

MATEUSZ STANISŁAW KONCZAL

*Instytut Nauk o Środowisku
Uniwersytet Jagielloński
Gronostajowa 7, 30-387 Kraków
E-mail: mateusz.konczal@uj.edu.pl*

GENOMIKA ADAPTACJI

WSTĘP

Od wielu lat biolodzy z uwagą badają cechy, które powstały pod wpływem doboru naturalnego i są przyczyną przystosowania organizmu do środowiska lub powodują zwiększony sukces reprodukcyjny osobnika. Wciąż jednak dalecy jesteśmy od odpowiedzi na wiele pytań związanych z powstaniem i naturą adaptacji. Obserwowana przez nas zmienność genetyczna w świecie organizmów żywych może wynikać zarówno z procesów deterministycznych (dobór naturalny), jak i stochastycznych (dryf genetyczny), a spór pomiędzy „adapcjonistami” a „neutralistami” pozostaje otwarty (NEI 2005). Fundamentalne pytanie, w jakim stopniu dobór naturalny wpływa na zmienność genetyczną populacji pozostaje póki co bez odpowiedzi (FRASER 2011). W ostatnich kilku latach zbliżamy się jednak szybciej niż kiedykolwiek wcześniej do wyjaśnienia przynajmniej części z tych wątpliwości. Gwałtowny rozwój technik sekwencjonowania pozwala na stosunkowo precyzyjne wnioskowanie o zróżnicowaniu genetycznym oraz o molekularnych podstawach adaptacji u coraz większej liczby gatunków, a całościowa analiza genotypu, fenotypu i dostosowania pozwoliła naukowcom na wysunięcie pierwszych wniosków o molekularnych podstawach adaptacji (STAPLEY i współaut. 2010).

Poniżej omawiam w jaki sposób naukowcy radzą sobie z niełatwą materią badania adaptacji na poziomie molekularnym. Mam nadzieję, że prezentacja różnych strategii badawczych, podparta konkretnymi przykła-

dami badań z ostatnich kilku, lat pozwoli czytelnikowi na zrozumienie podstawowych zagadnień związanych z genomiką adaptacji.

Adaptacja to cecha lub stan cechy, która przyczynia się do lepszego przeżywania i rozrodu w porównaniu z organizmami o innym fenotypie. Powstanie jej wyobrazić sobie można tak: wraz ze zmianą środowiska (np. ociepleniem klimatu), inne niż dotychczas cechy mogą stać się korzystne. Nowopowstały lub rzadki w populacji wariant genu warunkujący tę korzystną cechę zyskuje przewagę selekcyjną, jego częstość wzrasta i pod warunkiem utrzymywania się presji selekcyjnej wystarczająco długo, utrwała się w populacji. Tak powstałą cechę nazywamy adaptacją (FISHER 1930). Nawet intuicyjne w pewnej mierze podejście Fishera wyjaśniające powstanie adaptacji może budzić kontrowersje (HUGHES 2011). Są one uzasadnione o tyle, o ile nie dodamy do definicji aspektu historycznego. Jeżeli obecnie istniejąca w populacji cecha przyczynia się do wyższego dostosowania osobników ją posiadających, może być ona wynikiem zarówno doboru pozytywnego, jak i zaszłości historycznych czy ewolucji neutralnej (FUTUYMA 2008). W kontekście tego artykułu termin adaptacja będzie używany w węższym zakresie (Ramka 1) i określać będzie cechę, która wyewoluowała pod wpływem działania konkretnego czynnika selekcyjnego (HARVEY i PAGEL 1991). W tym znaczeniu, cecha, której utrwalenie w populacji związane jest z ewolucją neutralną, nie będzie nazywana adaptacją, nawet jeżeli

RAMKA 1. PODSTAWOWE POJĘCIA STOSOWANE W GENOMICIE ADAPTACJI

Adaptacja – cecha, która wyewoluowała pod wpływem działania określonego czynnika selekcyjnego.

Dobór pozytywny – dobór zwiększający częstość korzystnego allelu w następnym pokoleniu.

Dostosowanie – średni wkład fenotypu lub genotypu w pulę genową następnych generacji.

Dryf genetyczny – losowe zmiany częstości alleli w populacji.

Epistaza – interakcja między allelami znajdującymi się w różnych loci

Ewolucja eksperymentalna – testowanie hipotez biologii ewolucyjnej poprzez kontrolowane eksperymenty w laboratorium, wykorzystujące zróżnicowaną rozrodczość i przeżywania osobników w populacji pod wpływem ekspozycji na konkretny czynnik.

Ewolucja neutralna – zmiany częstości alleli w kolejnych pokoleniach powodowane przez mutacje (neutralne lub prawie neutralne) i dryf genetyczny.

Fenotyp – zbiór cech osobnika danego gatunku determinowanych przez jego genotyp, środowisko i ich interakcje.

Genomika – dziedzina zajmująca się badaniem genomów, kompleksów genów oraz interakcji pomiędzy genami.

Genotyp – zestaw alleli posiadany przez osobnika w interesujących nas loci.

Geny kandydackie – geny wytypowane do badań nad adaptacją na podstawie wcześniejszej wiedzy dotyczącej funkcji pełnionej przez produkty ich ekspresji

Geny ortologiczne – geny, których rozdzielenie i zróżnicowanie nastąpiło na skutek procesu dziedziczenia

GWAS (ang. Genome Wide Association Studies) – poszukiwanie związków pomiędzy wariantami genetycznymi a interesującymi nas cechami poprzez ogólnogenomowe genotypowanie w populacji osobników, dla których dostępna jest informacja o fenotypie

Indel – krótka delecja, insercja lub ich kombinacja

Konwergencja – niezależna ewolucja podobnych lub takich samych cech rozpoczynająca się z różnych punktów.

Ontologia genów – hierarchiczna baza danych grupująca geny i ich produkty według funkcji pełnionych w organizmie.

Paralelizm – niezależna ewolucja podobnych lub takich samych cech rozpoczynająca się z tego samego punktu.

Proteom – zbiór wszystkich białek ulegających ekspresji w konkretnym czasie i miejscu w organizmie.

Sekwencjonowanie nowej generacji (ang. next-generation sequencing) – technologie pozwalające stosunkowo szybko i tanio uzyskać sekwencje miliardów nukleotydów.

SNiP (ang. Single Nucleotide Polymorphism) – polimorfizm pojedynczego nukleotydu, występuje, gdy w badanej próbce w konkretnej pozycji na chromosomie mamy do czynienia z więcej niż jednym wariantem nukleotydowym.

Transkryptom – zbiór wszystkich cząsteczek RNA, transkrybowanych w konkretnym czasie i miejscu w organizmie.

Wymiatanie selekcyjne – proces zmniejszenia zmienności genetycznej w miejscu genomu sprzężonym z miejscem genomu znajdującym się pod wpływem działania doboru pozytywnego.

Zastana zmienność genetyczna (ang. standing genetic variation) – zmienność genetyczna istniejąca w danym momencie w populacji.

obecnie przyczynia się do wyższego dostosowania osobników w porównaniu z cechą alternatywną.

Ponieważ artykuł ten dotyczyć ma genomiki adaptacji, pozostaje nam jeszcze zdefiniowanie, czym jest genomika. Ta dyscyplina zajmuje się badaniem genomów organizmów. W przeciwieństwie do klasycznej

genetyki traktuje obserwowane w genomie zjawiska całościowo, a wnioski wysuwane są na podstawie często wyrafinowanych analiz bioinformatycznych. W szczególności genomika zajmuje się interakcjami pomiędzy poszczególnymi loci oraz tym, które geny odpowiedzialne są za wykształcenie konkretnych cech. Genomika porównawcza natomiast,

to badania porównawcze całych genomów różnych organizmów. W genomice wnioski możemy formułować na podstawie badań genomu, lub pośrednio, badając transkryptom (MORTAZAVI i współaut. 2008) czy proteom (DIZ i współaut. 2012).

Genomika adaptacji jest więc multidyscyplinarnym projektem badawczym, którego celem jest zidentyfikowanie wzorców genomowych związanych z pojawianiem się adaptacji. Nieco bardziej szczegółowe pytania, jakie stawiają sobie naukowcy zaangażowani w te badania, można sformułować następująco: (i) czy adaptacja jest wynikiem mutacji w kilku loci o dużych efektach czy w wielu genach małych efektów, (ii) czy adaptacja następuje z nowopowstałych mutacji czy z istniejącej w populacji zmienności genetycznej, (iii) czy źródłem adaptacji są mutacje w genach regulatorowych czy w sekwencjach kodujących, (iv) czy adaptacja jest powtarzalna na poziomie genetycznym, (v) jaka część obserwowanej pomiędzy populacjami zmienności jest zmiennością adaptatywną a jaka neutralną, (vi) jak bardzo możliwość zaadaptowania się do nowego środowiska wynika z ogólnogenomowej zmienności oraz poziomu interakcji pomiędzy poszczególnymi genami?

Aby odpowiedzieć na powyższe pytania stosowane są różne strategie badawcze. Część naukowców bada te zagadnienia prowadząc ewolucję eksperymentalną w laboratorium (TENAILLON i współaut. 2012), inni

skupiają się na sytuacjach, w których odseparowane od siebie populacje wielokrotnie zaadaptowały się do podobnego środowiska (TAYLOR i MCPHAIL 2000). Jeszcze inne badania dotyczą zmienności pomiędzy populacjami danego gatunku (człowiek, rzodkiewnik) i związku poszczególnych wariantów nukleotydów w genomie z wykształceniem konkretnych cech fenotypowych (STRANGER i współaut. 2011). Pomimo że ogólnogenomowe badania poszukujące związków konkretnych alleli z konkretnymi cechami (głównie chorobami) (ang. genome wide association studies, GWAS), prowadzone są zwykle dla innych celów niż badania molekularnych podstaw adaptacji, generowane przez nie dane mogą być w równym stopniu użyteczne dla genomiki adaptacji.

Trzy przedstawione powyżej strategie (badania w kontekście ewolucji eksperymentalnej, konwergencji/paralelizmu, GWAS), są dominującymi schematami, w jakich bada się adaptację oraz jej związek z konkretnym regionem genomu. Poniżej dokładniej omawiam każdą z tych strategii, przedstawiając zarówno jej wady, jak i zalety, dzięki czemu czytelnik może w bardziej krytyczny sposób odnieść się do przedstawionych w kolejnych częściach wyników. W dodatkowych ramach tłumaczę podstawowe pojęcia związane z artykułem, przedstawiam sygnatury doboru pozytywnego w genomie oraz testy wykorzystywane do jego wykrywania.

ADAPTACJA W LABORATORIUM

Adaptacja jako cecha, która wyewoluowała pod wpływem działania doboru naturalnego, badana być może zarówno w naturze, jak i w laboratorium. Umożliwia to koncepcja ewolucji eksperymentalnej, gdzie badany w laboratorium gatunek poddajemy działaniu czynnika, do którego przez wiele pokoleń populacja o określonej wielkości się adaptuje. Takie eksperymenty przez setki pokoleń w łatwy sposób można prowadzić na mikroorganizmach takich jak wirusy, bakterie czy grzyby (ROKYTA i współaut. 2005, BARRICK i współaut. 2009, ARAYA i współaut. 2010, HERRING i współaut. 2006, TENAILLON i współaut. 2012). Więcej wysiłku i czasu wymagają natomiast badania na bardziej złożonych organizmach, takich jak muszki owocowe, niecień czy myszy (TEOTÓNIO i współaut. 2009, GARLAND i KELLY 2006, MORRAN i współaut. 2011). W ciągu ostatnich dwudziestu lat wie-

le z takich badań prowadzonych jest na niespotykaną dotąd skalę, a dzięki technikom wielkoskalowego sekwencjonowania możliwe jest wnioskowanie o zmianach genetycznych leżących u podstaw zachodzącej w laboratorium ewolucji (DETTMAN i współaut. 2012).

Jedną z największych zalet takiego podejścia jest potencjalne kontrolowanie niemal wszystkich czynników związanych z ewoluującą populacją. Przede wszystkim osoba prowadząca badania nad ewolucją eksperymentalną sama określa, co jest czynnikiem selekcyjnym i tylko od poprawnego projektu eksperymentu zależy, czy zadany czynnik będzie jedynym aspektem, który różni selekcionowane i kontrole linie. W ewolucji eksperymentalnej ważnym jest również, aby linii eksperymentalnych i kontrolnych było kilka, tak, aby możliwe było badanie powtarzalności

RAMKA 2. SYGNATURY DOBORU POZYTYWNEGO W GENOMIE

Mając do dyspozycji ogólnogenomowe dane możemy poszukiwać w nich śladów działania doboru pozytywnego. Na podstawie takich śladów można przeprowadzić wnioskowanie o adaptacyjnych zmianach, które zaszły w badanym regionie. Jest to możliwe dzięki istnieniu matematycznej teorii ewolucji neutralnej, która w testach na dobór często traktowana jest jak hipoteza zerowa. Generuje ona oczekiwania, co do tego jak powinna wyglądać zmienność genetyczna w loci, w których nie działa dobór naturalny. Odstępstwa od tych predykcji pozwalają nam wnioskować o działaniu w przeszłości doboru pozytywnego.

Pierwszą z takich sygnatur doboru jest obniżona heterozygotyczność. Na skutek wymiatania selekcyjnego (Ryc. 1, 2) w populacji dochodzi do utrwalenia nie tylko selekcjonowanego wariantu, ale także fragmentu genomu będącego w bezpośredniej bliskości z tym wariantem. W porównaniu z innymi regionami genomu obniżona heterozygotyczność może być sygnaturą działania doboru. Może być ona związana zarówno z interesującym nas w tym przypadku doborem pozytywnym jak i z doborem oczyszczającym. Dobór oczyszczający polega na usuwaniu nowopowstałych i szkodliwych mutacji z populacji. Można wyobrazić sobie to w taki sposób, że mamy w populacji utrwalaony fragment genu kodującego białko, każda mutacja powodująca zmianę aminokwasu w sekwencji będzie powodować utratę właściwości białka i będzie usuwana z populacji. W ten sposób zarówno dobór oczyszczający jak i pozytywny będą powodować obniżenie zmienności. Sygnatura ta ma więc ograniczoną interpretację.

Kolejnym śladem działania doboru pozytywnego jest zmieniony kształt rozkładu frekwencji alleli w populacji. Po wymiataniu selekcyjnym i redukcji heterozygotyczności zaczynają pojawiać się nowe neutralne mutacje. Potrzeba dużo czasu, aby dryf genetyczny spowodował wzrost ich częstości. W takim regionie genomu obserwujemy w związku z tym nadmiar rzadkich alleli i niedobór wariantów o pośrednich częstościach w porównaniu z przewidywaniami generowanymi przez teorię neutralną. Drugą zmianą jest nadmiar pochodnych w stosunku do stanu ancestralnego alleli. Pojawiają się one na haplociepie podczas doboru i są one podwożone razem z selekcjonowanym wariantem. Pochodne allele są oczekiwane w genomie średnio w niższej frekwencji w stosunku do ich ancestralnych odpowiedników z powodu ich relatywnie młodsze go wieku. Wymiatanie selekcyjne tworzy sytuację, gdzie w porównaniu z oczekiwaniami, więcej pochodnych alleli ma wyższą częstość w populacji.

Ponieważ w różnych populacjach działają różne czynniki selekcyjne, lokalne adaptacje przyczyniają się także do zwiększenia zróżnicowania pomiędzy populacjami. Jeżeli w danym regionie genomu miało miejsce wymiatanie selekcyjne w dwóch różnych populacjach, to selekcjonowane warianty znajdowały się najprawdopodobniej na różnym tle genetycznym (o ile nie zostały odziedziczone po wspólnym przodku). Dobór pozytywny powoduje więc zarówno zmniejszenie zróżnicowania genetycznego wewnątrz populacji jak i zwiększenie dywergencji pomiędzy populacjami.

Śladem działania doboru pozytywnego są także długie i częste w populacji haplotypy. Jeżeli jakiś wariant genu ewoluuje neutralnie, to czas, w którym stanie się on częsty w populacji, jest długi, dużo dłuższy niż w przypadku doboru pozytywnego. Im dłuższy czas allel jest obecny w populacji tym więcej rekombinacji miało miejsce w regionie genomu związanym z tym wariantem. Natomiast im szybciej dany allel utrwała się w populacji tym dłuższy fragment genomu jest z nim podwożony.

Ostatnia ważna sygnatura działania doboru pozytywnego oparta jest na tempie podstawień synonimowych (ds) i niesynonimowych (dn). Podstawienia synonimowe to takie, które nie zmieniają sekwencji kodowanego białka. W związku z tym nie mają wpływu na fenotyp i powinny ewoluować neutralnie. Jeżeli nowopowstała mutacja zmieniająca strukturę białka ma takie samo prawdopodobieństwo zniknięcia z populacji jak ta, która zaszła w miejscu synonimowym, znaczy to, że mutacja taka jest neutralna, a proces ten jest stochastyczny. Jeżeli natomiast dobór faworyzuje taką mutację, prawdopodobieństwo utrzymania się mutacji w populacji jest większe. Wyższe tempo podstawień niesynonimowych w stosunku do synonimowych w jakimś regionie daje również możliwość wnioskowania o adaptatywnej ewolucji.

zachodzących procesów oraz testowanie hipotez stawianych w konkretnych badaniach. Dodatkowo w badaniach laboratoryjnych w łatwy sposób można określić wielkość populacji oraz częstości alleli w każdym pokoleniu. Genomowe sygnatury doboru (Ramka 2) często są podobne do efektów działania czynników demograficznych, które możemy kontrolować w laboratorium. Dzięki temu wnioskowanie o molekularnych podstawach zachodzących procesów staje się dużo prostsze.

Nie jest to jednak podejście idealne. Poważnymi problemami są dwie kwestie. Po pierwsze, ewolucja eksperymentalna, trwająca wystarczająco długo by pozostawić ślady w genomie, jest wykonalna głównie przy użyciu mikroorganizmów rozmnażających się bezpłciowo. Po drugie, ewolucja w laboratorium nie koniecznie wygląda tak jak ta, która miała miejsce w przyrodzie i prowadziła do wykształcenia obecnie obserwowanych adaptacji.

Klonalność i bardzo duża wielkość populacji mikroorganizmów sprawia, że ich ewolucja może wyglądać zgoła inaczej niż ewolucja organizmów rozmnażających się płciowo i żyjących w mniejszych populacjach. Klonalność sprawia, że brak w takich populacjach rekombinacji, a więc adaptatywne zmiany na poziomie genomu praktycznie muszą następować w sekwencji jedna po drugiej. Wyobrazić to sobie można w taki sposób, że osobnik zyskuje przewagę z powodu konkretnego wariantu genu, więc jego kłony stosunkowo szybko wypierają inne osobniki z populacji. Następnie, wśród dużej liczby klonów, jeden zyskuje przewagę z powodu nowopowstałej korzystnej mutacji i sytuacja się powtarza. U organizmów rozmnażających się płciowo, takie mutacje mogą występować niezależnie, po czym „spotkać się” w jednym osobniku. Przy rozmnażaniu płciowym ważnym zagadnieniem są więc interakcje pomiędzy korzystnymi wariantami różnych genów, natomiast u bakterii czy wirusów problem ten ma zupełnie inny wymiar (MULLER 1932).

Prawdopodobieństwo wystąpienia mutacji na parę zasad na generację nie różni się bardzo pomiędzy różnymi gatunkami organizmów zbudowanych na bazie DNA (KONDRAŠOV i KONDRAŠOV 2010). Różnice w liczbie genów, nawet pomiędzy bakteriami i ssakami, nie są większe niż jeden rząd wielkości. Mnożąc prawdopodobieństwo mutacji razy liczbę miejsc w genomie mających wpływ na fenotyp (fragmenty kodujące i regulatorowe)

razy liczbę osobników otrzymujemy liczbę nowopowstałych do populacji mutacji mających wpływ na dostosowanie. Mając tak uproszczony rachunek łatwo zauważyć, że ogromne populacje bakterii pozwalają na uzyskanie w każdym pokoleniu znacznie większej liczby nowych, wpływających na dostosowanie mutacji niż populacje innych organizmów. Duża liczba nowych mutacji pozwala populacjom mikroorganizmów ewoluować z wykorzystaniem nowopowstałych mutacji łatwiej niż innym organizmom. Oznacza to, że wewnątrzpopulacyjna zmienność genetyczna w świecie mikroorganizmów nie ma tak dużego znaczenia, jak ma to miejsce u bardziej złożonych organizmów, ponieważ w każdym pokoleniu mutacje dostarczają wystarczająco dużo zmienności do populacji.

Przy prowadzeniu ewolucji eksperymentalnej na bardziej złożonych organizmach można zauważyć odwrotny problem, a mianowicie populacje te są zazwyczaj mniejsze od naturalnych. Problemem może więc stać się wzrost wsobności i utrata zmienności w wyniku działania dryfu genetycznego. W końcu wielu autorów sugeruje, że ewolucja eksperymentalna ma niewiele wspólnego z tą zachodzącą w naturze, ponieważ, w przyrodzie nigdy nie działa tylko jeden czynnik selekcyjny, a dodatkowo występują zaniebawane w laboratorium interakcje genotyp-środowisko.

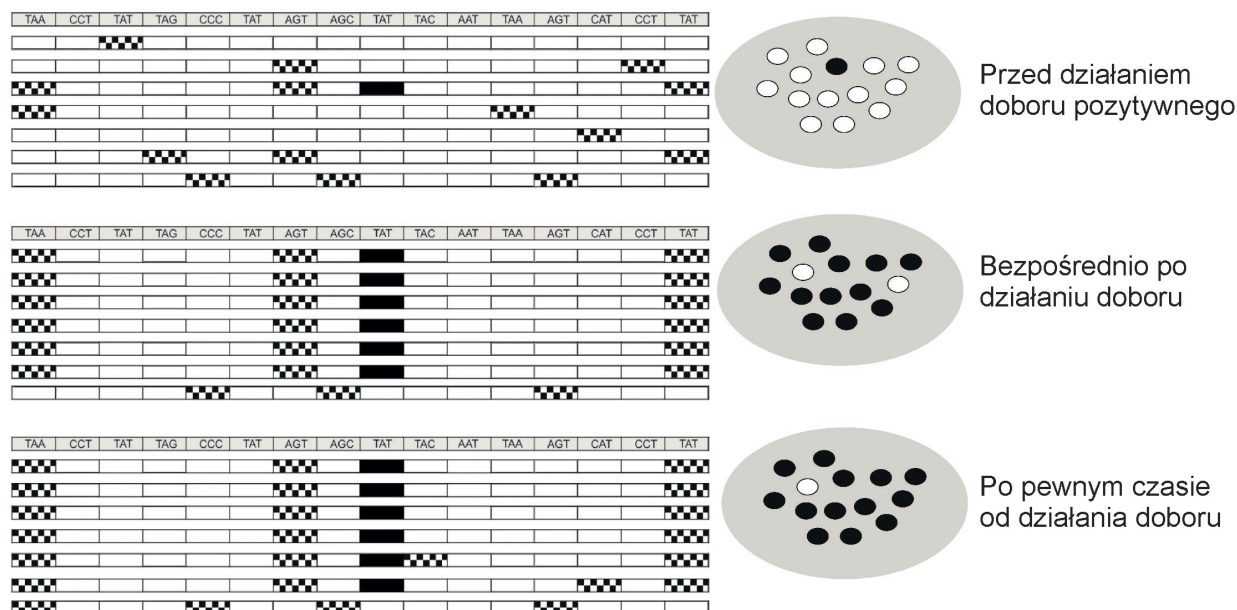
Pomimo ograniczeń badania eksperymentalne prowadzone w laboratoriach związane są obecnie praktycznie ze wszystkimi pytaniami genomiki adaptacji. Wady, które przed chwilą wymieniałem, powodują, że interpretacja tak otrzymanych wyników powinna być ograniczona. Mimo to eksperymenty selekcyjne są jednymi z najciekawszych w kontekście badania molekularnych podstaw adaptacji. Poniżej streszczam rezultaty prac prowadzonych na dwóch modelowych organizmach: bakterii (*Escherichia coli*) oraz na muszkach owocowych (*Drosophila melanogaster*).

Escherichia coli to jeden z najczęściej wybieranych do ewolucji eksperymentalnej organizmów (SHENDURE i współaut. 2005, BARRICK i współaut. 2009, CONRAD i współaut. 2009, LEE i PALSSON 2010, WANG i współaut. 2010, TENAILLON i współaut. 2012). Wynika to z faktu, że gatunek ten od wielu lat hodowany jest w laboratoriach i używany w różnego rodzaju badaniach. Mając na uwadze wymienione wcześniej ograniczenia eksperymentów selekcyjnych warto podkreślić,

że wiele z nich wniosło ciekawe obserwacje dla osób zajmujących się tematyką adaptacji. Na przykład TENAILLON wraz ze współpracownikami (2012) opublikowali niedawno ciekawą pracę, w której testują liczbę potencjalnie korzystnych mutacji pojawiających się w procesie adaptacji bakterii do wysokiej temperatury. Autorzy zauważyli, że wcześniejsze eksperymenty selekcyjne zawierały zazwyczaj tylko kilka równoległych linii, przez co możliwość wnioskowania o powtarzalności ewolucji była ograniczona (TEOTÓNIO i współaut. 2009, LENSKI i współaut. 1991). W ramach swojego eksperymentu autorzy przeprowadzili więc badania na ponad 100 liniach (populacjach) *E. coli* adaptowanych do życia w wysokiej temperaturze. Z oczywistych względów taki eksperyment jest niemal niemożliwy do wykonania dla większych organizmów. Po 2000 pokoleń autorzy zaobserwowali istotnie wyższe dostosowanie organizmów z linii selekcyjowanych, w porównaniu z kontrolnymi, podczas ekspozycji na wysoką temperaturę. Przystąpili więc do dalszych badań w celu określenia jakie zmiany genomowe odpowiedzialne są za zaobserwowane adaptacje. Zidentyfikowali ponad 600 pozycji, w których nowopowstałe mutacje były korzystne. Dodatkowo autorzy oszacowali, że liczba wszystkich potencjalnie korzystnych mutacji jest ogromna i trudna do bezpośredniego zaobserwowania. Natomiast udało im się oszacować zdecydowaną większość genów (loci) zaangażowanych w adaptatywną ewolucję. Elementy te były związane z dwiema alternatywnymi strategiami ewolucyjnymi: jedną związaną z kompleksem polimerazy RNA, a drugą z czynnikiem terminacyjnym rho. Badania te pokazują, że o ile ewolucja zachodzi z nowopowstałych mutacji, potencjał adaptacyjny populacji może być duży, a adaptacja zachodzić może w różny sposób, zmieniając przebieg różnych procesów w organizmie. Jeżeli występuje powtarzalność adaptatywnej ewolucji na poziomie molekularnym, to z przedstawionych badań wynika, że ma ona miejsce raczej na poziomie funkcjonalnych elementów (takich jak szlaki metaboliczne) aniżeli pojedynczych podstawień nukleotydowych.

Jedną z pierwszych prac prowadzonych na bardziej skomplikowanych organizmach, analizującą ogólnogenomową odpowiedź w eksperymencie ewolucyjnym opublikowała na łamach *Nature* BURKE i współaut. (2010). Zbadali oni genetyczne podstawy zmian zachodzących w eksperymencie prowadzonym

na *Drosophila melanogaster*. Wywodzące się z jednej ancestralnej populacji osobniki w 1980 r. zostały podzielone na 5 linii selekcyjowanych na szybszy rozwój osobniczy oraz 5 linii kontrolnych. Po ponad 600 pokoleniach selekcyjowane muszki rozwijały się o ok. 20% szybciej niż te z populacji wyjściowej. Aby zidentyfikować różnice genetyczne na przestrzeni całego genomu, zdecydowano się na wykorzystanie sekwencjonowania nowej generacji i zsekwencjonowanie zmieszanych próbek pochodzących z poszczególnych linii w obrębie tego samego kierunku selekcji. Na podstawie ponad 680 tys. pojedynczych polimorfizmów nukleotydowych (ang. single nucleotide polymorphism, SNP), wytypowano kilkadziesiąt regionów genomu, które różniły się istotnie frekwencją alleli pomiędzy liniami selekcyjowanymi a kontrolnymi. Regiony genomu, w których zaobserwowano różnice w częstości alleli, charakteryzowały się także obniżoną heterozygotycznością, co związane jest z wymiataniem selekcyjnym. Jest to zjawisko opisane przez MAYNARDA-SMITHA i HAIGHA (1974) i nazywane inaczej podwożeniem (ang. hitchhiking). Polega ono na obniżeniu zmienności genetycznej w miejscach, które są fizycznie sprzężone z genem, na który działa dobór pozytywny. Jeżeli konkretna kopia genu zyskuje przewagę selekcyjną i utrwała się w populacji, fragmenty genomu znajdujące się w tym samym rejonie utrwalą (podwoją) się razem z nią, dlatego, że rekombinacja „nie zdąży” zróżnicować tła genetycznego selekcyjowanego genu w tej części genomu (Ryc. 1). Zgodnie z oczekiwaniami, analiza ontologii genów różnicujących selekcyjowane i kontrolne linie wykazała, że regiony te związane były z rozwojem osobniczym. Udało się więc uzyskać geny kandydackie (Ramka 1), które najprawdopodobniej odpowiedzialne były za zaobserwowaną fenotypową adaptację. Dokonano tego bez wcześniejszych założeń (takich jak badanie tylko wybranych na podstawie wcześniejszej wiedzy genów, mogących mieć znaczenie w adaptacji), dzięki czemu udało się uniknąć stronniczości i pominięcia niektórych potencjalnie ważnych fragmentów genomu. Zaskakujący okazał się natomiast fakt, że selekcja nie spowodowała utrwalenia potencjalnie korzystnych alleli w ewoluujących niezależnie populacjach (liniach). Pomimo silnego kierunkowego doboru trwającego przez kilkaset pokoleń nie zaobserwowano klasycznego obrazu zaniku zmienności w regionie związanym z adapta-



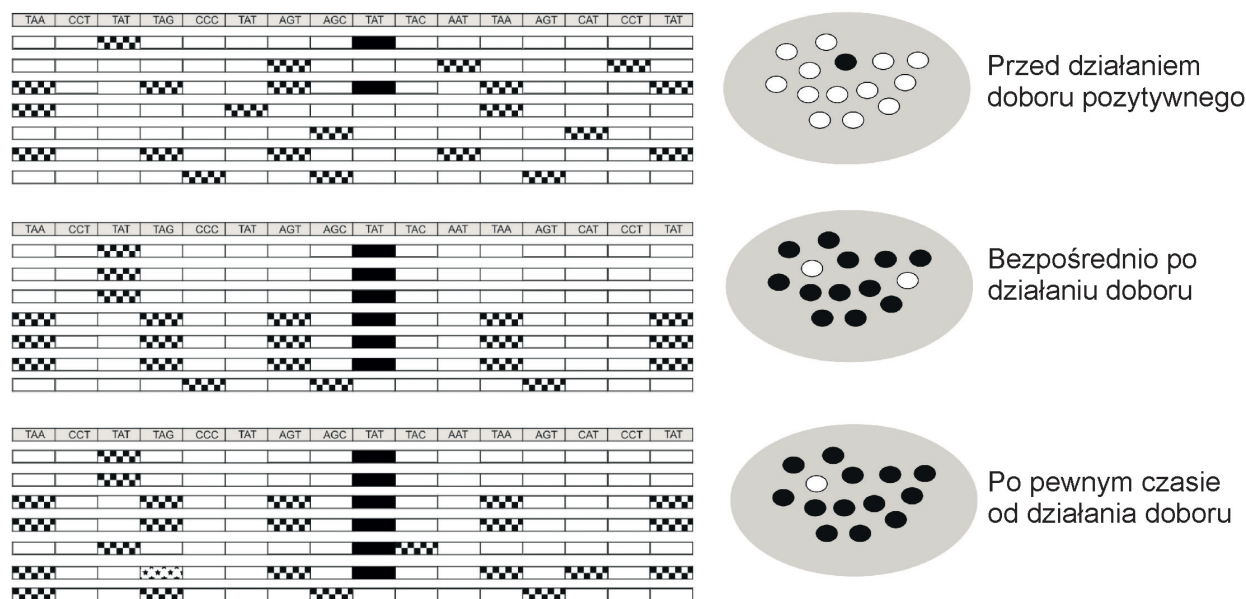
Ryc. 1. Efekt działania doboru pozytywnego na zmienność w populacji. Po lewej stronie przedstawione są fragmenty chromosomów obecne w populacji. Białe komórki oznaczają taki sam kodon, jak ten w sekwencji referencyjnej, a zacieniowane, to takie, w których kodon jest różny. Taki sam deseń oznacza taką samą różnicę. Na czarno zaznaczony jest różny od referencyjnego kodon, powodujący korzystną zmianę w fenotypie. Po prawej przedstawiona jest zmienność fenotypowa w populacji. Czarnym kolorem zaznaczony jest selekcjonowany fenotyp. Sytuacja przedstawia efekt klasycznego wymiatania selekcyjnego i zanik heterozygotyczności w loci sprzężonych z selekcjonowanym wariantem.

tywnym genem. Taki zanik zmienności znany jest powszechnie z populacji rozmnażających się bezpłciowo mikroorganizmów. Przeprowadzone badania pokazują, że wykrywanie adaptacji na poziomie molekularnym u rozmnażających się płciowo organizmów jest trudniejsze niż u bakterii. Ewolucja adaptacyjna z zastanej zmienności genetycznej, dobór zmienny w czasie, antagonistyczne działanie różnych genów mogą spowodować, że nie będziemy obserwować zupełnego zaniku heterozygotyczności w regionach związanych z genem znajdującym się pod wpływem działania doboru pozytywnego. W przypadku rozmnażania płciowego i działania doboru na złożone cechy warunkowane przez wiele genów, proces adaptacji może wpływać na zmienność genetyczną w inny sposób niż ma to miejsce u bakterii.

Jednym z możliwych wyjaśnień wyników uzyskanych przez Burke jest model łagodnego wymiatania (ang. soft sweep) (HERMISSON i PENNING 2005). Opisuje on dwa scenariusze, z których każdy jest dużo trudniejszy do wykrycia na poziomie molekularnym, w porównaniu z klasycznym modelem wymiatania selekcyjnego. Pierwszy zakłada, że faworyzo-

wanych przez dobór jest wiele niezależnych mutacji pojawiających się w tym samym genie. Dobór pozytywny działa do momentu aż suma frekwencji korzystnych wariantów w populacji wyniesie 1. Ponieważ mutacje pojawiały się niezależnie, występują na różnych łach genetycznych. Zmienność genetyczna w tym rejonie zostanie obniżona, ale żaden z wariantów się nie utrwali. Drugi scenariusz zakłada adaptacyjną ewolucję z zastanej zmienności genetycznej. Obecny, ale rzadki w populacji allel zyskuje przewagę i utrwała się. Ponieważ wcześniej wariant ten był przez wiele pokoleń obecny w populacji, na skutek mutacji oraz procesów rekombinacji występuje on na różnych łach genetycznych. W tym wypadku również nie zaobserwujemy zupełnego zaniku zmienności genetycznej w rejonie selekcjonowanego allelu (Ryc. 2) (KARASOV i współaut. 2010).

Zaprezentowane prace to tylko przykłady tego, jak próbuje się odpowiadać na przedstawione pytania przy użyciu koncepcji ewolucji eksperymentalnej. Podobne projekty badawcze prowadzone na mikroorganizmach podsumowane zostały w pracy DETTMANA i współaut. (2012). Prace te dotyczyły m. in.



Ryc. 2. Efekt działania doboru pozytywnego na zmienność w populacji. Po lewej stronie przedstawione są fragmenty chromosomów obecne w populacji. Białe komórki oznaczają taki sam kodon, jak ten w sekwencji referencyjnej, a zaciemnione, to takie, w których kodon jest różny. Taki sam desień oznacza taką samą różnicę. Na czarno zaznaczony jest różny od referencyjnego kodon, powodujący korzystną zmianę w fenotypie. Po prawej przedstawiona jest zmienność fenotypowa w populacji. Czarnym kolorem zaznaczony jest selekcionowany fenotyp. Sytuacja przedstawia efekt łagodnego wymiatania selekcyjnego i ewolucję z zastanej zmienności genetycznej. Korzystna mutacja w tym przypadku obecna jest na dwóch różnych tłach genetycznych.

epistazy; pokazano, że mutacje w konkretnych genach mogą mieć pozytywny wpływ na dostosowanie tylko pod warunkiem występowania konkretnego allelu w innym genie (LEE i PALSSON 2010), a prezentowane na tle genetycznym populacji ancestralnej mogą mieć nawet negatywny wpływ na dostosowanie (CONRAD i współaut. 2009). Inne badania pokazują, że więcej mutacji oraz mutacje w genach tzw. dużych efektów pojawiają się na wczesnym etapie procesu adaptacji (BARRICK i współaut. 2009, CONRAD i współaut. 2009). Większość badań na muszkach owocowych skupiała się natomiast na poszukiwaniu rozkładu wielkości efektów genów warunkujących cechy ilościowe, takie jak masa

ciała (TURNER i współaut. 2011), długość życia (WILSON i współaut. 2006), kształt skrzydeł (MEZEY i HOULE 2005) czy odporność na głód (HARBISON i współaut. 2005). Cytowane prace potwierdzają wcześniej zaproponowany przez ROBERTSONA (1967) wykładniczy rozkład wielkości efektów, z kilkoma loci dużych lub średnich efektów oraz rosnącą liczbą loci malejących efektów. Jest to ważne również w kontekście genomiki adaptacji, gdyż bardzo często to cechy ilościowe znajdują się pod silnym działaniem doboru pozytywnego (ELLEGREN i SHELDON 2008). Dobór ten w różny sposób będzie wpływał na genom w zależności od rozkładu efektów genów warunkujących ewoluującą cechę.

BADANIA NAD PARALELIZMEM I KONWERCENCJĄ W NATURALNYCH POPULACJACH

Alternatywnym podejściem dla prac laboratoryjnych są badania naturalnych populacji (NADEAU i JIGGINS 2010). Jeżeli interesuje nas architektura genetyczna ważnych ekologicznie i ewolucyjnie cech (ORR 1998)

oraz zależności, jakie występują pomiędzy tą architekturą a procesem adaptacji (MERILÄ i SHELDON 1999), to oczywiście najlepszymi miejscami do badania tych zależności są naturalne populacje. Jednym ze stosowanych od

wielu lat podejść jest badanie niezależnych od siebie populacji przystosowujących się do podobnego środowiska i wykształcających podobne cechy (ELLEGREN i SHELDON 2008, ELMER i MEYER 2011).

Procesy, które dla niezależnych populacji doprowadziły do powstania podobnych lub takich samych cech nazywamy konwergencją i paralelizmem. Granica pomiędzy tymi dwoma pojęciami jest nieostra i bywa interpretowana w różny sposób. Przyjmijmy, że paralelizm to niezależna ewolucja podobnych cech, rozpoczynająca się z tego samego punktu początkowego, a konwergencja to powtarzalna ewolucja rozpoczynająca się niezależnie z różnych punktów. Badania naturalnych populacji są o tyle problematyczne, że zazwyczaj trudno oddzielić od siebie procesy, które związane były z adaptacją, od tych determinowanych przez zaszłości historyczne i ewolucję neutralną (FUTUYMA 2008). Wykształcenie takich samych cech pod wpływem działania podobnego czynnika selekcyjnego (w takich samych warunkach środowiska) w niezależnych populacjach skłania do stwierdzenia, że cechy te są adaptatywne (LOSOS 2011). Prawdopodobieństwo takiego zjawiska obserwowanego na skutek procesów stochastycznych (ewolucja neutralna) jest bardzo niskie.

Zasada ta nie jest jednak uniwersalna. Cechy, które nie mają wpływu na dostosowanie w nowym środowisku mogą ewoluować powtarzalnie, zwłaszcza, jeżeli ewolucja ta związana jest z utratą wcześniej pełnionej funkcji. Taka sytuacja ma miejsce na przykład u ryb żyjących w jaskiniach, gdzie neutralne mutacje doprowadziły do utraty pigmentacji ciała niezależnie w wielu populacjach (JEFFERY 2009). Jeżeli cecha jest wielogenowa, to prawdopodobieństwo pojawienia się mutacji, która spowoduje utratę tej cechy jest duże, a więc prawdopodobieństwo, utrwalenia się takiej neutralnej mutacji także jest duże [zgodnie z neutralną teorią ewolucji molekularnej KIMURAY (1983)].

Neutralna teoria ewolucji molekularnej Kimury jest niezwykle ważna w poszukiwaniu śladów działania doboru pozytywnego, gdyż prawie zawsze stanowi hipotezę zerową w przeprowadzanych analizach. Stwierdza ona, że niewiele mutacji jest korzystnych i zostaje utrwalona na skutek doboru, przeciwnie, wiele mutacji jest szkodliwych i usuwanych jest z populacji w ramach działania doboru oczyszczającego. Większość mutacji utrwalanych jest natomiast na skutek loso-

wych zmian częstości alleli w populacji (dryf genetyczny). W związku z tym większość zmienności genetycznej obecnej w populacji jest neutralna i obojętna dla procesu adaptacji. Zmienność w populacji jest również przejściowa, nowopowstały allel znajduje się na drodze albo ku utrwaleniu albo ku zniknięciu z populacji. Motoo Kimura przyjmując te założenia zbudował niezwykle wyrafinowane modele matematyczne opisujące ewolucję na poziomie molekularnym. Jednym z ważniejszych wyników jest stwierdzenie, że liczba mutacji, które powstają w każdym pokoleniu i pewnego dnia ulegną utrwaleniu jest równa tempu mutacji i nie zależy od wielkości populacji. Ten i inne wyniki wykorzystywane są w testach na obecność doboru pozytywnego w genomie (Ramka 3).

Wracając do badań adaptacji w modelach związanych z paralelizmem pierwszym i najważniejszym ograniczeniem jest skupienie się tylko na tych cechach, co do których nie ma wątpliwości, że związane są z adaptatywną ewolucją do konkretnego środowiska. Jest to podejście oparte na badaniu populacji pod kątem pewnej wybranej cechy tak, aby upewnić się, że cecha ta ma dodatni wpływ na dostosowanie. Następnie na podstawie istniejącej wiedzy próbuje się wytypować geny, które mogą być odpowiedzialne za wykształcenie tej cechy, a następnie na poziomie molekularnym bada się właśnie te geny. Może to doprowadzić do pominięcia ważnych genów i procesów zaangażowanych w adaptację oraz ograniczyć postrzeganie adaptacji tylko do cech, które łatwo zauważyć i zmierzyć na poziomie fenotypu (OLEKSYK i współaut. 2010, BARRETT i HOEKSTRA 2011).

Dużą część rewolucji, którą przyniosło ze sobą wykorzystanie sekwencjonowania nowej generacji w badaniach adaptacji związana jest z rozwiązaniem wyżej wymienionych problemów (STAPLEY i współaut. 2010). Po pierwsze, nie trzeba już skupiać się na typowaniu genów, które prawdopodobnie odpowiedzialne są za adaptacje, można zsekwencjonować całe transkryptomy, a nawet genomy organizmów i szukać tych fragmentów genomu bez wcześniejszych założeń. Dzięki śladom pozostawionym przez dobór na poziomie genetycznym (Ramka 2) jesteśmy w stanie za pomocą testów statystycznych lokalizować geny, na które działał dobór pozytywny (Ramka 3). W taki sposób zdolni jesteśmy do wykrycia nie tylko łatwo zauważalnych adaptacji, jak na przykład wielkość ciała czy pigmentacja, ale także dużo bar-

RAMKA 3. TESTY NA OBECNOŚĆ DOBORU POZYTYWNEGO W GENOMIE

Statystyka D Tajimy – populacyjne tempo mutacji można oszacować za pomocą dwóch estymatorów. Pierwszy z nich to różnorodność nukleotydów, czyli średnia liczba różnic na pozycję nukleotydową między parami sekwencji. Drugi to proporcja miejsc zmiennych, czyli udział miejsc polimorficznych w porównywanych sekwencjach. Oszacowane na podstawie tych wartości szacunkowych populacyjnego tempa mutacji powinny być sobie równe w przypadku neutralności (Tajima 1989). Jeżeli wyższa jest wartość szacunkowa oparta na proporcji miejsc zmiennych, znaczy to, że mamy do czynienia z nadmiarem rzadkich alleli, który jest jedną z sygnatur działania doboru pozytywnego.

Statystyka F_{ST} – koncepcyjnie jest to spadek heterozygotyczności w wyniku podziału na subpopulacje (Weir i Cockerham 1984). F_{ST} można obliczyć według wzoru $(H_T - H_S)/H_T$, gdzie H_T to oczekiwana heterozygotyczność w całej populacji obliczona z frekwencji alleli (dla dwóch alleli: $2pq$), a H_S to średnia oczekiwana heterozygotyczność dla subpopulacji. W przypadku neutralności F_{ST} zależy od dryfu genetycznego i migracji, ale lokalne adaptacje mogą przyspieszyć różnicowanie się populacji w loci znajdujących się pod wpływem działania doboru pozytywnego i tym samym zwiększyć F_{ST} w miejscu związanym z adaptacją.

Statystyka H Faya i Wu – statystyka wykrywająca nadmiar pochodnych w stosunku do stanu ancestralnego alleli o wysokiej częstości w populacji, co zazwyczaj związane jest z działaniem doboru pozytywnego (Fay i Wu 2000).

Test dn/ds – test wykonywany dla sekwencji kodujących; jest to stosunek podstawień niesynonimowych na miejsce niesynonimowe do podstawień synonimowych na miejsce synonimowe (Nei i Gojobori 1986). Gdy jako wynik testu otrzymamy $dn/ds > 1$ możemy wnioskować o działaniu doboru pozytywnego.

Test Hudsona-Kreitmana-Aguade – wykorzystuje tylko te allele, które są allelami pochodnymi w stosunku do stanu ancestralnego i porównuje proporcję utrwalonych i segregujących wariantów pomiędzy dwoma regionami genomu (Hudson i współaut. 1987). W badaniach ogólnogenomowych można zastosować to podejście testując kolejne fragmenty genomu i porównywać je ze średnią dla całego genomu lub chromosomu (Begun i współaut. 2007). Regiony, w których jest istotnie więcej utrwalonych fragmentów mogą być związane z działaniem doboru pozytywnego. Efekt ten związany jest z działaniem wymiatania selekcyjnego, gdzie podczas doboru pozytywnego utrwalany jest nie tylko allel znajdujący się pod doбором, ale także fragmenty sekwencji znajdujące się z nim w fizycznym sprzężeniu.

Test McDonalda Kreitmana – test wykorzystywany do wykrywania doboru pozytywnego pomiędzy dwoma gatunkami. Wykorzystuje utrwalone różnice pomiędzy gatunkami (D), oraz te, które są polimorficzne (P) (McDonald i Kreitman 1991). Dodatkowo do testu dane dzielone są na różnice synonimowe (D_s , P_s) i niesynonimowe (D_n , P_n). W przypadku neutralności $P_n/P_s = D_n/D_s$, natomiast dobór pozytywny powoduje zwiększenie dywergencji niesynonimowej, czego efektem jest $D_n/D_s > P_n/P_s$.

Test oparty na strukturze haplotypów – test ten polega na analizowaniu długości haplotypów w genomie. W pierwszym etapie identyfikujemy w interesującym nas locus (rdzeń haplotypu). Następnie szacuje się sprzężenie pomiędzy konkretnymi rdzeniami haplotypów a loci coraz bardziej oddalonymi od tego miejsca na chromosomie. W ten sposób szacuje się wydłużoną homozygotyczność haplotypów (ang. Extended Haplotype Homozygosity, EHH). Rdzenie haplotypów, które mają wysokie wartości EHH oraz są częste w populacji opisują te mutacje, których częstość wzrosła szybciej niż wynikałoby to z teorii neutralnej (Sabeti i współaut. 2002).

Dokładne omówienie sposobu testowania neutralności na poziomie genomowym można znaleźć w Biswas i Akey 2006, Nielsen i współaut. 2007.

dziej subtelnych zmian (Nielsen i współaut. 2007). Możemy więc badać niezależne, ale żyjące w podobnym środowisku populacje bez wcześniejszych założeń i hipotez

(ang. discovery mode of research) (Korona 2009).

Niemniej jednak również to podejście ma swoje ograniczenia. Pomimo tego, że adapta-

cja mogła zajść w podobny sposób u dwóch siostrzanych populacji, różna demografia może zatrzeć obraz molekularnych zmian spowodowanych takimi procesami. Innym problemem jest to, że nie jesteśmy w stanie dowiedzieć się, jak proces adaptacji wpływał na wzorce genomowe w kolejnych pokoleniach, zazwyczaj obserwujemy tylko efekt końcowy zaistniałych procesów.

Pomimo tych wątpliwości, wiele prac badających molekularne podstawy adaptacji w kontekście konwergencji i paralelizmu przyniosło interesujące odkrycia, chociaż jak zostało to przedstawione wcześniej, należy je interpretować w nieco inny sposób niż wyniki badań laboratoryjnych.

W tym kontekście na szeroką skalę prowadzone są badania nad kryptycznym ubarwieniem u gryzoni w Ameryce Północnej. Dowodzą one, że podobne fenotypowo, adaptatywne zmiany mogą zachodzić na skutek mutacji w różnych genach, nawet u blisko ze sobą spokrewnionych gatunków myszaków (LINNEN i współaut. 2009). Z drugiej strony, można zauważyć konwergencję spowodowaną dokładnie taką samą mutacją w genach ortologicznych (MANCEAU i współaut. 2010). Tego typu wyniki jeszcze raz potwierdzają brak prostego schematu adaptatywnej ewolucji, nawet w blisko ze sobą spokrewnionych liniach.

Modelowym gatunkiem w badaniach paralelizmu jest ciernik (*Gasterosteus aculeatus*). Był to jeden z pierwszych gatunków ryb, dla którego udało się zsekwencjonować cały genom, co bez wątpienia ułatwiło badania nad molekularnym podłożem adaptacji do życia w słodkiej wodzie (http://www.ensembl.org/Gasterosteus_aculeatus/Info/Index) (JONES i współaut. 2012). Badania nad ewolucją słodkowodnych populacji dają klarowny obraz tego, jak wyglądały zmiany demograficzne i struktura populacji u tego gatunku (FOSTER i BAKER 2004). Po ustąpieniu lądolodu cierniki skolonizowały słodkie wody Ameryki Północnej i Europy. Od wielu lat stawiana była hipoteza, że ewolucja ryb do słodkich wód następowała niezależnie (TAYLOR i MCPHAIL 2000). Ilościowe potwierdzenie tej hipotezy udało się uzyskać dzięki użyciu narzędzi związanych z genomiką (HOHENLOHE i współaut. 2010). Cierniki nazywane są w taki właśnie sposób, ponieważ osobniki tego gatunku posiadają kolce. Dodatkowo, cechuje je posiadanie wzdłuż boków szeregu płytek kostnych. Te właśnie cechy uległy redukcji lub całkowitemu zanikowi w wielu słodkowodnych populacjach, co związane

było z kilkoma czynnikami, takimi jak niewielka liczba drapieżników, ograniczony dostęp do wapnia, drapieżnictwo ze strony owadów (BELL i współaut. 1993), szybszy wzrost czy wcześniejsza reprodukcja (BARRETT i współaut. 2008).

Ostatnie badania słodkowodnych populacji ciernika wskazały kluczową w redukcji kolców rolę delekcji występującej w elemencie *cis*, regulującym ekspresję genu *Pitx1* (CHAN i współaut. 2010). Co ciekawe, delekcje te miały różne rozmiary, w zależności od tego, z jakiej populacji pochodziły osobniki, potwierdzając przypuszczenia o niezależnej ewolucji cechy w różnych populacjach. Autorzy pracy zbadali także ślady doboru pozytywnego w regionie odpowiedzialnym za redukcję omawianej cechy. Wykorzystując statystykę H (Ramka 3) zaobserwowali oni nadmiar alleli, w których zaszła mutacja w regionie związanym z genem *Pitx1*, oraz obniżoną w tym regionie heterozygotyczność (Ramka 2). Te efekty nie mogły być spowodowane przez efekt wąskiego gardła mającego miejsce podczas kolonizacji potoków, ponieważ heterozygotyczność była istotnie niższa u słodkowodnych populacji ze zredukowanymi kolcami, w porównaniu z podobnymi populacjami, w których ta cecha była obecna. Dodatkowo ogólnogenomowe badania HOHENLOHE i współaut. (2010) również wskazały na gen *Pitx1*, jako na jedno z miejsc w genomie, które związane są z adaptatywną ewolucją do słodkowodnego środowiska. Genomowe badania zwróciły uwagę także na kilka innych wcześniej niebadanych regionów genomu, które również wykazują ślady działania doboru pozytywnego.

Badania nad ciernikiem pokazują, że adaptacja do nowego środowiska nie zawsze związana jest z mutacjami we fragmentach kodujących. Specyficzne dla tkanki wyciszenie homeoboksu *Pitx1* jest doskonałym przykładem przewidywalnej mutacji w genie regulatorowym, powodującej adaptację w konkretnym środowisku. Badania genomu ciernika pozwoliły wysunąć wnioski, że zdecydowana większość adaptatywnych zmian u tego gatunku była wynikiem mutacji w genach regulatorowych (JONES i współaut. 2012). Również badania prowadzone na rybach z rodziny łososiowatych wskazują na istotną rolę powtarzalnych zmian ekspresji genów w procesie adaptacji (DEROME i BERNATCHEZ 2006). Niewiadomą jednak pozostaje, czy zmiany te związane są głównie z utratą pełnionej funkcji lub redukcją jakiejś cechy (STERN i ORGOGOZO 2008).

OGÓLNOGENOMOWE BADANIA ASOCJACYJNE

Kolejnym podejściem do badania molekularnych podstaw adaptacji w naturalnych populacjach jest wykorzystanie danych generowanych przez ogólnogenomowe badania asocjacyjne. Są to prace, oparte na poszukiwaniu korelacji, pomiędzy genotypem a fenotypem. Dla organizmów, dla których znany jest stan danej cechy (np. choroba, wzrost) sprawdza się genotyp w milionach polimorficznych miejsc nukleotydowych. Jeżeli widoczna jest korelacja pomiędzy genotypem w konkretnym miejscu w genomie a fenotypem badanych organizmów, można wnioskować o istnieniu związku pomiędzy allelem a stanem badanej cechy. Znalezienie markerów genetycznych predysponujących osobniki do zapadnięcia na konkretną chorobę było głównym czynnikiem determinującym rozpoczęcie takich badań (GIBSON 2012). Obecnie badania tego typu dotyczą nie tylko chorób, ale także wielu innych, złożonych i odziedziczalnych cech (np. wzrost) oraz prowadzone są nie tylko na populacji ludzkiej, ale także na innych organizmach (np. *Arabidopsis thaliana*) (FLINT i MACKAY 2009, CAO i współaut. 2011). GWAS umożliwiają także poszukiwanie ogólnogenomowych śladów presji selekcyjnej, co może pomóc odpowiedzieć na pytania o zależność pomiędzy cechami genomu i procesem adaptacji (OLEKSYK i współaut. 2010).

Pierwszym poważnym problemem, który napotkały ogólnogenomowe badania asocjacyjne u ludzi było to, że pomimo dużej próby osobników i skanowania całego genomu, w wielu przypadkach nie udało się wyjaśnić więcej niż połowy zmienności genetycznej oszacowanej na podstawie badań odziedziczalności (MANOLIO i współaut. 2009). Debatą na temat „brakującej odziedziczalności” pozostaje otwarta, a dwa główne modele próbujące wyjaśnić ten fenomen mówią albo o rzadkich wariantach genów dużych efektów, które to warianty nie są wykrywane przez badania prowadzone na tak dużą skalę, albo o dużej liczbie częstych wariantów bardzo małych efektów (GIBSON 2012). Wyjaśnienie tego sporu ma o tyle ważne dla genomiki adaptacji konsekwencje, że obserwując dobór na złożoną cechę fenotypową, będziemy mogli przewidzieć ile genów jakich efektów mogło znajdować się pod presją selekcyjną. Posiadając taki model precyzyjniej niż dzisiaj będzie można poszukiwać genów mogących

ewoluować pod wpływem działania konkretnego czynnika selekcyjnego.

Niemniej jednak projekty asocjacyjne oraz ogólnogenomowe skany w poszukiwaniu śladów doboru pokazują pewne ciekawe zależności, jakie zachodzą na poziomie genetycznym podczas procesu adaptacji. Jeden z takich programów prowadzony jest na rzodkiewniku pospolitym (ATWELL i współaut. 2010, CAO i współaut. 2011). W 2010 r. opublikowano badania, w których przeanalizowano 107 różnych fenotypów tego gatunku (ATWELL i współaut. 2010). W zamyśle tak szerokie badania mają dać możliwość zidentyfikowania genów czy nawet konkretnych alleli odpowiadających za zmienność fenotypową. Przebadane przez autorów fenotypy związane były z kwitnieniem, rozwojem osobniczym, odpornością na patogeny i koncentracją pierwiastków w organizmie. Dzięki zaprojektowanej mikromacierzy SNIpów autorzy odnaleźli wiele częstych alleli dużych efektów warunkujących badane fenotypy. Problemem okazała się natomiast struktura populacji, która uniemożliwiła oddzielenie prawdziwych asocjacji od efektu izolowanych populacji.

Z tym wyzwaniem postanowili zmierzyć się inni badacze, którzy od 2008 r. prowadzą projekt mający na celu zidentyfikowanie zmienności genetycznej odpowiadającej za adaptację w 1001 szczepach rzodkiewnika pochodzących z różnych populacji. W 2011 r. ukazała się praca podsumowująca pierwszą część tego projektu (CAO i współaut. 2011). Przeanalizowano w niej genomy 80 linii rzodkiewnika pochodzących z 8 regionów Europy i Środkowej Azji. Zidentyfikowano niemal 5 mln polimorfizmów pojedynczych podstawień (SNIpów) i ponad 800 tys. krótkich indeli. Ta ogromna zmienność pokazała, że wcześniejsze analizy ograniczające się do pewnej części genomu (np. mikromacierze, czyli genotypowanie milionów wcześniej wyselekcjonowanych SNIpów) pokazują tylko niewielką część funkcjonalnie ważnych polimorfizmów. W przebadanych szczepach rzodkiewnika scharakteryzowano jeden region, gdzie oprócz obniżonej różnorodności haplotypów, zgodnie z przewidywaniami dla niedawnego silnego wymiatania, znaleziono także nadmiar rzadkich alleli (Ramka 2). Był to region związany z genem *FRIGIDA*, głównym genem odpowiadającym za regulację czasu kwitnienia. Gen ten już wcześniej był

przedstawiany jako ewoluujący pod wpływem doboru pozytywnego (TOOMAJIAN i współaut. 2006), niemniej jednak nie został wykryty w badaniach GWAS opartych na macierzy SNIpów (ATWELL i współaut. 2010).

Badania Atwell i współpracowników, jak i Cao i współpracowników pokazują, że struktura populacji może utrudnić wnioskowanie o działaniu ewolucji adaptatywnej na podstawie danych molekularnych. Obie prace potwierdzają natomiast, że u rzodkiewnika pospolitego za zmienność fenotypową odpowiadają geny dużych efektów, które nie są jednak w stanie wyjaśnić całej szacowanej na podstawie badań odziedziczalności zmienności genetycznej.

Podobne badania prowadzi się także od kilku lat na ludziach. Z wielu względów najciekawsze dla nas samych jest poznanie historii i ewolucji własnego gatunku. Obecnie na świecie żyje ponad 7 miliardów ludzi, różnych od siebie pod wieloma względami. Naukowcy próbują odpowiedzieć na pytanie, jaka część, z obserwowanej zmienności pomiędzy populacjami jest wynikiem adaptacji, a co pozostaje zaszłością związaną ze stochastycznymi zmianami w izolowanych grupach. Do niedawna prace takie prowadzone były na pojedynczych genach (geny kandydackie). Wybierano loci, które na podstawie dotychczas posiadanej wiedzy mogły być związane z obserwowanymi w środowisku przystosowaniami i o których sądzono, że znajdowały się pod presją doboru pozytywnego. Dzięki tym badaniom udało się wykryć ślady działania doboru pozytywnego na gen *LCT*, związany z tolerancją laktozy u dorosłych osobników (BERSAGLIERI i współaut. 2004), czy geny związane z odpornością na malarię (HAMBLIN i DI RIENZO 2000, HEDRICK 2011). Niemniej jednak badania przy użyciu genów kandydackich mają wiele przedstawionych już wcześniej wad. Dlatego też od kilku lat prowadzone są badania mające na celu odnalezienia ogólnogenomowych śladów doboru pozytywnego. Podsumowania kilku lat takich badań dokonał AKEY (2009). W dwudziestu jeden pracach, które przeanalizował, zidentyfikowano 5110 odległych regionów, które wykazywały ślady działania doboru pozytywnego. Sumarycznie, miejsca te związane były z 23% wszystkich ludzkich genów. Ten obraz budzi jednak pewne zastrzeżenia. Tylko 722 regiony zidentyfikowane zostały w więcej niż jednej pracy, a zaledwie 129 w więcej niż 3. Nie jest to zaskakujące, gdyż autorzy używali różnych testów statystycznych do wykrywa-

nia doboru pozytywnego oraz sami sugerowali, że takie analizy generują dużo fałszywie negatywnych i fałszywie pozytywnych wyników. Niemniej jednak zaskakujący jest i tak duży udział genów związanych z ewolucją adaptatywną, który zauważalny jest także u innych gatunków (FAY i współaut. 2002). Ogólnogenomowe skany wykazały nie tylko wcześniej badane geny związane z adaptacją, ale także nowe loci, które niekoniecznie mogłyby być silnymi kandydatami *a priori* do badania doboru pozytywnego. Udało się wysunąć też nieco bardziej ogólne wnioski, np. takie, że znaczna jest rola doboru pozytywnego działającego na procesy metaboliczne, co sugeruje ich ważną rolę w niedawnej historii ewolucyjnej naszego gatunku. Ponadto, dane sugerują, że najważniejszym czynnikiem selekcyjnym w niedawnej historii były patogeny (FUMAGALLI i współaut. 2011). Średnio w rejonach związanych z adaptacją zaobserwowano natomiast rzadkie występowanie oznak klasycznego wymiatania selekcyjnego kosztem prawdopodobnie modelu łagodnego wymiatania (HERNANDEZ i współaut. 2011). Badania wskazują również na istotną rolę działania doboru oczyszczającego (LOHMUELLER i współaut. 2011).

Reasumując, technologie sekwencjonowania nowej generacji umożliwiły wnioskowanie o adaptacji na niespotykaną dotąd skalę. Możliwe stało się sekwencjonowanie całych genomów, genotypowanie tysięcy loci u tysięcy osobników czy wykorzystanie w badaniach genetycznych organizmów modelowych dla ekologii. Nie jest to rewolucja intelektualna, ale rewolucja technologiczna, porównywalna do tej, jaką przyniosło wykorzystywanie reakcji łańcuchowej polimerazy pod koniec lat 80. XX w. Póki co nie udało się jednak znaleźć wyczerpujących odpowiedzi na postawione w pierwszej części artykułu pytania. Próby zrozumienia molekularnych i ogólnogenomowych podstaw adaptacji podejmowane są poprzez trzy wyżej opisane strategie. Żadne z tych podejść nie jest wolne od założeń i ograniczeń, a rozwój w genomice adaptacji może się odbyć tylko poprzez równoległe kontynuowanie badań we wszystkich trzech kierunkach.

Jednym z celów osób zajmujących się genomiką adaptacji było zweryfikowanie hipotezy o istnieniu przewidywalnych wzorców genetycznych związanych z fenotypową ewolucją (STERN i ORGOGOZO 2009). Po przeanalizowaniu dotychczasowych wyników badań nie mamy podstaw, aby sądzić, że powta-

rzalność zmian na poziomie pojedynczych nukleotydów jest powszechna. Da się zauważyć natomiast pewne wzorce, które pomagają nam zrozumieć naturę adaptacji. U złożonych organizmów bardzo często następuje ona z zastanej zmienności genetycznej. Ma to ważne implikacje dla ochrony gatunkowej, ponieważ zachowanie zmienności genetycznej populacji zwiększy prawdopodobieństwo zaadaptowania się jej do zmieniającego się środowiska. Często adaptatywne zmiany zachodzą w regionach regulatorowych zwłaszcza, jeżeli związane są z utratą jakiejś funkcji lub cechy. Coraz więcej danych wskazuje na to, że spora część obserwowanej zmienności genetycznej jest skutkiem działania doboru pozytywnego. Natomiast adaptacja związana jest z kilkoma mutacjami dużych i średnich efektów pojawiających się w pierwszym etapie oraz coraz większą liczbą zmian zmniejszających się efektów pojawiających się później. Część z tych wyników nie jest zaskakujących inne wręcz przeciwnie (np. duża liczba śladów doboru pozytywnego w geno-

mie), wszystkie natomiast są empirycznymi wynikami, testującymi stawiane dotychczas hipotezy.

Większość, o ile nie wszystkie z tych odpowiedzi możliwe był dzięki zidentyfikowaniu alleli wpływających na fenotyp. Obecnie liczba zidentyfikowanych funkcjonalnie ważnych nukleotydów jest całkiem pokaźna, choć duża część zmienności genetycznej pozostała niewyjaśniona. Brak odpowiedzi na część pytań niektórzy naukowcy wiążą z tym problemem. Główna debata wokół genomiki adaptacji toczy się obecnie wokół pytania czy identyfikacja kolejnych alleli wpływających na fenotyp pozwoli te wątpliwości rozwiązać (ROCKMAN 2012). Bez wątpienia w przeciągu kilku najbliższych lat będziemy świadkami odkryć umożliwiających wyjaśnienie całej zmienności genetycznej albo wieszczanej przez niektórych naukowców klęski związanej z identyfikacją większości ważnych fenotypowo mutacji. Dopiero ten krok pozwoli w pełni wnioskować o architekturze genetycznej dostosowania.

GENOMIKA ADAPTACJI

Streszczenie

Genomika adaptacji to multidyscyplinarny projekt badawczy, którego celem jest zidentyfikowanie wzorców genomowych związanych z procesem adaptacji. W poniższym artykule staram się zapoznać czytelników z podstawowymi pojęciami związanymi z zagadnieniem genomiki adaptacji. Badania w tym obszarze można podzielić na trzy grupy: laboratoryjne eksperymenty naśladujące naturalne procesy ewolucyjne, badania nad konwergencją i paralelizmem oraz badania prowadzone na całych genomach poszukujące związku alleli z wybranymi cechami. Różne podejścia badawcze wiążą się z koniecznością

zróżnicowanej interpretacji uzyskiwanych wyników. W poniższym artykule omawiam szczegółowo te trzy podejścia oraz podaję przykłady prac związanych z konkretną strategią badawczą. Następnie uogólniam uzyskane wyniki, czego wynikiem jest obraz adaptacji jako skomplikowanego na poziomie molekularnym procesu. Powtarzalność adaptacji można zaobserwować na poziomie szlaków metabolicznych; liczba fragmentów genomu ewoluująca pod wpływem działania doboru pozytywnego jest większa niż sądzono wcześniej; niestety inne problemy pozostają jednak wciąż bez satysfakcjonującej odpowiedzi.

ADAPTATION GENOMICS

Summary

Adaptation genomics is a multidisciplinary research program, which aims at detecting genomic patterns resulting from the process of adaptation. Studies in this field can be divided into 3 groups: laboratory experiments that model evolutionary processes, studies on convergence and parallelism and genome wide association studies. Results from particular scientific approaches should be interpreted in different ways. The three approaches accompa-

nied by examples are discussed in detail and then I try to draw some generalizations, which indicate that adaptation is a complex process, also at the molecular level. Repeatability of adaptation is evident at the level of metabolic pathways and large number of genes in genome evolve under positive selection. Unfortunately other problems remain unresolved. Basic terminology related to the field of adaptation genomics is introduced.

LITERATURA

- AKEY J. M., 2009. *Constructing genomic maps of positive selection in humans: Where do we go from here?* Genome Res. 19, 711–722.
- ARAYA C. L., PAYEN C., DUNHAM M. J., FIELDS S., 2010. *Whole-genome sequencing of a laboratory-evolved yeast strain.* BMC Genomics 11.
- ATWELL S., HUANG Y. S., VILHJÁLMSSON B. J., WILLEMS G., HORTON M., LI Y., MENG D., PLATT A., TARONE A. M., HU T. T., JIANG R., MULIYATI N. W., ZHANG X., AMER M. A., BAXTER I., BRACHI B. i współaut., 2010. *Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines.* Nature 465, 627–631.
- BARRETT R. D. H., HOEKSTRA H. E., 2011. *Molecular spandrels: tests of adaptation at the genetic level.* Nat. Rev. Genet. 12, 767–780.
- BARRETT R. D. H., ROGERS S. M., SCHLUTER D., 2008. *Natural selection on a major armor gene in threespine stickleback.* Science 322, 255–257.
- BARRICK J. E., YU D. S., YOON S. H., JEONG H., OH T. K., SCHNEIDER D., LENSKE R. E., KIM J. F., 2009. *Genome evolution and adaptation in a long-term experiment with Escherichia coli.* Nature 461, 1243–1247.
- BEGUN D. J., HOLLOWAY A. K., STEVENS K., HILLIER L. W., POH Y. P., HAHN M. W., NISTA P. M., JONES C. D., KERN A. D., DEWEY C. N., PACHTER L., MYERS E., LANGLEY C. H., 2007. *Population genomics: whole-genome analysis of polymorphism and divergence in Drosophila simulans.* PLoS Biology 5.
- BELL M. A., ORTI G., WALKER J. A., KOENINGS J. P., 1993. *Evolution of Pelvic Reduction in Threespine Stickleback Fish – a Test of Competing Hypotheses.* Evolution 47, 906–914.
- BERSAGLIERI T., SABETI P. C., PATTERSON N., VANDER-PLOEG T., SCHAFFNER S. F., DRAKE J. A., RHODES M., REICH D. E., HIRSCHHORN J. N., 2004. *Genetic signatures of strong recent positive selection at the lactase gene.* Am. J. Human Genet. 74, 1111–1120.
- BISWAS S., AKEY J. M., 2006. *Genomic insights into positive selection.* Trends Genet. 22, 437–446.
- BURKE M. K., DUNHAM J. P., SHAHRESTANI P., THORNTON K. R., ROSE M. R., LONG A. D., 2010. *Genome-wide analysis of a long-term evolution experiment with Drosophila.* Nature 467, 587–590.
- CAO J., SCHNEEBERGER K., OSSOWSKI S., GUENTHER T., BENDER S., FITZ J., KOENIG D., LANZ C., STEGLE O., LIPPERT C., WANG X., OTT F., MUELLER J., ALONSO-BLANCO C., BORGWARDT K., SCHMID K. J., WEIGEL D., 2011. *Whole-genome sequencing of multiple Arabidopsis thaliana populations.* Nat. Genet. 43, 956–963.
- CHAN Y. F., MARKS M. E., JONES F. C., VILLARREAL G. JR., SHAPIRO M. D., BRADY S. D., SOUTHWICK A. M., ABSHER D. M., GRIMWOOD J., SCHMUTZ J., MYERS R. M., PETROV D., JONSSON B., SCHLUTER D., BELL M. A., KINGSLEY D. M., 2010. *Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer.* Science 327, 302–305.
- CONRAD T. M., JOYCE A. R., APPLEBEE M. K., BARRETT C. L., XIE B., GAO Y., PALSSON B. T., 2009. *Whole-genome resequencing of Escherichia coli K-12 MG1655 undergoing short-term laboratory evolution in lactate minimal media reveals flexible selection of adaptive mutations.* Genome Biol. 10.
- DEROME N., BERNATCHEZ L., 2006. *The transcriptomics of ecological convergence between 2 limnetic coregonine fishes salmonidae.* Mol. Biol. Evol. 23, 2370–2378.
- DETTMAN J. R., RODRIGUE N., MELNYK A. H., WONG A., BAILEY S. F., KASSEN R., 2012. *Evolutionary insight from whole-genome sequencing of experimentally evolved microbes.* Mol. Ecol. 21, 2058–2077.
- DIZ A. P., MARTÍNEZ-FERNÁNDEZ M., ROLÁN-ALVAREZ E., 2012. *Proteomics in evolutionary ecology: Linking the genotype with the phenotype.* Mol. Ecol. 21, 1060–1080.
- ELLEGREN H., SHELDON B. C., 2008. *Genetic basis of fitness differences in natural populations.* Nature 452, 169–175.
- ELMER K. R., MEYER A., 2011. *Adaptation in the age of ecological genomics: Insights from parallelism and convergence.* Trends Ecol. Evol. 26, 298–306.
- FAY J. C., WU C. I., 2000. *Hitchhiking under positive Darwinian selection.* Genetics 155, 1405–1413.
- FAY J. C., WYCKOFF G. J., WU C. I., 2002. *Testing the neutral theory of molecular evolution with genomic data from Drosophila.* Nature 415, 1024–1026.
- FISHER R. A., 1930. *Genetical theory of natural selection.* Oxford University Press, Oxford.
- FLINT J., MACKAY T. F. C., 2009. *Genetic architecture of quantitative traits in mice, flies, and humans.* Genome Res. 19, 723–733.
- FOSTER S. A., BAKER J. A., 2004. *Evolution in parallel: new insights from a classic system.* Trends Ecol. Evol. 19, 456–459.
- FRASER H. B., 2011. *Genome-wide approaches to the study of adaptive gene expression evolution: Systematic studies of evolutionary adaptations involving gene expression will allow many fundamental questions in evolutionary biology to be addressed.* BioEssays 33, 469–477.
- FUMAGALLI M., SIRONI M., POZZOLI U., FERRER-ADMETTLA A., PATTINI L., NIELSEN R., 2011. *Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution.* PLoS Genetics 7.
- FUTUYMA D. J., 2008. *Ewolucja.* Wydawnictwo Uniwersytetu Warszawskiego, Warszawa.
- GARLAND T. JR., KELLY S. A., 2006. *Phenotypic plasticity and experimental evolution.* J. Exp. Biol. 209, 2344–2361.
- GIBSON G., 2012. *Rare and common variants: Twenty arguments.* Nat. Rev. Genet. 13, 135–145.
- HAMBLIN M. T., DI RIENZO A., 2000. *Detection of the signature of natural selection in humans: Evidence from the Duffy blood group locus.* Am. J. Human Genet. 66, 1669–1679.
- HARBISON S. T., CHANG S., KAMDAR K. P., MACKAY T. F., 2005. *Quantitative genomics of starvation stress resistance in Drosophila.* Genome Biology 6.
- HARVEY P. H., PAGEL M. D., 1991. *The comparative method in evolutionary biology.* Oxford University Press, Oxford.
- HEDRICK P. W., 2011. *Population genetics of malaria resistance in humans.* Heredity 107, 283–304.
- HERMISSE J., PENNINGS P. S., 2005. *Soft sweeps: Molecular population genetics of adaptation from standing genetic variation.* Genetics 169, 2335–2352.
- HERNANDEZ R. D., KELLEY J. L., ELYASHIV E., MELTON S. C., AUTON A., MCVEAN G., SELLA G., PRZEWORSKI M., 2011. *Classic selective sweeps were rare in recent human evolution.* Science 331, 920–924.
- HERRING C. D., RAGHUNATHAN A., HONISCH C., PATEL T., APPLEBEE M. K., JOYCE A. R., ALBERT T. J., BLATTNER F. R., VAN DEN BOOM D., CANTOR C. R., PALSSON B. O., 2006. *Comparative genome se-*

- quencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nature Genet.* 38, 1406–1412.
- HOHENLOHE P. A., BASSHAM S., ETTER P. D., STIFFER N., JOHNSON E. A., 2010. *Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags.* *PLoS Genetics* 6.
- HUDSON R. R., KREITMAN M., AGUADÉ M., 1987. *A test of neutral molecular evolution based on nucleotide data.* *Genetics* 116, 153–159.
- HUGHES A. L., 2011. *Evolution of adaptive phenotypic traits without positive Darwinian selection.* *Heredity* 108, 347–353.
- JEFFERY W. R., 2009. *Regressive evolution in *astyanax cavefish*.* *Ann. Rev. Genet.* 43, 25–47.
- JONES F. C., GRABHERR M. G., CHAN Y. F., RUSSELL P., MAUCELI E., JOHNSON J., SWOFFORD R., KINGSLEY D. M., 2012. *The genomic basis of adaptive evolution in threespine sticklebacks.* *Nature* 484, 55–61.
- KARASOV T., MESSER P. W., PETROV D. A., 2010. *Evidence that Adaptation in *Drosophila* Is Not Limited by Mutation at Single Sites.* *PLoS Genetics* 6.
- KIMURA M., 1983. *The neutral theory of molecular evolution.* Cambridge University Press, New York.
- KONDRASHOV F. A., KONDRASHOV A. S., 2010. *Measurements of spontaneous rates of mutations in the recent past and the near future.* *Philosoph. Transact. Royal Soc. B: Biol. Sci.* 365, 1169–1176.
- KORONA R., 2009. *Granice adaptacji i konwergencji.* *Kosmos* 58, 395–402.
- LEE D. H., PALSSON B. O., 2010. *Adaptive evolution of *Escherichia coli* K-12 MG1655 during growth on a nonnative carbon source, L-2-propanediol.* *Appl. Environ. Microbiol.* 76, 4158–4168.
- LENSKI R. E., ROSE M. R., SIMPSON S. C., TADLER S. C., 1991. *Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2000 generations.* *Am. Natural* 138, 1315–1341.
- LINNEA C. R., KINGSLEY E. P., JENSEN J. D., HOEKSTRA H. E., 2009. *On the origin and spread of an adaptive allele in deer mice.* *Science* 325, 1095–1098.
- LOHMEIJER K. E., ALBRECHTSEN A., LI Y., KIM S. Y., KORNELIUSSEN T., VINCKENBOSCH N., TIAN G., HUERTA-SANCHEZ E., FEDER A. F., GRARUP N., JØRGENSEN T., JIANG T., WITTE D. R., SANDBÆK A., HELLMANN I., LAURITZEN T., HANSEN T., PEDERSEN O., WANG J., NIELSEN R., 2011. *Natural selection affects multiple aspects of genetic variation at putatively neutral sites across the human genome.* *PLoS Genetics* 7.
- LOSOS J. B., 2011. *Convergence, adaptation, and constraint.* *Evolution* 65, 1827–1840.
- MANCEAU M., DOMINGUES V. S., LINNEA C. R., ROSENBLUM E. B., HOEKSTRA H. E., 2010. *Convergence in pigmentation at multiple levels: Mutations, genes and function.* *Philosoph. Transact. Royal Soc. B: Biol. Sci.* 365, 2439–2450.
- MANOLIO T. A., COLLINS F. S., COX N. J., GOLDSTEIN D. B., HINDORFF L. A., HUNTER D. J., MCCARTHY M. I., RAMOS E. M., CARDON L. R., CHAKRAVARTI A., CHO J. H., GUTTMACHER A. E., KONG A., KRUGLYAK L., MARDIS E., ROTIMI C. N., SLATKIN M., VALLÉ D., WHITTEMORE A. S., BOEHNKE M., CLARK A. G., EICHLER E. E., GIBSON G., HAINES J. L., MACKAY T. F. C., MCCARROLL S. A., VISSCHER P. M., 2009. *Finding the missing heritability of complex diseases.* *Nature* 461, 747–753.
- MAYNARD SMITH J., HAIGH J., 1974. *The hitch-hiking effect of favourable gene.* *Genet. Res.* 23, 23–35.
- MCDONALD J. H., KREITMAN M., 1991. *Adaptive protein evolution at the *Adh* locus in *Drosophila*.* *Nature* 351, 652–654.
- MERILÄ J., SHELDON B. C., 1999. *Genetic architecture of fitness and nonfitness traits: Empirical patterns and development of ideas.* *Heredity* 83, 103–109.
- MEZEY J. G., HOULE D., 2005. *The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*.* *Evolution* 59, 1027–1038.
- MORRAN L. T., SCHMIDT O. G., GELARDEN I. A., PARRISH II R. C., LIVELY C. M., 2011. *Running with the Red Queen: Host-parasite coevolution selects for biparental sex.* *Science* 333, 216–218.
- MORTAZAVI A., WILLIAMS B. A., MCCUE K., SCHAEFFER L., WOLD B., 2008. *Mapping and quantifying mammalian transcriptomes by RNA-Seq.* *Nature Methods* 5, 621–628.
- MULLER H. J., 1932. *Some genetic aspects of sex.* *Am. Natural.* 66, 118–138.
- NADEAU N. J., JIGGINS C. D., 2010. *A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations.* *Trends Genet.* 26, 484–492.
- NEI M., 2005. *Selectionism and neutralism in molecular evolution.* *Mol. Biol. Evol.* 22, 2318–2342.
- NEI M., GOJOBORI T., 1986. *Simple methods for estimating the numbers of synonymous and non-synonymous nucleotide substitutions.* *Mol. Biol. Evol.* 3, 418–426.
- NIELSEN R., HELLMANN I., HUBISZ M., BUSTAMANTE C., CLARK A. G., 2007. *Recent and ongoing selection in the human genome.* *Nature Rev. Genet.* 8, 857–868.
- OLEKSYK T. K., SMITH M. W., O'BRIEN S. J., 2010. *Genome-wide scans for footprints of natural selection.* *Philosoph. Transact. Royal Soc. B: Biol. Sci.* 365, 185–205.
- ORR H. A., 1998. *The population genetics of adaptation: The distribution of factors fixed during adaptive evolution.* *Evolution* 52, 935–949.
- ROBERTSON A., 1967. *The nature of quantitative genetic variation.* [W:] *Heritage from Mendel.* BRINK R. A., STYLES E. D., University of Wisconsin Press, Madison, 265–280.
- ROCKMAN M. V., 2012. *The QTN program and the alleles that matter for evolution: All that's gold does not glitter.* *Evolution* 66, 1–17.
- ROKYTA D. R., JOYCE P., CAUDLE S. B., WICHMAN H. A., 2005. *An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus.* *Nature Genet.* 37, 441–444.
- SABETI P. C., REICH D. E., HIGGINS J. M., LEVINE H. Z. P., RICHTER D. J., SCHAFNER S. F., GABRIEL S. B., PLATKO J. V., PATTERSON N. J., MCDONALD G. J., ACKERMAN H. C., CAMPBELL S. J., ALTSHULER D., COOPER R., KWIATKOWSKI D., WARD R., LANDER E. S., 2002. *Detecting recent positive selection in the human genome from haplotype structure.* *Nature* 419, 832–837.
- SHENDURE J., PORRECA G. J., REPPAS N. B., LIN X., MCCUTCHEON J. P., ROSENBAUM A. M., WANG M. D., ZHANG K., MITRA R. D., CHURCH G. M., 2005. *Molecular biology: Accurate multiplex polony sequencing of an evolved bacterial genome.* *Science* 309, 1728–1732.
- STAPLEY J., REGER J., FEULNER P. G. D., SMADJA C., GALINDO J., EKBLOM R., BENNISON C., BALL A. D., BECKERMAN A. P., SLATE J., 2010. *Adaptation genomics: The next generation.* *Trends Ecol. Evol.* 25, 705–712.
- STERN D. L., ORGOGOZO V., 2008. *The loci of evolution: How predictable is genetic evolution?* *Evolution* 62, 2155–2177.

- STERN D. L., ORGOGOZO V., 2009. *Is genetic evolution predictable?* Science 323(5915): 746–751.
- STRANGER B. E., STAHL E. A., RAJ T., 2011. *Progress and promise of genome-wide association studies for human complex trait genetics.* Genetics 187, 367–383.
- TAJIMA F., 1989. *Statistical method for testing the neutral mutation hypothesis by DNA polymorphism.* Genetics 123, 585–595.
- TAYLOR E. B., MCPHAIL J. D., 2000. *Historical contingency and ecological determinism interact to prime speciation in sticklebacks, Gasterosteus.* Proc Royal Soc London, Ser. B: Biol. Sci. 267, 2375–2384.
- TENAILLON O., RODRÍGUEZ-VERDUGO A., GAUT R. L., McDONALD P., BENNETT A. F., LONG A. D., GAUT B. S., 2012. *The molecular diversity of adaptive convergence.* Science 335, 457–461.
- TEOTÓNIO H., CHELO I. M., BRADIĆ M., ROSE M. R., LONG A. D., 2009. *Experimental evolution reveals natural selection on standing genetic variation.* Nature Genetics 41, 251–257.
- TOOMAJIAN C., HU T. T., ARANZANA M. J., LISTER C., TANG C., ZHENG H., ZHAO K., CALABRESE P., DEAN C., NORDBOG M., 2006. *A nonparametric test reveals selection for rapid flowering in the Arabidopsis genome.* PLoS Biology 4, 732–738.
- TURNER T. L., STEWART A. D., FIELDS A. T., RICE W. R., TARONE A. M., 2011. *Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in Drosophila melanogaster.* PLoS Genetics 7.
- WANG L., SPIRA B., ZHOU Z., FENG L., MAHARJAN R. P., LI X., LI F., MCKENZIE C., REEVES P. R., FERENCI T., 2010. *Divergence involving global regulatory gene mutations in an Escherichia coli population evolving under phosphate limitation.* Genome Biol. Evol. 2, 478–487.
- WEIR B. S., COCKERHAM C. C., 1984. *Estimating F-statistics for the analysis of population structure.* Evolution 38, 1358–1370.
- WILSON R. H., MORGAN T. J., MACKAY T. F. C., 2006. *High-resolution mapping of quantitative trait loci affecting increased life span in Drosophila melanogaster.* Genetics 173, 1455–1463.