

STANISŁAW CEBRAT, DOROTA MACKIEWICZ, WOJCIECH WAGA

Zakład Genomiki

Wydział Biotechnologii Uniwersytetu Wrocławskiego

S. Przybyszewskiego 63/77, 51-148 Wrocław

E-mail: cebrat@smorfland.uni.wroc.pl

KOMPUTEROWE MODELOWANIE EWOLUCJI GENOMÓW

WSTĘP

Chcąc przybliżyć biologom problem ewolucji genomów, używając w tym celu modeli komputerowych, nie można uniknąć wprowadzenia dotyczącego sensu samego modelowania. Biologowie bardzo często uzasadniają swoją niechęć do matematycznych, fizycznych lub komputerowych modeli zjawisk biologicznych ich zbyt dużym upraszczaniem. Nie zdają sobie najczęściej sprawy z tego, że sami są mistrzami modelowania. Eksperyment wykonany na stole biologa molekularnego ma na ogół dostarczyć cegiełkę do naszego wyobrażenia o tym, co dzieje się na przykład w żywej komórce – przecież modelu tego, co się tam dzieje! Uproszczenie biologa najczęściej polega na tym, że jego model jest małym wycinkiem całej sieci zależności między wielką liczbą elementów. Z drugiej strony należy sobie zdać sprawę z tego, że uproszczenie – jeżeli właściwe – nie jest słabością modelu, ale na ogół stanowi o jego sile. Świetną ilustracją takiego twierdzenia mogą być prawa Keplera. Zostały one stworzone przy założeniu, że planety są punktami. Absurd, nawet my nie jesteśmy punktami, a co dopiero Ziemia, po której stąpamy!

Model musi być prosty i musi mieć możliwie mało parametrów. Każde zjawisko można by opisać dowolnie dokładnie wprowadzając do modelu dowolnie wielką liczbę parametrów, ale taki model byłby bezużyteczny, ponieważ nie potrafilibyśmy zinterpretować znaczenia tych parametrów, nie mówiąc o tym, że nie byłibyśmy w stanie przeprowadzić samego modelowania. Modele mogą opisywać zjawiska mniej skomplikowane, jak na

przykład dyfuzję przez błony, lub nieco bardziej skomplikowane, jak konformację białek w otoczeniu innych cząsteczek. Mogą też być stosowane do opisywania skrajnie skomplikowanych procesów, jak ewolucji genomów i populacji. Można jeszcze zapytać, po co nam model ewolucji genomów, skoro znamy już setki pełnych sekwencji genomów organizmów o różnym stopniu pokrewieństwa i możemy już sobie wyobrazić ich ewolucję.

Nasza intuicja i wyobrażenia opierają się często na ogólnie przyjętych za paradygmaty wynikach modelowania neo-darwinowskiej teorii ewolucji. Jak bardzo jesteśmy przyzwyczajeni do posługiwania się „klasycznym” modelowaniem populacji mendlowskiej może świadczyć fakt, że nie możemy pogodzić się z wynikami analiz genetycznych przeprowadzonych na Islandii stwierdzającymi, że liczba dzieci w małżeństwach maleje wraz ze spadkiem pokrewieństwa między małżonkami – dane zebrane z niemal 200 lat (HELGASON i współaut. 2008). Ten fenomen biologiczny jest faktem, chociaż niezgodnym z naszymi intuicyjnymi wyobrażeniami (modelami?) ewolucji populacji. Tym gorzej dla faktów?!

Należy wspomnieć o jeszcze jednym interesującym aspekcie modelowania – pokazuje ono, że wiele zjawisk biologicznych można opisywać relacjami znalezionymi dla zupełnie innych zjawisk fizycznych – podobne równania potęgowe opisują wielkość i częstość występowania trzęsień ziemi, lawin śnieżnych w górach i ... częstości rekombinacji między homologicznymi chromosomami.

W artykule tym będą przedstawione wyniki modelowania zakładającego przestrzenną organizację populacji i genomów, które często nie odpowiadają wynikom uzyskiwanym przez dotychczasowe modelowanie (średniopolewe), stosowane do opisu ewolucji populacji mendlowskiej. Zostanie przedstawiony jeden model ewolucji populacji ze strukturą wiekową, od najprostszych jego postaci, i wręcz trywialnych wyników, do prób odpo-

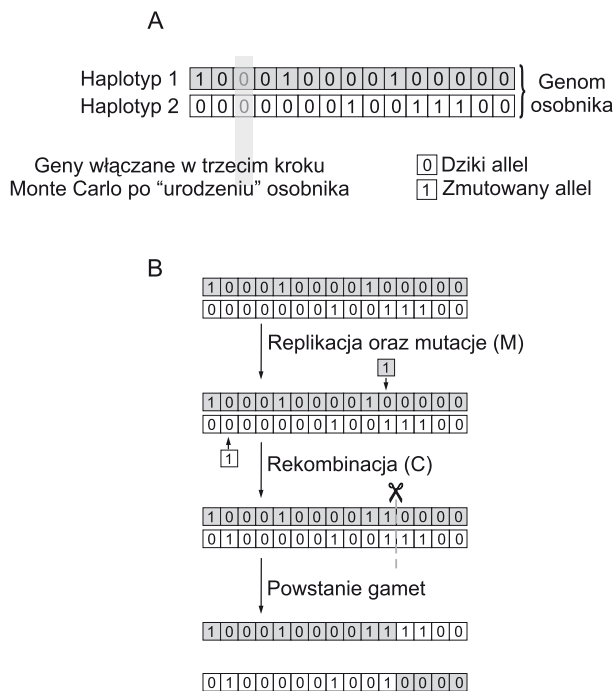
wiadania na coraz bardziej skomplikowane pytania i porównywania uzyskanych wyników modelowania z wynikami analiz genomowych baz danych. Takie podejście pozwala na ilościowe opisywanie zjawisk biologicznych, przybliża możliwość matematycznego opisu tych zjawisk i stwarza możliwości wytłumaczenia wielu fenomenów przyrodniczych.

MODEL PENNA – KOMPUTEROWY MODEL EWOLUCJI POPULACJI ZE STRUKTURĄ WIEKOWĄ

Model PENNA (1995) jest modelem typu Monte Carlo (MC), w którym każde zdarzenie zachodzi z określonym prawdopodobieństwem i w wyniku przeprowadzonego losowania albo zajdzie, albo nie. Populacja składa się z osobników reprezentowanych przez ich genomy, w naszej wersji – zawsze diploidalne i składające się z jednej lub więcej par autosomów, ewentualnie dodatkowo z chromosomów płciowych. Każdy autosom (haplotyp) reprezentowany jest przez ciąg bitów – zer i jedynek. Jeden bit reprezentuje jeden gen; jeżeli jego wartość wynosi 0, to reprezentuje on allel dziki – funkcjonalny. Jeżeli jego wartość wynosi 1, to reprezentuje on allel zmutowany – niefunkcjonalny. Wszystkie zmutowane allele są recesywne, to znaczy, że obydwa allele w odpowiadających sobie pozycjach w obydwu chromosomach muszą być defektywne, żeby funkcja tego locus była również defektywna, czyli żeby defekt wyraził się fenotypowo. Minimalną liczbę fenotypowych defektów powodujących śmierć genetyczną określa parametr T . Najbardziej charakterystyczną cechą modelu Penna jest chronologiczne włączanie funkcji genów. W pierwszym kroku czasowym (określanym krokiem Monte Carlo) włączane są allele w pierwszym locus, w drugim kroku allele drugiego locus itd. Tak więc, zarówno liczba włączonych loci, jak i liczba kroków MC, którą „przeżył” dany osobnik odpowiadają jego wiekowi i po odpowiednim przeskalowaniu możemy posługiwać się skalą czasową mierzoną w dniach (na przykład dla muszki owocowej) lub latach dla populacji ludzkich. Strukturę genomu i przyjętą terminologię przedstawiono na Ryc. 1. Szczegółowy opis modelu i jego zastosowanie w badaniach biologicznych i demograficznych można znaleźć w pracach przeglądowych (MOSS de OLIVEIRA i współaut. 1999, ŁASZKIEWICZ i CEBRAT 2005, STAUFFER i współaut. 2006).

Osobniki w modelu Penna rozmnażają się płciowo, chociaż w podstawowej wersji ich genomy nie posiadają genetycznej informacji odpowiedzialnej za warunkowanie płci. Wprowadzenie tej informacji w specjalnej wersji modelu prowadzi do zaskakujących wniosków na temat ewolucji chromosomów płciowych, „kurczącego” się chromosomu Y i altruizmu osobników męskich (BIECEK i CEBRAT 2008), co zostanie opisane poniżej.

Symulacja ewolucji populacji zaczyna się od określonej liczby osobników N_0 , których płeć losuje się z równym prawdopodobieństwem – żeńską albo męską. Nazwijmy je odpowiednio kobietami i mężczyznami. Ich wiek jest losowany od 1 do wieku odpowiadającego liczbie bitów w ciągu bitowym (L). Do ciągów bitowych wprowadza się w losowych miejscach defektywne bity (1) z określonym prawdopodobieństwem – rzędu kilku procent (genomy osobników w początkowej populacji mogą być też doskonałe, to znaczy, że wszystkie bity ustawione są na 0). Każdemu osobnikowi sprawdza się jego fenotyp, zliczając liczbę loci odpowiadających jego wiekowi, które zawierają obydwa allele defektywne. Jeżeli ta liczba osiągnie wartość T , to osobnik jest eliminowany (poddawany śmierci genetycznej). Jako jeden z parametrów modelu przyjmuje się minimalny wiek rozrodczy osobników (R). Osobniki, które osiągnęły ten wiek mogą się rozmnażać. Kobieta losowo wybiera mężczyznę w wieku rozrodczym i oboje produkują gamety. W tym celu genomy obu osobników są replikowane. W losowo wybranym miejscu każdego chromosomu zachodzi mutacja z prawdopodobieństwem M . Jeżeli wybrany do mutacji bit ma wartość 0, to zamieniany jest na 1, jeżeli posiada wartość 1, to pozostaje 1 – nie ma rewersji. Kopie chromosomów rekombinują w procesie odpowiadającym crossing-over z prawdopodobieństwem C w losowo wybra-



Ryc. 1. Procesy genetyczne uwzględniane w diploidalnym modelu Penna.

A – genom osobnika złożony z dwóch ciągów bitowych (haplotypów). Allele w poszczególnych loci włączane są w kolejnych krokach Monte Carlo. B – Osobnik przed przystąpieniem do rozmnażania kopiuje obydwa ciągi bitowe, wprowadza przy tym do każdej kopii w losowo wybranym miejscu mutację, a następnie przeprowadzany jest crossing-over z określonym prawdopodobieństwem w losowo wybranym miejscu.

nym miejscu. Losowo wybrany jeden produkt rekombinacji, jeżeli ona zaszła, albo jedna z kopii chromosomów, jeżeli do rekombinacji

nie doszło, odpowiada gamecie. Gamety pary osobników łączą się tworząc zygotę. Jej płeć ustala się losowo z prawdopodobieństwem 0,5. Pierwsza para genów tej zygoty zostanie sprawdzona w następnym kroku MC. W ten sposób każda kobieta w populacji w wieku rozrodczym w jednym kroku MC produkuje dziecko z prawdopodobieństwem B lub liczbę B dzieci. Zależnie od parametrów ewolucji populacja ma tendencję do nieograniczonego wzrostu, na przykład przy wysokim B, albo obumiera, na przykład przy wysokiej częstotliwości mutacji. Aby uniknąć nieograniczonego wzrostu populacji, wprowadza się tak zwany czynnik Verhulsta. Jego wielkość określa równanie logistyczne:

$$V = 1 - N_t / N_{\max}$$

gdzie:

V – prawdopodobieństwo przeżycia osobnika w danym kroku. W tym przypadku śmierć osobnika następuje z powodów losowych, niezależnie od jego statusu genetycznego i aktualnego fenotypu;

N_t – aktualna liczba osobników w populacji;

$N_{\max}$ – maksymalna liczba osobników, jaką może pomieścić środowisko, zwana pojemnością środowiska.

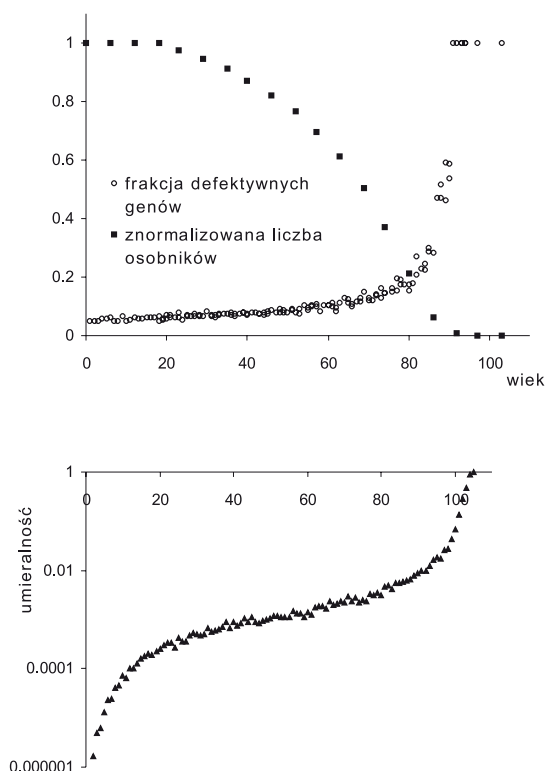
Łatwo zauważyć, że jeżeli liczba osobników osiągnie wartość $N_{\max}$, to prawdopodobieństwo przeżycia osobnika równa się zero. Populacja nie może więc całkowicie zapełnić środowiska. W wersjach modelu opisywanych w tej pracy czynnik Verhulsta ogranicza szansę przeżycia jedynie noworodków (S'A MARTINS i CEBRAT 2000). Osobniki starsze mogą umrzeć tylko z powodu śmierci genetycznej.

METODY OPISYWANIA POPULACJI W SYMULACJACH

Symulacje ewolucji populacji pozwalają na ilościowe opisywanie struktury genetycznej poszczególnych osobników oraz całej puli genetycznej, struktury wiekowej populacji, umieralności, pozwalają wartościować populację z punktu widzenia ewolucyjnego (jakie populacje wygrywają w warunkach konkurencyjnych), a nawet pozwalają na obserwowanie samoorganizacji pewnych mechanizmów genetycznych.

Na Ryc. 2 przedstawiono wykresy opisujące strukturę wiekową populacji, strukturę genetyczną w formie frakcji defektywnych genów w zależności od ich pozycji na chromosomie oraz zależność umieralności od

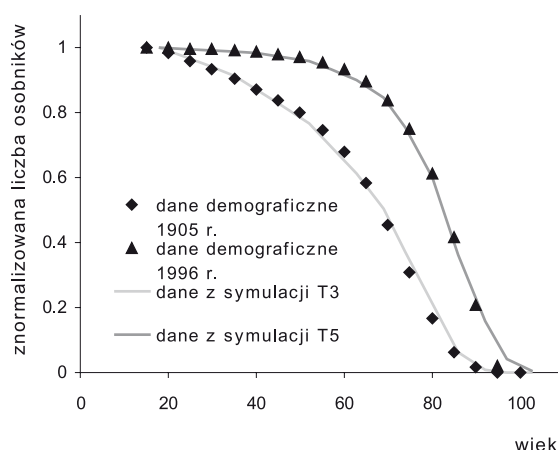
wieku. Ten ostatni schemat opisu jakości populacji został zaproponowany przez B. Gomperta niemal 200 lat temu (GOMPERTZ 1825). Proszę zwrócić uwagę na skalę logarytmiczną osi y. W przybliżeniu liniowa zależność obserwowana w centralnej części wykresu świadczy o wykładniczym wzroście umieralności wraz z wiekiem osobników. Na Ryc. 3 porównano wyniki modelowania struktury wiekowej populacji symulowanych z rzeczywistymi danymi demograficznymi populacji niemieckiej. Zgodność tych dwóch struktur jest zaskakująco wysoka. W tym miejscu należy zwrócić uwagę na możliwość interpretacji wyników modelowania. Struktura wiekowa



Ryc. 2. Graficzny opis populacji.

Na górnym wykresie przedstawiono rozkład defektywnych genów wzdłuż chromosomu, w zależności od wieku osobnika, w którym te geny ulegają ekspresji. Rozkładowi defektywnych genów odpowiada struktura wiekowa populacji. Na dolnym wykresie przedstawiono umieralność w populacji przedstawionej powyżej (należy zwrócić uwagę na skalę logarytmiczną osi y).

populacji niemieckiej zmieniła się dramatycznie w ciągu ostatniego stulecia. Aby uzyskać odpowiednie zmiany w strukturze wiekowej symulowanych populacji, należało zmienić wartość tylko jednego parametru – T. Jest to parametr określający minimalną liczbę defektów fenotypowych, które zabijają osobnika. Można więc przyjąć, że jest to parametr opisujący naszą relację ze środowiskiem – stopień podporządkowania sobie środowiska naszym potrzebom albo dopasowania się nas do środowiska, nasz styl życia, opiekę medyczną itp. Na uwagę zasługuje również struktura puli genetycznej populacji (Ryc. 2).



Ryc. 3. Struktura wiekowa populacji wirtualnych i rzeczywistych populacji niemieckich na początku i końcu dwudziestego wieku.

Zmiana struktury demograficznej Niemiec może być uzyskana również w populacjach komputerowych przy jedynym założeniu, że śmierć osobników następuje na skutek większej liczby (T) ekspresjonowanych niedoborów genetycznych (NIEWCZAS i współaut. 2000).

Frakcja defektywnych genów, ulegających ekspresji przed minimalnym wiekiem rozrodczym, jest niska i są one równomiernie rozłożone. Frakcja defektywnych genów wśród genów ulegających ekspresji po minimalnym wieku rozrodczym rośnie wraz z wiekiem i osiąga dla najpóźniej włączanych genów wartość 1. To znaczy, że osobniki nie dysponują informacją pozwalającą im na przeżycie w późnej starości. Model ten potwierdza hipotezę starzenia się MEDAWARA (1952) zakładającą, że za starzenie się osobników odpowiedzialne są defekty akumulowane w genach ulegających ekspresji w wieku późniejszym. Za prawdziwością tej hipotezy przemawiają jeszcze inne wyniki symulacji i cechy puli genetycznej. Według modelu, zwiększenie parametru T w ewoluującej populacji powoduje, że osobniki dożywają późniejszego wieku, ale za śmierć osobników odpowiedzialne są inne defekty niż było to wcześniej, „odslaniają się” nowe choroby, wcześniej rzadkie lub nieznanne – w naturalnej populacji ludzkiej np. choroba Parkinsona, Alzheimer, nowotwory.

DŁUGOŚĆ TELOMERÓW I STARZENIE SIĘ

Jedna z hipotez starzenia się, niewykluczająca się z hipotezą Medawara, zakłada, że

za starzenie się odpowiada liczba możliwych podziałów komórkowych, która ograniczo-

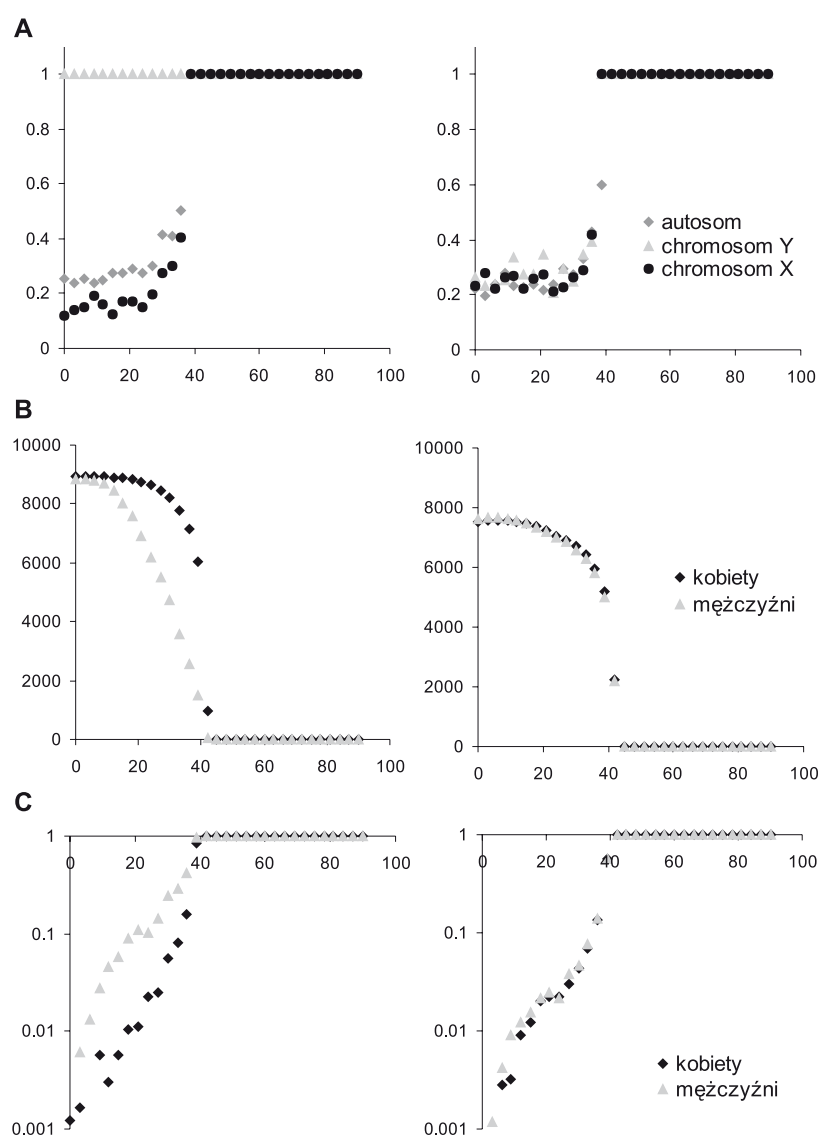
na jest długością telomerów, a właściwie ich skracaniem się przy kolejnych replikacjach chromosomów (HAYFLICK, MOORHEAD 1961). Modelowanie hipotezy telomerowej procesu starzenia się prowadzi do wniosku, że wydłużając telomery, na przykład w procesie klonowania, można by wydłużyć życie osobników. Jeżeli owieczka Dolly żyła krócej, ponieważ już w chwili urodzenia miała skró-

cone telomery, to ich wydłużenie mogłoby zapobiec jej wcześniejszej śmierci. Niestety, próby odsuwania starości i śmierci przez wydłużanie telomerów powyżej ich długości ustalonej ewolucyjnie zwiększałyby prawdopodobieństwo śmierci z powodu nowotworów we wczesnych okresach rozwoju (MASA i współaut. 2006).

EWOLUCJA CHROMOSOMÓW PŁCIOWYCH

Jednym z bardziej interesujących zjawisk w ewolucji genomów jest ewolucja chromosomów płciowych. W różnych liniach filogenetycznych pojawiają się różne sposoby na kontrolowanie płci osobników. W wielu przypadkach, osobniki odrębnych płci różnią się kariotypem. U ssaków, samice mają dwa identyczne chromosomy X, natomiast samce mają jeden chromosom X i jeden chromosom Y. Chromosom Y niesie relatywnie mało informacji genetycznej, przede wszystkim jest to informacja odpowiedzialna za rozwój męskich cech płciowych. Biologowie raczej zgodnie przyjmują hipotezę, że obydwie chromosomy mają to samo, autosomalne pochodzenie. Nie są jednak już tak zgodni, co do mechanizmów ewolucyjnych, które doprowadziły do tak dramatycznej utraty przez samce informacji z jednego chromosomu. Ta degeneracja chromosomu Y zdestabilizowała męski genotyp, przez co stał się on bardziej podatny na zgubny efekt mutacji w chromosomie X. Wszystkie defekty w genach chromosomu X mogą być widoczne w fenotypie męskim, podczas gdy u kobiet te same defekty mogą być maskowane przez funkcjonalne allele z drugiego chromosomu X. Może to być przyczyną tego, że mężczyźni żyją krócej niż kobiety. Większą umieralność mężczyzn zauważono już trzysta lat temu winiąc za to ich hulaszczy tryb życia (ARBUTHNOT 1710). Autor przyjął, że sposobem na wyrównanie liczby kobiet i mężczyzn w wieku rozrodczym byłoby większe prawdopodobieństwo urodzenia chłopca, co potwierdził wynikami analiz przeprowadzonych dla kolejnych 85 roczników. Jedną z hipotez ewolucji chromosomów płciowych u ssaków zakłada, że kilka kolejnych aberracji w obrębie chromosomu Y spowodowało utratę możliwości jego rekombinowania z chromosomem X (FERGUSON-SMITH 2007). To wywołało jego postępującą degradację, zgodnie z zasadą działania zapadki Mullera. Zgodnie z tą zasadą, kolejne mu-

tacje w chromosomie Y, komplementowane przez geny w chromosomie X, ulegały akumulacji, a ponieważ nie mogły być eliminowane przez rekombinację – mogło być tylko gorzej, nie mogło być lepiej. Symulacje komputerowe z użyciem modelu Penna sugerują jednak inny scenariusz. Jeżeli przyjmiemy, że początkowo chromosomy płciowe były identyczne, a jedyną różnicą między nimi były geny odpowiedzialne za determinowanie płci męskiej ulokowane na chromosomie Y, to chromosom ten tracił pozostałą informację genetyczną, a presja selekcyjna faworyzowała utratę zdolności rekombinacji chromosomu Y z chromosomem X (LOBO i ONODY 2005). Za taki przebieg ewolucji chromosomu Y odpowiedzialne są zachowania seksualne i strategia rozrodcza ssaków (BIECEK i CEBRAT 2008). Na ogół populacje ssaków są poligamiczne, a ich potencjał rozrodczy zależy od wielkości frakcji samic będących w wieku rozrodczym. Mniejsza frakcja samców niż samic nie zmniejsza tego potencjału, a nawet, przyjmując ograniczoną pojemność niszy ekologicznej, mniejsza liczba samców może sprzyjać zwiększeniu konkurencyjności populacji. Tak więc, recesywne mutacje na chromosomie Y mogą być łatwiej tolerowane przez selekcję niż mutacje genów na chromosomie X. Akumulacja tych mutacji prowadzi do sytuacji, w której defekty na chromosomie X w genomie męskim są widoczne (początkowo były to układy homozygotyczne, a następnie, po delecji genu z chromosomu Y, hemizygotyczne). Te same recesywne mutacje w genomach kobiet mogą być komplementowane przez dzikie kopie genów z drugiego chromosomu X. Gdyby rekombinacja między chromosomami X i Y była zachowana, to defektywne geny z chromosomu Y mogłyby być przenoszone na chromosom X, co powodowałoby większą umieralność samic i obniżenie potencjału rozrodczego populacji. Selekcja faworyzuje więc wyłączenie możliwości rekombinacji



Ryc. 4. Ewolucja chromosomów płciowych ssaków.

W panelu lewym przedstawiono wyniki symulacji, przy założeniu, że mężczyzna może być wybierany jako partner seksualny przez wiele kobiet w tym samym kroku Monte Carlo. W prawym panelu przedstawiono wyniki symulacji, gdy partnerzy byli sobie wierni. A – w chromosomie Y wszystkie geny zostały uszkodzone, chromosom X zawiera mniej defektów niż autosom, ponieważ są one usuwane przez selekcję z genomów męskich. Wierność partnerów zapewnia tę samą strukturę chromosomów niezależnie od płci. B – struktura wiekowa jest konsekwencją struktury genetycznej, umieralność mężczyzn w każdej grupie wiekowej w populacji poligamicznej jest wyższa niż kobiet (widoczne również na wykresie C).

między chromosomami X i Y (BIECEK i CEBRAT 2008). Stosując modelowanie komputerowe można łatwo pokazać przebieg ewolucji chromosomów płciowych potwierdzający opisany scenariusz (patrz seria wykresów na Ryc. 4). Można również podjąć próbę falsyfikacji tej hipotezy, deklarując w modelu inną strategię rozrodczą – na przykład, że samiec jest niezbędnie potrzebny do wychowania potomstwa i nie może mieć potomstwa równocześnie z dwiema lub większą liczbą samic. Przekładając to na stosunki w populacji człowieka – mężczyzna i kobieta są sobie wierni co najmniej przez jeden rok, a w szczególności kobieta nie będzie uwodzić mężczyzn, którzy już mają dziecko z inną kobietą. Ta strategia zakłada taką samą presję selekcyjną na obie płcie – utrzymanie liczebności kobiet i mężczyzn w wieku rozrodczym na tym

samym poziomie. Po zadeklarowaniu w modelu takiej strategii rozrodczej, chromosom Y nie traci informacji genetycznej, struktura wiekowa populacji kobiet i mężczyzn jest taka sama i selekcja nie dopuszcza do utraty zdolności rekombinacji między różnymi chromosomami płciowymi (Ryc. 4 prawy panel). Przyjęcie hipotezy ewolucji chromosomów płciowych, wynikającej z komputerowego modelowania, wyjaśnia wiele innych cech ewolucyjnych populacji ssaków. Ponieważ decydującym parametrem określającym potencjał rozrodczy populacji jest frakcja samic a nie samców, samce mogą zachowywać się altruistycznie, ryzykując nawet własnym życiem. Ponadto, wyższy wiek samca zdaje się wskazywać na jego lepszy genotyp, czego najlepszym dowodem jest fakt, że przeżył do późniejszego wieku. W związku z tym sami-

ce mogą mieć tendencje do wyboru partnerów starszych od siebie. W wielu kulturach jest to dopuszczalne, podczas gdy sytuacja odwrotna uważana jest za mezalians. Z tego wynika też, że bardziej altruistycznie powinny zachowywać się młode samce, dlatego na froncie pojawiają się młodzi mężczyźni, a na tyłach starsi generałowie. Kobiety w ogóle na wojnę nie idą, ale są cennymi brankami, zwiększającymi potencjał rozrodczy zwycięskiej populacji.

Opisana strategia ewolucji chromosomów płciowych ma jeszcze jedną charakterystyczną cechę – wybór tej strategii jest nieodwra-

calny. Takson (gatunek), który raz wszedł na tę drogę ewolucji nie może już odmienić roli płci i płć męska jest już skazana na to by być płcią biologicznie słabszą. Jeżeli rola płci może się zmieniać okresowo, to bardziej „ekonomiczną” i bezpieczną staje się strategia, w której płć męska jest homogametyczną, a płć żeńska heterogametyczną, jak u ptaków. Wtedy nawet poligamia czy polianndria (ogólnie promiskuityzm) nie prowadzą do eliminacji informacji genetycznej jednego z chromosomów.

CZĘSTOŚĆ REKOMBINACJI WZAJEMNEJ A EWOLUCJA GENOMÓW DIPLOIDALNYCH

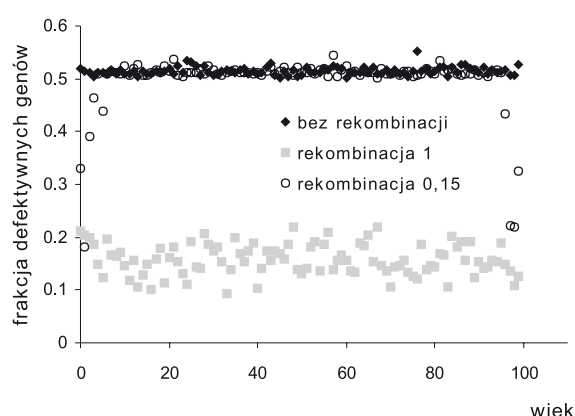
Ogólnie przyjmowany jest pogląd, że rekombinacja jest zjawiskiem niezbędnym dla utrzymania bioróżnorodności populacji oraz, że brak rekombinacji musi prowadzić do śmierci populacji ze względu na działanie wcześniej już wspomnianej zapadki Mullera. Genomy w czasie replikacji muszą ulegać mutacji z pewną skończoną częstością, a ponieważ częstość mutacji powrotnych (rewersji) albo mutacji korzystnych jest znikoma, można się spodziewać, że będzie tylko gorzej. Rekombinacje stwarzają natomiast możliwość utworzenia konfiguracji, która może być lepsza od każdej z form rodzicielskich. Do badania i opisu ewolucji populacji, od niemal 80 lat używano modeli średniopoloowych, zakładających, że populacje są olbrzymie i panmiktyczne (każda samica może szukać samca w całej populacji). W takich populacjach równie uzasadnionym jest przyjęcie, że allele różnych genów dziedziczą się niezależnie. Na takich założeniach oparte jest prawo Hardy'ego-Weinberga i model Fishera-Wrighta (HARDY 1908, FISHER 1930). Nic więc dziwnego, że w takich populacjach dochodzi do „homogenizacji” puli genetycznej i trudno wyobrazić sobie w nich specjację sympatryczną, kry-

tykowaną przez kilkadziesiąt lat przez Mayra i sprowadzaną do zjawiska pomijalnego w ewolucji (MAYR 1942). Specjacja sympatryczna jest powstawaniem nowych gatunków w obrębie jednej populacji, na tym samym terenie, w tej samej niszy ekologicznej. Dopiero w ostatnich latach specjacja sympatryczna zaczyna być zauważalna i dopuszczalna w rozważaniach ewolucyjnych (JIGGINS 2006). W rzeczywistości powinniśmy uwzględnić przestrzenne rozmieszczenie zarówno genów w genomach, jak i osobników w przestrzeni, którą aktualnie zajmują. Wiadomo, że geny tworzą grupy sprzężeniowe (na przykład chromosomy), a duże populacje nie są panmiktyczne. Nawet przy założeniu całkowitego braku preferencji seksualnych, osobniki mogą być skazane na wybór partnera znajdującego się w ograniczonej odległości, w przypadku ludzi mogą to być wybory ograniczone do tej samej miejscowości, grupy etnicznej, językowej czy nawet zawodowej. Zasadnym jest więc pytanie, jak tworzą się grupy sprzężeniowe genów w zależności od gęstości kodowania lub częstości rekombinacji i czy zależy to od stopnia pokrewieństwa między rodzicami.

SELEKCJA OCZYSZCZAJĄCA A STRATEGIA KOMPLEMENTOWANIA HAPLOTYPÓW

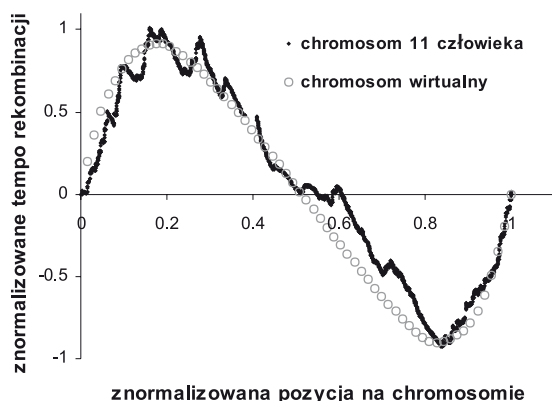
Wyobraźmy sobie na wstępie model prostszy od opisanego wcześniej modelu Penna. Genom złożony jest z jednej pary chromosomów o długości 100 genów każdy. Mutacje genów są zawsze letalne, ale recesywne, nie ma rewersji. Organizmy giną, jeżeli co najmniej w jednym locus obydwu allele są defektywne albo losowo z prawdo-

podobieństwem 2%. W czasie produkcji gamet dochodzi do jednej mutacji i jednego crossing-over w losowym miejscu. Możemy więc powiedzieć, że odległość między genami wynosi 1 centyMorgan. Po dłuższym czasie symulacji częstość defektywnych genów w puli genetycznej ustala się na poziomie około 15%, a defekty są rozmieszczone rów-



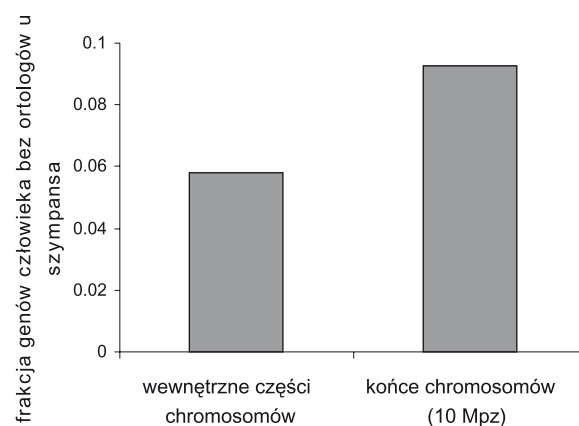
Ryc. 5. Rozkład defektywnych genów na chromosomach w zależności od częstości rekombinacji.

Przy wysokiej częstości rekombinacji (1 na chromosom o zawartości 100 bitów) selekcja oczyszczająca utrzymuje relatywnie niską liczbę defektów i są one równomiernie rozłożone. Wyłączenie rekombinacji powoduje nagromadzenie defektów, ale i powstanie haplotypów komplementujących. Przy pośredniej częstości rekombinacji (0,15) rozkład defektów staje się nierównomierny, wewnątrz utrzymuje strukturę komplementującą, podczas gdy na końcach chromosomów obserwuje się efekty selekcji oczyszczającej.



Ryc. 6. Nierównomierne rozłożenie defektów w komplementujących chromosomach w pobliżu krytycznych wartości rekombinacji można wytłumaczyć zróżnicowaną zdolnością do akceptowania crossing-over w zależności od miejsca na chromosomie.

Zarówno w chromosomach wirtualnych, jak i w rzeczywistych chromosomach ssaków obserwuje się zwiększone prawdopodobieństwo rekombinacji w rejonach subtelomerowych (rejon z rosnącymi wartościami na wykresach).



Ryc. 7. Różne strategie ewolucji chromosomów w zależności od liczby genów na jednostkę długości (1cM).

Prawa kolumna przedstawia frakcję genów ludzkich w sekwencjach subtelomerowych (ostatnich 10 Mbp od końca chromosomu), które nie posiadają genów ortologicznych w genomie szympansa. Lewa kolumna przedstawia frakcję genów w pozostałych częściach chromosomów ludzkich, które nie posiadają ortologów w genomie szympansa. Liczby wynoszą odpowiednio 403 i 921, a różnice we frakcjach są statystycznie istotne.

nomiennie (Ryc. 5). Jest to wynikiem równowagi dynamicznej między selekcją oczyszczającą, eliminującą defektywne geny a presją mutacyjną wprowadzającą nowe defekty. Sytuacja zmienia się dramatycznie jeżeli rekombinacja zostanie wyłączona. Frakcja defektywnych genów w całej puli genetycznej rośnie wtedy do 0,5. Gdyby przyjąć, że dystrybucja defektywnych genów na chromosomach jest losowa, to przy losowym doborze gamet prawdopodobieństwo utworzenia przeżywającego osobnika byłoby rzędu 10^{-13} , taka populacja musiałaby wymrzeć. W symulacjach komputerowych jednak nie wymiera. Analizując pulę genetyczną tej populacji można zauważyć, że w poszczególnych genomach wszystkie defekty są komplementowane przez dzikie allele oraz, że liczba haplotypów w całej populacji maleje i w skrajnych warunkach w całej puli są tylko dwa różne, komplementujące haplotypy (ZAWIERTA i współaut. 2007, WAGA i współaut. 2007, ZAWIERTA i współaut. 2008). W takiej populacji prawdopodobieństwo utworzenia przeżywającego osobnika wynosi 0,5, a nie 10^{-13} , jak wynikałoby z losowego rozkładu defektów. Tak się dzieje w przypadku, kiedy rekombinacja między chromosomami nie zachodzi. Jeżeli zachodzi jeden crossing-over na gene-

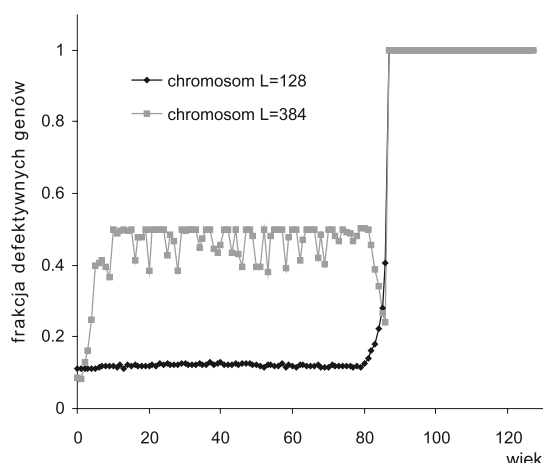
rację, to frakcja defektywnych genów jest niska, a ich rozkład jest losowy. Zmieniając częstość rekombinacji między wartością 0 i 1 można znaleźć taką krytyczną jej wartość, w której populacja „nie może się zdecydować” jaką strategię wybrać – komplementowania haplotypów czy strategię oczyszczającą. Wybór między tymi strategiami ma charakter przejścia fazowego (DE OLIVEIRA i współaut. 2008, ZAWIERTA i współaut. 2008). Różnica w częstości rekombinacji rzędu 0,01 decyduje o przyjęciu jednej z tych strategii. W pobliżu wartości krytycznej rekombinacji tworzy się charakterystyczny rozkład defektywnych genów (Ryc. 5, puste kółka dla rekombinacji 0,15) i częstości akceptowanych rekombinacji – akceptowanych, to znaczy takich crossing-over, które w efekcie dają gametę tworzącą przeżywającą zygotę. Skumulowany, pozbawiony trendu wykres rozkładu częstości tych rekombinacji przedstawiono na Ryc. 6. Regiony, w których wartości rosną odpowiadają tym regionom chromosomów, w których częstość rekombinacji jest wyższa niż średnia dla danego chromosomu. Regiony, w których wartości się zmniejszają, odpowiadają tym częściom chromosomu, w których rekombinacja jest niższa niż średnia. Dla porównania, na tej samej rycinie przedstawiono rozkład rekombinacji na jednym z ludzkich chromosomów. Dla wszystkich innych badanych chromosomów ludzkich uzyskano podobne wykresy. W częściach subtelerowych częstości rekombinacji są wyższe niż w częściach centralnych chromosomów, zarówno w chromosomach rzeczywistych, jak i wirtualnych. Te rozkłady sugerują, że częstości rekombinacji na eukariotycznych chromosomach mogą być bliskie wartości krytycznych.

Sprawdzono więc, jak zmieniają się wartości krytyczne rekombinacji dla chromosomów o liczbie genów odpowiadających chromosomom ludzkim. Symulacje wykazały, że dla chromosomu o długości około 2000 genów, krytyczna wartość rekombinacji wynosi między dwa a trzy crossing-over. Odpowiada to parametrom najdłuższego ludzkiego chromosomu. Dla innych chromosomów wartości te również są bliskie wartościom krytycznym. Rozważając, czy chromosomy ludzkie są bliskie stanu krytycznego należy wziąć również pod uwagę, w jak wielkich efektywnych populacjach ewoluował gatunek. Z analiz wynika, że, aby przyjąć, iż ewolucja chromosomów ludzkich przebiegała w pobliżu warunków krytycznych należy założyć, że efektywna wielkość populacji powinna być rzędu kilkuset osób. Wydaje się to być warunkiem do przyjęcia. Istnieją jednak pewne konsekwencje przyjęcia hipotezy, że chromosomy są bliskie stanu krytycznego, o czym świadczy już rozkład częstości rekombinacji. Konsekwencją tego jest również stabilność informacji genetycznej zawartej w chromosomach. Część centralna chromosomów, zawierająca komplementujące elementy, powinna być bardziej konserwatywna, rejony subtelerowe powinny być bardziej zmienne. Sprawdzono to porównując genomy różnych gatunków z genomem ludzkim. W szczególności, jak często geny ludzkie, w zależności od położenia na chromosomie, tracą ortologów w innych genomach (ortologów, są to geny o tym samym pochodzeniu filogenetycznym, występujące w różnych gatunkach). Okazało się, że najszybciej w czasie filogenezy eliminowane są geny leżące w rejonach subtelerowych (Ryc. 7).

CZY RÓŻNICE W FUNKCJONALNEJ STRUKTURZE CHROMOSOMÓW SĄ ISTOTNE?

Wykazano, że wartość krytyczna rekombinacji zależy nie tylko od genetycznej pojemności chromosomu (liczby genów), ale również od liczby genów przypadających na jednostkę rekombinacji (dane niepublikowane, Mackiewicz i współaut.), a tym chromosomy różnią się znacznie, na przykład ludzki chromosom 18 zawiera średnio 2,5 genu na 1cM, a chromosom 19 – 13,6 genu na 1 cM. W chromosomie 18 prawdopodobieństwo separacji dwóch sąsiednich genów przez rekombinację jest więc wielokrotnie wyższe niż w chromosomie 19. Jeżeli częstość rekombinacji w naszym genomie jest bliska

wartości krytycznej, to możliwe jest, że jeden chromosom jest bliższy selekcji oczyszczającej, a drugi bliższy strategii komplementowania haplotypów. Aby wykazać, że istnienie w jednym genomie chromosomów o tak zróżnicowanych właściwościach genetycznych jest możliwe, przeprowadzono symulacje zgodnie z modelem Penna, gdzie osobniki reprezentowane były przez genomy zawierające dwie pary autosomów o różnej długości – 128 i 384 bitów. Częstość mutacji na bit w obu chromosomach była taka sama i taka sama częstość rekombinacji na chromosom (dłuższy chromosom miał więc trzy razy więcej bi-

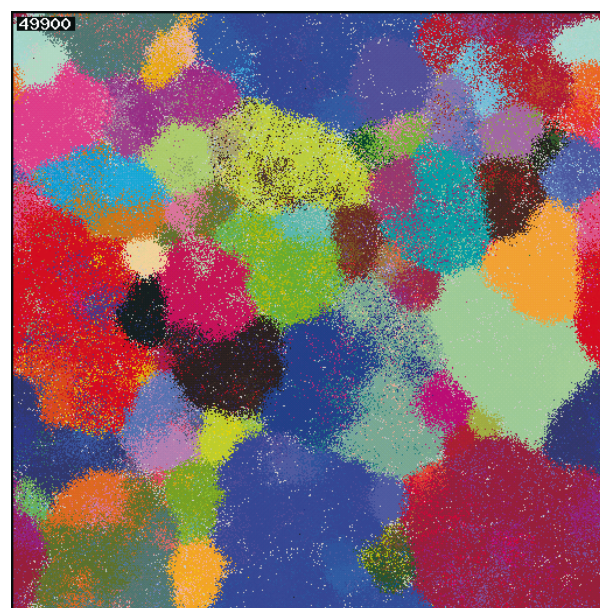


Ryc. 8. Stabilność ortologów w genomie ludzkim w zależności od ich położenia na chromosomie.

Genomy osobników złożone były z dwóch par chromosomów o różnej długości (128 i 384 bity). Częstość mutacji na bit i częstość rekombinacji na chromosom były takie same. Chromosom o dużej gęstości kodowania (liczba bitów na 1 cM) wchodzi w strategię komplementowania, podczas gdy chromosom o małej liczbie bitów na 1cM pozostaje pod selekcją oczyszczającą.

tów na 1 cM, między chromosomami 18 i 19 ta różnica jest niemal sześciokrotna). Wyniki symulacji pokazano na Ryc. 8. Chromosom krótki znajdował się pod selekcją oczyszczającą, natomiast chromosom dłuższy ewoluował zgodnie ze strategią komplementującą.

Należy pamiętać, że między tymi dwoma strategiami jest ostre przejście fazowe, co oznacza, że przekroczenie granicy tego przejścia w obie strony jest bardzo kosztowne z punktu widzenia ewolucyjnego – musi się wiązać ze zmniejszeniem potencjału rozrodczego populacji. Z drugiej strony, częstość rekombinacji związana jest bezpośrednio ze współczynnikiem inbredu – jeżeli osobniki w populacji są blisko ze sobą spokrewnione, to prawdopodobieństwo, że doszło do rekombinacji w obrębie grupy sprzężonych genów dziedziczonych od wspólnego przodka jest niskie. Populacja, której część genomu znajduje pod się wpływem selekcji oczyszczającej, a część pod wpływem strategii komplementującej, znajduje się jakby w dołku potencjału, jeden chromosom „pilnuje”, żeby współczynnik inbredu nie był zbyt wysoki, a drugi, żeby nie był on zbyt niski. Innymi słowami da się to powiedzieć, że ewolucja do-



Ryc. 9. Rozkład osobników na siatce kwadratowej o wymiarach 512×512 po 49900 krokach symulacji.

Partnerka szukała partnera i lokowała dziecko w odległości nie większej niż 4. Kwadraty na siatce kolorowane są według struktury genomu osobnika, który zajmuje dane pole. Osobniki o różnych kolorach mogą się krzyżować, ale prawdopodobieństwo przeżycia potomstwa jest bardzo małe (szczegółowy opis ZAWIERTA i współaut. 2008).

prowadziła do takich częstości rekombinacji między homologicznymi chromosomami, aby potencjał rozrodczy był optymalny przy osiągalnych wielkościach efektywnych populacji. A to znaczy, że sukces reprodukcyjny (również człowieka) nie powinien maleć monotonicznie wraz ze stopniem pokrewieństwa między rodzicami. Powinniśmy raczej obserwować maximum, odpowiadające współczynnikowi inbredu charakterystycznemu dla ewolucji gatunku. Takie właśnie maximum obserwowano analizując dzietność populacji Islandii. Wyniki badań w obrębie innych gatunków również wskazują na to, że przewaga selekcyjna potomstwa maleje wraz z odległością genetyczną rodziców (EDMANDS 2007). Nie da się takich wyników wytłumaczyć stosując analizy i modele średniopoloowe, używane dotychczas do opisu ewolucji populacji.

Opisana ewolucja chromosomów, możliwość komplementowania całych grup genów, a wreszcie niemonotoniczna zależność sukcesu reprodukcyjnego od pokrewieństwa między rodzicami wskazuje na równoczesne występowanie dwóch zjawisk

w populacjach, obniżenia płodności na skutek zarówno zbyt wysokiego, jak i zbyt niskiego współczynnika inbredu (BOŃKOWSKA i współaut. 2007). Jest to związane z nielosowym rozmieszczeniem aktywności genów w komplementujących grupach sprzężeniowych i jest bezpośrednią przyczyną sympatrycznej specjacji. Na Ryc. 9 przedstawiono obraz populacji na siatce kwadratowej o wymiarach 512×512 . Każdy osobnik na siatce zajmuje tylko jeden kwadrat i nadaje temu kwadratowi kolor, który odpowiada strukturze centralnej części jego chromosomu. Osobniki o różnych kolorach mogą się krzyżować, ale z bardzo dużym prawdopodobieństwem umrą przed osiągnięciem wieku rozrodczego, różne kolory odpowiadają,

więc różnym gatunkom. Ta zróżnicowana populacja powstała na skutek wirtualnej ewolucji początkowo homogennej populacji przy ograniczeniu odległości, w której dobierają się partnerzy i gdzie lokują swoje dzieci. Analizując rozmieszczenie gatunków na przedstawionej siatce można stwierdzić, że osobniki z jednego gatunku nie powinny szukać partnerów w zbyt dużej odległości, ponieważ prawdopodobieństwo sukcesu reprodukcyjnego będzie wtedy mniejsze. Ten efekt zanika całkowicie, jeżeli zadeklarujemy w modelu, że populacja jest panmiktyczna. Więcej ilustracji ewolucji populacji można znaleźć na stronie: <http://www.smorfland.uni.wroc.pl/sympatry/>

PODSUMOWANIE

Wyniki komputerowego modelowania ewolucji chromosomów i genomów sugerują inne niż przyjmowane dotychczas scenariusze ewolucji chromosomów płciowych. To przyjęta przez ssaki strategia rozrodcza doprowadziła do utraty genów z chromosomu płciowego Y, co z kolei zmieniło rolę społeczną osobników płci męskiej i różnice w umieralności kobiet i mężczyzn. Wprowadzenie do modeli przestrzennej dystrybucji genów w genomach i osobników w środowisku zastąpiło średniopolewe teorie neo-darwinowskie

i pozwoliło wytłumaczyć zarówno zjawisko specjacji sympatrycznej, jak i obserwowany fakt wzrostu dzietności wraz ze wzrostem pokrewieństwa między rodzicami.

Praca została wykonana w ramach programu: COST Action MP0801 i Katedry Studiów Interdyscyplinarnych UNESCO, Uniwersytetu Wrocławskiego. Symulacje komputerowe były przeprowadzone we Wrocławskim Centrum Sieciowo Superkomputerowym (<http://www.wcss.wroc.pl>), grant#102.

COMPUTER MODELING OF GENOME EVOLUTION

Summary

A dozen years of computer simulations of age structured populations composed of individuals represented by their diploid genomes show how evolution of the genetic pool of populations depends on the population size, intragenomic recombination rate and promiscuity. The cross-over rate and the effective population size decide about the probability of separation of genes located on one chromosome during the reproduction. If this probability is low, the genes are inherited as a cluster. Purifying selection, which tries to minimise the number of mutations by eliminating defective genes from a cluster, seems to be the more costly strategy and genomes may chose the strategy of complementation. Switching between the two strategies – purifying selection and complementation of haplotypes – has a character of transition. Results of the human chromosome analyses suggest that our chromosomes evolve in conditions close to this transition and formation of clusters

and their complementation should be expected. The distribution of genes in the complementing clusters is not random and it is specific for evolving populations. Sympatric speciation, where one species splits into several within the same territory, should be considered as a very common phenomenon in spatially distributed populations and, in fact, it is observed during the computer simulations. In neo-Darwinian theory of evolution, sympatric speciation has been considered as an improbable and negligible phenomenon just because in the mean field models of very large Mendelian populations (panmictic, with very high intragenomic recombination rate) these effects cannot be observed. Computer modeling also showed that the shrinking of the Y chromosome observed during genome evolution of mammals is connected with promiscuity in the strategy of their reproduction.

LITERATURA

- ARBUTHNOT J., 1710. *An argument for Divine Providence, taken from the constant regularity observed in the births of both sexes*. Royal Society, London.
- BIECEK P., CEBRAT S., 2008. *Why Y chromosome is shorter and women live longer?* Europ. Physical J. B. 65, 149–153.
- BOŃKOWSKA K., KULA M., CEBRAT S., STAUFFER D., 2007. *Inbreeding and outbreeding depressions in the Penna model as a result of crossover frequency*. Int. J. Mod. Phys. C 18, 1329–1338.
- de OLIVEIRA P. M. C., MOSS DE OLIVEIRA S., STAUFFER D., CEBRAT S., PEKALSKI A., 2008. *Does Sex Induce a Phase Transition?* Europ. Phys. J. B. 63, 245–254.
- EDMANDS S., 2007. *Between a rock and a hard place: Evaluating the relative risks of inbreeding and outbreeding for conservation and management*. Mol. Ecol. 16, 463–475.
- FERGUSON-SMITH M., 2007. *The Evolution of Sex Chromosomes and Sex Determination in Vertebrates and the Key Role of DMRT1*. Sex Dev. 1, 2–11.
- FISHER, R. A., 1930. *The Genetical Theory of Natural Selection*, Clarendon Press, Oxford
- GOMPERTZ B., 1825. *On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies*. Phil. Trans. Roy. Soc. (London) 115, 513–585.
- HARDY G. H., 1908. *Mendelian proportions in a mixed population*. Science 28, 49–50.
- HAYFLICK L., MOORHEAD P. S., 1961. *The serial cultivation of human diploid cell strains*. Exp. Cell Res. 25, 585–621.
- HELGASON A., PÁLSSON S., GUÐBJARTSSON D. G., KRISTJÁNSSON P., STEFÁNSSON K., 2008. *An association between the kinship and fertility of human couples*. Science 319, 813–816.
- JIGGINS C. D., 2006. *Sympatric speciation: why the controversy?* Curr. Biol. 16, R333–R334.
- LOBO M. P., ONODY R. N., 2005. *Degeneration of the Y chromosome In evolutionary aging models*. Eur. Phys. J. B 45, 533–537.
- ŁASZKIEWICZ A., CEBRAT S., 2005. *Monte Carlo Simulations of the Age Structure of the Human Population*. J. Insur. Med. 37, 3–12.
- MASA M., CEBRAT S., STAUFFER D., 2006. *Does telomere elongation lead to a longer lifespan if cancer is considered?* Physica A 364, 324–330.
- MAYR E., 1942. *Systematics and the Origin of Species*. Columbia Press, New York.
- MEDAWAR P. B., 1952. *An unsolved problem of biology*. H. K. Lewis & Co., London.
- MOSS DE OLIVEIRA S., DE OLIVEIRA P. M. C., STAUFFER D., 1999. *Evolution, Money, War and Computers*. Stuttgart-Leipzig: Teubner.
- NIEWCZAS E., CEBRAT S., STAUFFER D., 2000. *The influence of the Medical Care on the Human Life Expectancy in 20th Century and the Penna Ageing Model*. Theory Biosci. 119, 122–131.
- PENNA T. J. P., 1995. *A bit-string model for biological aging*. J. Stat. Phys. 78, 1629–1633.
- S'A MARTINS J. S., CEBRAT S., 2000. *Random Deaths in a Computational model for age-structured populations*. Theory Biosci. 119, 156–165.
- STAUFFER D., MOSS DE OLIVEIRA S., DE OLIVEIRA P. M. C., S'A MARTINS J.S. 2006. *Biology, Sociology, Geology by Computational Physicists*. Amsterdam: Elsevier
- WAGA W., MACKIEWICZ D., ZAWIERTA M., CEBRAT S., 2007. *Sympatric speciation as intrinsic property of expanding populations*. Theory Biosci. 126, 53–59.
- ZAWIERTA M., BIECEK P., WAGA W., CEBRAT S., 2007. *The role of intragenomic recombination rate in the evolution of population's genetic pool*. Theory Biosci. 125, 123–132.
- ZAWIERTA M., WAGA W., MACKIEWICZ D., BIECEK P., CEBRAT S., 2008. *Phase Transition in Sexual Reproduction and Biological Evolution*. Int. J. Mod. Phys. C 6, 917–936.

SŁOWNIK UŻYWANYCH TERMINÓW GENETYCZNYCH

- Allel – jedna z dwóch albo wielu postaci danego genu odpowiedzialnego za pełnienie konkretnej funkcji, zajmującego zwykle to samo miejsce w genomie (locus).
- Allel dziki – funkcjonalna forma genu, allel funkcjonalny, występujący w naturalnej populacji.
- Autosom – chromosom inny, niż chromosom płciowy.
- centyMorgan (cM) – genetyczna jednostka długości chromosomu – prawdopodobieństwo zajścia crossing over na odcinku jednego centyMorgana, w czasie jednej mejozy wynosi 1%.
- Chromosomy płciowe – chromosomy, których występowanie się różni u obu płci i które są zaangażowane w różnicowanie płciowe (pozostałe chromosomy, to autosomy).
- Crossing-over – genetyczna rekombinacja wzajemna. Wymiana homologicznych fragmentów chromosomów w czasie mejozy, pozwala na pojawienie się chromosomów o strukturze genetycznej różnej, niż chromosomy rodzicielskie.
- Fenotyp – obserwowalna cecha osobnika, będąca wynikiem oddziaływania genotypu ze środowiskiem.
- Filogeneza – droga rozwoju, historia ewolucyjna, pochodzenie grupy organizmów (zwykle gatunku).
- Gameta – dojrzała komórka płciowa, zdolna do fuzji z gametą płci przeciwnej, w wyniku której powstaje zygota.
- Gen – brak dobrej definicji genu. Znaczenie terminu zmienia się zależnie od kontekstu. Tutaj, jednostka genetyczna odpowiedzialna za określoną funkcję.
- Genom – całkowita informacja genetyczna zawarta w organizmie. W przypadku eukariotów, termin jest często stosowany tylko dla pojedynczego zestawu chromosomów.
- Genom diploidalny – informacja genetyczna zawarta w obydwu kompletach chromosomów organizmu (diploidalnego).
- Genom haploidalny – informacja genetyczna zawarta w jednym komplecie chromosomów, gamety, bądź organizmu (haploidalnego).
- Genotyp – (1) – całkowita informacja genetyczna organizmu (odpowiada pojęciu genomu). (2) – genetyczna struktura organizmu pod względem jednego genu lub małej grupy genów.
- Haplotyp – (1) – termin używany zamiennie z genomem haploidalnym. (2) – specyficzny układ

- genów na jednym chromosomie lub w jednej gamecie.
- Hemizygota – organizm posiadający w swoim genomie tylko jedną kopię danego genu, serii genów lub chromosomu (na przykład samiec ssaka posiada tylko jedną kopię chromosomu X w swoim genomie).
- Heterogametyczna (płeć) – osobniki tej płci produkują dwa rodzaje gamet (samiec ssaka produkuje plemniki z chromosomem X albo z chromosomem Y).
- Heterozygota – osobnik, który co najmniej w jednym locus posiada dwa różne allele.
- Homogametyczna (płeć) – płeć, której osobniki produkują tylko jeden rodzaj gamet (wszystkie komórki jajowe samicy ssaka posiadają chromosom X).
- Homozygota – osobnik, który w konkretnym locus posiada identyczne allele, albo osobnik, który we wszystkich loci posiada identyczne allele.
- Locus (lm. loci) – miejsce w genomie lub na chromosomie, zajmowane przez dany gen (allel).
- Mutacja – zmiana w sekwencji nukleotydowej genu, najczęściej powodująca jego uszkodzenie i utratę funkcji.
- Mutacja letalna – mutacja powodująca uszkodzenie genu prowadzące do śmierci organizmu.
- Mutacja recesywna – mutacja, której efekt może być komplementowany (uzupełniany) przez dziki allel (w organizmie diploidalnym).
- Ortolog – gen homologiczny do danego genu, występujący w innym organizmie. Geny ortologiczne pochodzą od jednego genu występującego w organizmie wspólnego przodka obu organizmów.
- Poliandria – związek małżeński jednej kobiety z wieloma mężczyznami.
- Poligamia – związek małżeński, w której może być więcej niż jeden przedstawiciel jednej płci.
- Populacja panmiktyczna – populacja, w której dominującą strategią rozrodczą jest promiskuityzm. W opisywanym przypadku modelu w populacji występuje poliandria.
- Promiskuityzm – strategia rozrodcza, w której każdy osobnik populacji może kopulować z dowolnym innym osobnikiem płci odmiennej tej populacji.
- Pula genetyczna – zespół informacji genetycznej populacji (gatunku).
- Rekombinacja – wymiana, łączenie, wyłączanie, przemieszczanie informacji genetycznej w obrębie jednego genomu lub między genomami.
- Rekombinacja wzajemna – wymiana homologicznych odcinków chromosomów (patrz crossing over).
- Replikacja – podwajanie informacji genetycznej poprzez budowanie dwóch siostrzanych cząsteczek DNA na matrycy cząsteczki macierzystej.
- Rewersja – powrót zmutowanej formy genu do jego pierwotnej postaci na skutek kolejnej mutacji (bardzo rzadkie zjawisko).
- Śmierć genetyczna – śmierć organizmu, wynikająca z jego struktury genetycznej (na przykład na skutek mutacji letalnej).
- Specjacja – pojawienie się nowego gatunku.
- Specjacja sympatryczna – wyłonienie się nowego gatunku w obrębie istniejącej populacji bez występowania barier fizycznych, geograficznych lub ekologicznych.
- Subtelomerowe (rejony) – części chromosomów położone w pobliżu telomerów.
- Telomer – końcowa część chromosomu, zbudowana z wielu krótkich powtórzonych sekwencji nukleotydowych, które zapobiegają zlepianiu się chromosomów końcami.
- Współczynnik inbrodu – wartość określająca średni stopień pokrewieństwa między rodzicami.
- Zygota – komórka (zwykle diploidalna), która powstaje na skutek fuzji gamety żeńskiej i męskiej.