentowane przez czerwone romby). D) Zsekwencjonowana zostaje tylko część cząsteczek DNA. Są one zwykle krótkie i noszą ślady modyfikacji chemicznych powodujących nieprawidłowy odczyt nukleotydu podczas sekwencjonowania DNA (takie zmiany reprezentowane są tutaj przez żółte znaczniki). Zmiany te częściej występują w pobliżu pęknięć nici i zwykle powodują fałszywy odczyt T (tymina) zamiast właściwego C (cytozyna).
Przygotowanie: Martyna Molak

tylko przez podwójne drzwi; z oddzielającym je pomieszczeniem, w którym przechowywane są sterylne fartuchy laboratoryjne, maseczki na twarz i buty.

Materiałem używanym do analizy aDNA są zwykle zęby i kości. Jednak aDNA udało się wyekstrahować także z takich materiałów jak: włosy, zmumifikowane tkanki miękkie, pióra, skorupki jaj, gleba, a nawet koprolity (subskamieniałe odchody).

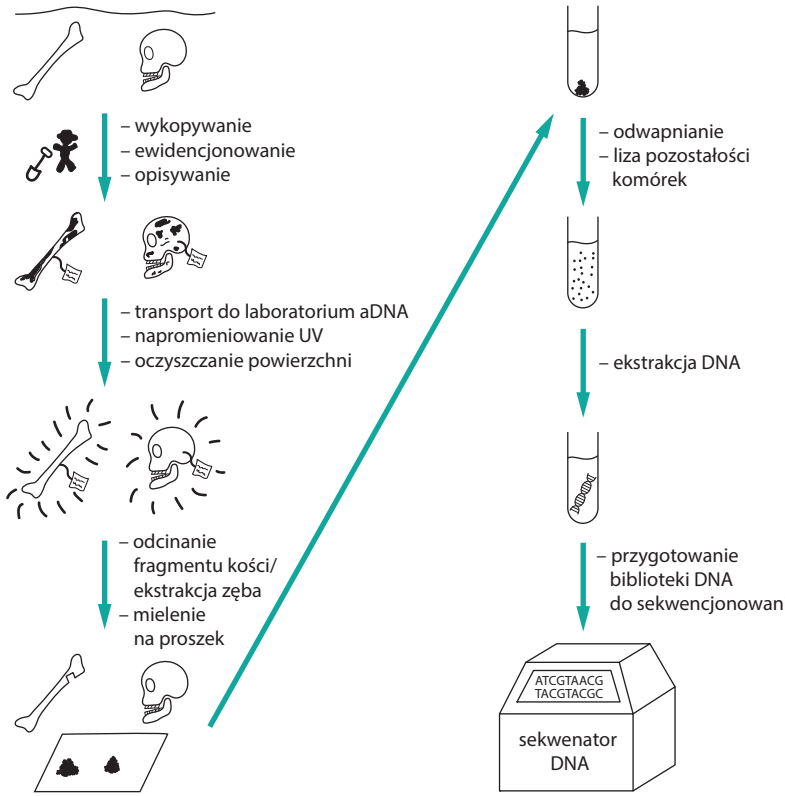


Rycina 3. Laboratorium antycznego DNA. Badaczka przygotowująca próbki kości do ekstrakcji DNA ubrana jest w odzież ochronną, aby zminimalizować ryzyko zanieczyszczenia próbki współczesnym DNA. Praca wykonywana jest w komorze z przepływem laminarnym usytuowanej w wydzielonym, przeznaczonym wyłącznie do pracy z aDNA, pomieszczeniu.

Zdjęcie: Martyna Molak

Kości, ze względu na ich powszechność, są najczęstszym źródłem antycznego DNA. Pomyślnie ekstrahowano materiał genetyczny nawet z kości mających ponad 700 000 lat (Orlando i in., 2013). Zawsze jednak należy brać pod uwagę fakt, że w kościach poziom zanieczyszczenia przez DNA drobnoustrojów jest wysoki, a w większości kości ponad 99% DNA to DNA egzogenny. Jednak moc analityczna, uzyskana dzięki zwiększeniu liczby odczytów do milionów niezależnych sekwencji DNA, znacznie zmniejsza problemy wynikające z zanieczyszczenia antycznych próbek przez współczesny DNA.

Dziedzina tych badań rozwinęła się jeszcze bardziej, gdy zdano sobie sprawę, że włosy stanowią niezwykłe źródło wysokiej jakości aDNA, który można poddać skutecznym procedurom dekontaminacji, nie mającym



Rycina 4. Schemat działań podczas analizy aDNA. Materiał organiczny do analizy aDNA często pochodzi z wykopalisk archeologicznych lub paleontologicznych. Po badaniach i opisie archeologicznym, paleontologicznym i/lub antropologicznym wybrane fragmenty materiału są transportowane do laboratorium aDNA w celu przeprowadzenia analiz genetycznych. Powierzchnia materiału jest zwykle czyszczona podchlorynem (wybielaczem) i/lub poddana promieniowaniu UV, aby maksymalnie usunąć wszelkie powierzchniowe zanieczyszczenia współczesnym DNA. Docelowy fragment próbki – zwykle gęsty fragment kości, w którym pośmiertne uszkodzenia DNA są stosunkowo najmniejsze – zostaje rozdrobniony. Proszek jest następnie poddawany działaniu odczynników chemicznych, które uwalniają cząsteczki DNA z pozostałości komórek i struktur tkankowych. Uwolniony DNA jest ekstrahowany metodą wybraną z szeregu dostępnych metod w zależności od rodzaju i stanu materiału. Oczyszczony izolat DNA poddaje się dalszej obróbce w celu uzyskania sekwencji DNA do analiz genetycznych. Zwykle proces ten obejmuje przekształcanie cząsteczek DNA w bibliotekę DNA, która umożliwia amplifikację całego DNA zawartego w izolacie, dzięki czemu cząsteczki są dostępne do wykrywania oraz sekwencjonowania (odczytu sekwencji nukleotydowej DNA poszczególnych cząsteczek) za pomocą platformy Illumina.

Przygotowanie: Martyna Molak

zastosowania do kości. DNA można również ekstrahować ze skorupek jaj. Antyczny DNA uzyskany ze skorupek jaj jest stosunkowo bogaty w endogenny DNA i znacznie mniej zanieczyszczony przez DNA bakteryjny niż kości. Innymi źródłami aDNA są koprolity i gleba, ale ekstrakcja w ich przypadku jest trudna, a aDNA zwykle silnie zmodyfikowany i zdegradowany.

W celu wyekstrahowania niewielkiej ilości endogennego DNA powierzchnię próbki zazwyczaj albo fizycznie usuwa się przez piaskowanie, albo przynajmniej naświetla promieniowaniem UV i czyści roztworem podchlorynu, co niszczy powierzchniowy DNA, pochodzący głównie z zanieczyszczeń.

Następnie, aby umożliwić odczynnikom wchodzenie w interakcję z materiałem, próbka musi zostać sproszkowana lub przynajmniej rozdrobniona na małe kawałki. W przypadku materiału kostnego i zębów często stosuje się specjalne młynki, służące do rozdrabniania twardej tkanki na proszek, który jest następnie zwykle odwapniany przy użyciu roztworu chelatującego (np. EDTA). Pozostałości błon komórkowych są rozrywane przez surfaktanty, a białka trawione przez enzymy proteolityczne. DNA ekstrahowany jest z takiego lizatu przy użyciu jednej z wielu metod opartych m.in. na rozpuszczalności DNA w różnych substancjach (metoda fenol : chloroform) lub na hydrofobowym wiązaniu między DNA a krzemionką (roztwór krzemionki i kolumny krzemionkowe) albo na wiązaniu DNA z powierzchniami magnetycznymi w określonych warunkach (ekstrakcja za pomocą kulek magnetycznych). Otrzymany czysty ekstrakt DNA następnie poddaje się amplifikacji (powieleniu) metodą PCR (ang. *polimerase chain reaction*) wybranych markerów genetycznych (w celu umożliwienia odczytania ich sekwencji) lub przygotowywana jest z niego biblioteka DNA, czyli dokonywana jest obróbka enzymatyczna cząsteczek DNA pozwalająca na sekwencjonowanie całych, ale przeważnie krótkich cząsteczek DNA z ekstraktu (zwana również unieśmiertelnianiem DNA). Patrz: rycina 4 przedstawiająca pozyskiwanie informacji genetycznej z antycznych materiałów.

Markery genetyczne

Chcąc oszacować zakres różnic genetycznych lub podobieństw między osobnikami lub populacjami, opieramy się na markerach genetycznych. Jako markery wykorzystuje się różne rodzaje polimorfizmów genetycznych, w tym

polimorfizmy pojedynczego nukleotydu (ang. *single nucleotide polymorphisms; SNPs*), proste polimorfizmy długości sekwencji (ang. *simple sequence length polymorphisms; SSLPs*) i restrykcyjne polimorfizmy długości fragmentów (ang. *restriction fragment length polymorphisms; RFLPs*). SSLPs obejmują sekwencje powtórzeń zwane minisatelitami (zmienna liczba powtórzeń tandemowych, ang. *variable number of tandem repeats; VNTRs),* mikrosatelity (proste powtórzenia tandemowe, ang. *simple tandem repeats; STRs*) oraz insercje/delecje (indele). W ludzkim genomie najczęstsze typy markerów to SNP, STR i indele. SNPs dotyczą pojedynczej zasady we fragmencie DNA. STRs to markery, w których fragment sekwencji powtarza się kilka razy z rzędu, a liczba powtórzeń jest zmienna w obrębie poszczególnych osobników i pomiędzy nimi. Indele to polimorfizmy, w przypadku których fragment sekwencji DNA występuje u niektórych osobników w populacji (allel insercyjny), podczas gdy u innych nie występuje (allel delecyjny).

Dopóki nie doszło do szybkiego postępu w technikach sekwencjonowania DNA, głównym przedmiotem badań dotyczących DNA były wirusy i plazmidy bakteryjne. Wynikało to po prostu z ich niewielkich rozmiarów i łatwości izolacji. W przypadku organizmów wyższych z tych samych powodów intensywnie badany był mitochondrialny DNA (mtDNA). Każda komórka zawiera od setek do tysięcy mitochondriów, które znajdują się w cytoplazmie. U ludzi mitochondrialny DNA składa się z 16 569 pz i zawiera 37 genów, z których wszystkie są niezbędne do prawidłowego funkcjonowania mitochondriów. Trzynaście z tych genów koduje enzymy zaangażowane w fosforylację oksydacyjną. Pozostałe geny kodują transferowy RNA (tRNA) i rybosomalny RNA (rRNA), cząsteczki zaangażowane w syntezę białek.

Ostatnio badania nad aDNA ułatwił intensywny rozwój technologii sekwencjonowania DNA drugiej generacji. Technika HTS (sekwencjonowanie o wysokiej przepustowości, ang. *high-throughput sequencing*) pozwala na sekwencjonowanie bibliotek aDNA na platformach takich jak Illumina NovaSeq. Są to techniki przystępne cenowo nawet dla małych laboratoriów, chociaż wiele laboratoriów decyduje się na wysyłanie swoich bibliotek do sekwencjonowania wyspecjalizowanym firmom, takim jak Beijing Genomics Institute. Możliwe jest użycie techniki HTS do sekwencjonowania tylko jednego wybranego locus w całej populacji. To podejście nosi nazwę „metabarkodowania" (ang. *metabarcoding*) i jest stosowane do sekwencjonowania

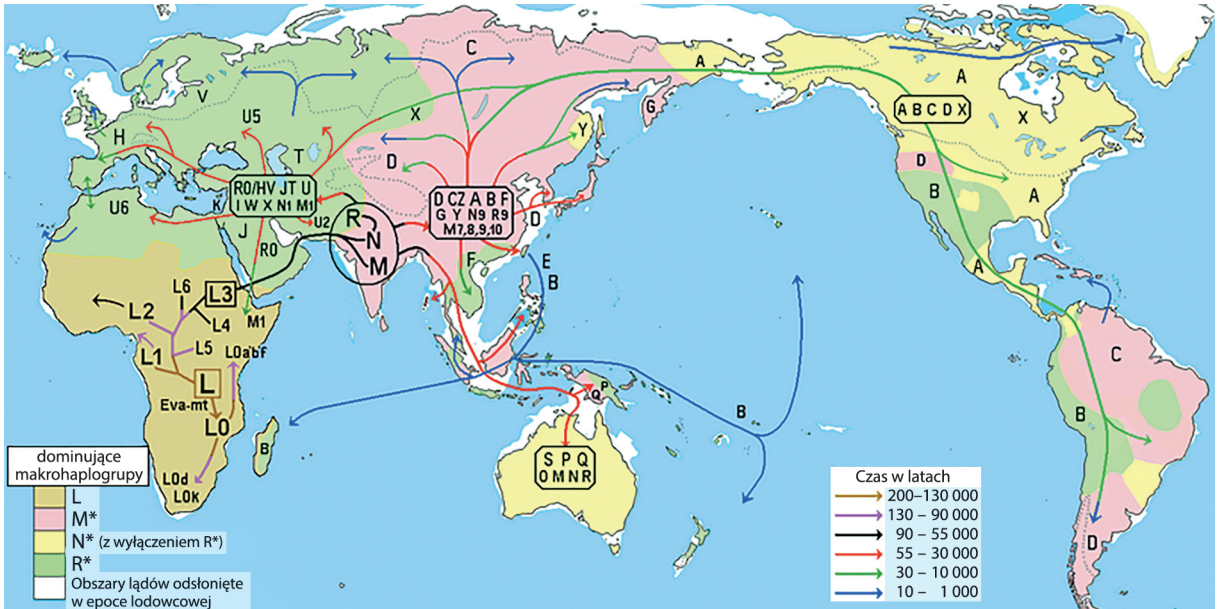
próbek pobranych np. z wody jeziora, osadów itp. Główną zaletą metabarkodowania jest możliwość jednoczesnej analizy dużej liczby próbek przy ograniczonych kosztach.

Według Gnerre'a i in. (2011) technologie masowego równoległego sekwencjonowania DNA rewolucjonizują genomikę, umożliwiając generowanie miliardów stosunkowo krótkich (~100 zasad) odczytów sekwencji przy bardzo niskich kosztach. Autorzy ci przedstawili algorytm (ALLPATHS-LG) do składania genomu ludzkiego i mysiego, wygenerowanych na platformie Illumina. Badacze ci uważają, że połączenie ulepszonej technologii sekwencjonowania i ulepszonej metody obliczeniowej umożliwi drastyczny wzrost sekwencjonowania *de novo* dużych genomów. Może to być szczególnie przydatne w badaniach aDNA organizmów, dla których nie dysponuje się zsekwencjonowanym genomem żadnego z bliskich krewnych, co jest częste w przypadku wielu wymarłych gatunków.

Mitochondrialny DNA

Przez długi czas mitochondrialny DNA był głównym materiałem wykorzystywanym w badaniach antycznego DNA. Wynikało to z faktu, że genom mitochondrialny ma liczne kopie w każdej komórce, a zatem celując w dowolny marker genetyczny, znajdujący się w genomie mitochondrialnym, można było liczyć na większe prawdopodobieństwo jego udanej amplifikacji niż w przypadku markerów jądrowych.

Fragment niekodujący genomu mitochondrialnego – pętla D (ang. *D-loop*) – ze względu na niski poziom presji selekcyjnej, a zatem praktycznie niemal nieograniczone tempo mutacji, jest bardzo zmienny u większości gatunków i dlatego często wykorzystuje się go w analizach wewnątrzgatunkowej struktury genetycznej. Inne markery mitochondrialne, takie jak cytochrom *b*, są bardziej konserwowane i są przydatne do analiz z udziałem wielu gatunków. Ponadto mitochondrialny DNA jest zasadniczo dziedziczony ściśle po matce, co oznacza, że zawiera tylko część historii genetycznej osobnika, a to znacznie ułatwia śledzenie genealogii i tym samym interpretację wyników badań opartych na markerach mitochondrialnych w porównaniu do markerów autosomalnych. Te właśnie cechy mitochondrialnego DNA sprawiły, że stał się on źródłem markerów pierwszego wyboru nie tylko dla



Rycina 5. Drogi rozprzestrzeniania się linii matczynych (reprezentowane przez haplogrupy genomu mitochondrialnego) podczas ekspansji człowieka poza Afrykę. Główne haplogrupy i miejsce, w których dominują we współczesnych populacjach, oznaczono nazwami haplogrup (pojedyncze litery, czasem z dodaną cyfrą). Ustalone trasy i czas poszczególnych wędrówek oznaczono strzałkami i kolorami. Ostatnia wspólna przodkini wszystkich współczesnych genomów mitochondrialnych (tzw. „mitochondrialna Ewa" należąca do makrohaplogrupy L) żyła prawdopodobnie w Afryce Południowej około 200 tys. lat temu. Stamtąd wraz z migracjami ludzi rozprzestrzeniały się następnie potomne linie rodowe. Część linii L3 opuściła Afrykę około 60–70 tys. lat temu i zaludniła resztę świata, docierając do Azji Południowo-Wschodniej i Australii około 50 tys. lat temu, Azji Środkowej nieznacznie później, Europy Zachodniej około 40 tys. lat temu i obu Ameryk nie później niż 14 tys. lat temu. Rozmieszczenie haplogrup mitochondrialnych na całym świecie z reguły odzwierciedla szlaki migracji ludzi. Określenie haplogrupy może więc pomóc w poznaniu historii rodowej danej osoby od strony matki.

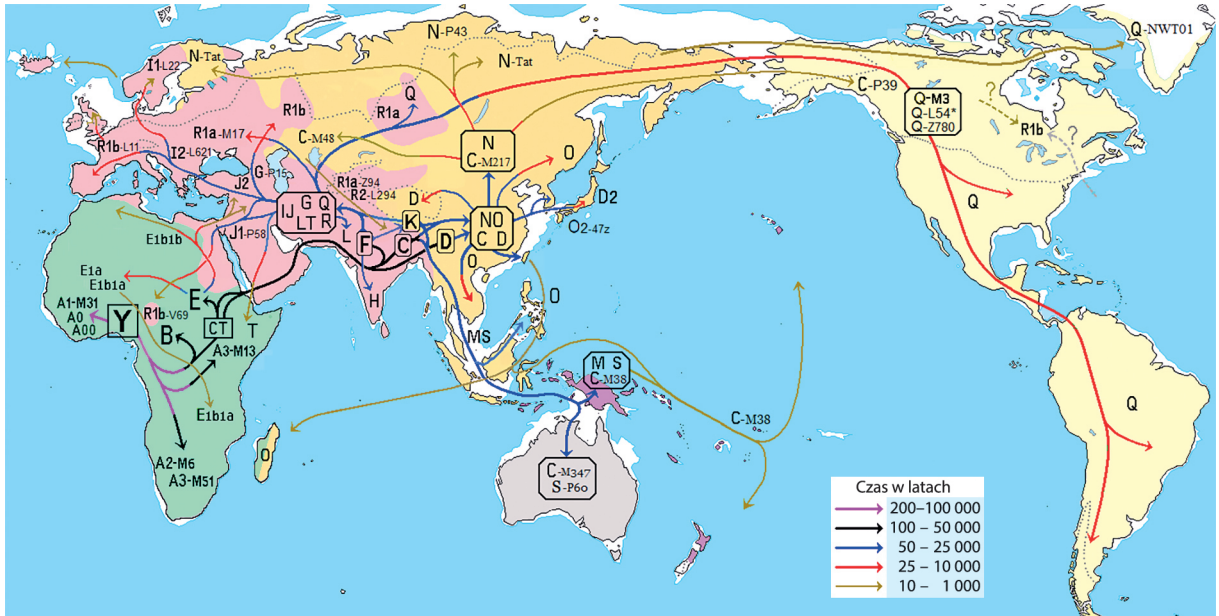
Autor: Maulucioni CC 3.0 Źródło: https://commons.wikimedia.org/wiki/File:Human_migrations_and_mitochondrial_haplogroups.PNG

badań nad antycznym DNA, ale także dla dużej części filogenetycznych i populacyjnych badań w ogóle. Wczesne badania genetyczne populacji ludzkiej z zastosowaniem DNA, chociaż zapoczątkowane w latach 80. przez pionierskie analizy Luki Cavalli-Sforzy z użyciem różnych markerów genetycznych, opierały się głównie na danych mitochondrialnych. Linie mitochondrialne powstawały w toku migracji ludzkich, które zaczęły się prawdopodobnie gdzieś w południowej Afryce i zaowocowały zaludnieniem prawie każdego kawałka lądu na Ziemi, zróżnicowane rozmieszczenie geograficzne linii mitochondrialnych odzwierciedla więc historię tych migracji. Linie pogrupowano w haplogrupy według wspólnego pochodzenia i przypisano im oznaczenia według liter alfabetu dla łatwej identyfikacji. Dzięki markerom mitochondrialnym zrekonstruowano drogi i czas ekspansji i różnicowania się poszczególnych linii (ryc. 5). Wnioski wyciągnięte na podstawie analizy danych mitochondrialnych stały się punktem odniesienia dla dalszych analiz genomowych.

Chromosom Y

Ludzki chromosom Y to struktura składająca się z około 60 Mpz (milionów par zasad), odpowiedzialna za ekspresję 45 białek, w tym niektórych związanych z płcią i płodnością. Oprócz tych genów chromosom Y zawiera wiele powtórzeń sekwencji bez znanej funkcji, których usunięcie nie pociąga za sobą widocznych konsekwencji fenotypowych. Geny na chromosomie Y można podzielić na dwie kategorie: te ulegające ekspresji w całym ciele we wszystkich rodzajach komórek i te, które najczęściej lub wyłącznie ulegają ekspresji w jądrach (gonadach męskich), prawdopodobnie zaangażowane w rozwój jąder i/lub spermatogenezę. Jeden z tych genów (*Sry*), koduje białko regulatorowe, kontrolujące ekspresję genów determinujących płeć. Gen *Sry* jest niezbędny, aby włączyć ekspresję *Sox9*, determinującą rozwój gonad męskich. Samice XX nie mają regionu *Sry*, a zatem nie zachodzi u nich aktywacja *Sox9* i w zamian rozwijają się jajniki. Ewolucję i strukturę chromosomu Y w pracy przeglądowej opisała Graves (2006).

W większości badań aDNA pochodzącego z antycznych skamieniałości wykorzystywano mtDNA. Postęp w technikach ekstrakcji i sekwencjonowania DNA sprawił, że możliwe stało się wykorzystanie do analiz DNA jądrowego.



Rycina 6. Drogi rozprzestrzeniania się linii ojcowskich (reprezentowanych przez haplogrupy chromosomu Y) w okresie ekspansji poza Afrykę. Główne haplogrupy i miejsce ich dominacji we współczesnych populacjach są oznaczone nazwami haplogrup (alfanumerycznie, czasami z nazwą SNP określającego daną linię). Ustalone trasy i czas poszczególnych wędrówek oznaczono strzałkami i kolorami. Ostatni wspólny przodek wszystkich współczesnych chromosomów Y (tzw. „Y-chromosomalny Adam", oznaczony tutaj jako „Y") mieszkał prawdopodobnie w środkowo-północno-zachodniej Afryce około 200 tys. lat temu. Drogi rozprzestrzeniania się są w dużej mierze zgodne z mitochondrialnym DNA (ryc. 5), chociaż niektóre szczegóły wyraźnie się różnią. Rozmieszczenie haplogrup chromosomu Y na całym świecie w większości odzwierciedla drogi migracji ludzi, a tym samym poznanie haplogrupy może pomóc w poznaniu historii rodowej danej osoby od strony ojca.

Autor: Maulucioni CC BY-SA 3.0
Źródło: https://commons.wikimedia.org/wiki/File:Migraciones_humanas_en_haplogrupos_de_ADN-Y.PNG)