

Barkoding DNA – nowe narzędzie do opisu bioróżnorodności

Zbigniew MIREK, Wojciech BIENIEK, Agnieszka SZTORC

MIREK Z., BIENIEK W., SZTORC A. 2007. **DNA barcoding – a new tool for description of biodiversity.** *Wiadomości Botaniczne* 51(3/4): 41–50.

DNA barcoding is currently developing useful tool for species delimitation, description and identification. It uses a short gene sequence from a standardized position in the genome to identify species. Since DNA barcoding relies on DNA, identifying a specimen requires only a small amount of tissue and it may be carried on by non-specialists. So far DNA barcodes based on cytochrome c oxidase 1 (*cox1*) sequence have been applied to a diverse groups of animals. Additional sequence (ITS) have been proposed for fungi. On the other hand there is still not agreed DNA barcode standard for plants.

The main international initiative created to develop and promote DNA barcoding is Consortium for the Barcode of Life (CBOL). It coordinates activities of museums, herbaria, zoos, conservation research institutes, government agencies and other organizations dealing with biodiversity from all over the world. In 2005 the network of four institutions has been created in order to initiate DNA barcoding in Poland. W. Szafer Institute of Botany PAS in Krakow is here responsible for plant and fungi DNA barcoding.

KEY WORDS: biodiversity, DNA barcoding, new tool, systematics, taxonomy

Zbigniew Mirek, Wojciech Bieniek, Agnieszka Sztorc, Instytut Botaniki im. W. Szafera, Polska Akademia Nauk, ul. Lubicz 46, 31-512 Kraków, e-mail: ibmirek@ib-pan.krakow.pl

WSTĘP

Kody paskowe (ang. *bar codes*) obecne są już niemalże wszędzie. W sklepach pozwalają na szybką identyfikację produktów, firmy i instytucje wykorzystują je do inwentaryzacji mebli i urządzeń, w bibliotekach ułatwiają ewidencję książek.

Świat organizmów żywych można porównać do ogromnej biblioteki, która nie posiada jeszcze kompletnego katalogu. Liczba pozycji zgromadzonych w tej bibliotece – gatunków – jest szacowana na ponad 10 milionów (a niektóre szacunki mówią o 30, a nawet 100 milionach), a opisano i „skatalogowano” zale-

dwie ok. 1,7 mln z nich (Stoeckle et al. 2003). W tak ogromnym zbiorze odpowiedni system znakowania wydaje się niezbędny do sprawnego identyfikacji już opisanych oraz ułatwienia opisu nowych dla nauki gatunków. Klasyczny sposób opisu i binominalna nomenklatura linneuszowska – aczkolwiek obowiązujące – okazują się niewystarczające wobec ogromnego tempa wymierania gatunków, braku specjalistów oraz nowych bieżących potrzeb i wyzwań jakie niesie życie. Stworzenie katalogu „kodów paskowych” („barkodów”) opisujących różnorodność biologiczną na poziomie gatunkowym jest, wbrew pozorom, pomysłem realnym (choć posiadającym swoje wyraźne ograniczenia). Każdy bowiem

organizm żywy posiada w swym genomie informację genetyczną zapisaną jako sekwencja nukleotydów, która stanowi swego rodzaju znacznik organizmu i może być wykorzystana do jego identyfikacji. Niektóre regiony genomu lepiej od innych pozwalają identyfikować organizmy i grupować je na różnych poziomach zróżnicowania (gatunkowym, rodzajowym, wewnątrzgatunkowym). Do identyfikacji gatunków wybrać trzeba obszar genomu jednolity u osobników reprezentujących tę właśnie rangę a wysoce zmienny na poziomie międzygatunkowym. Stworzenie „kodu paskowego DNA” to nic innego jak ustalenie specyficznej dla danego gatunku sekwencji nukleotydów w określonym *locus*.

DEFINICJE I USTALENIA TERMINOLOGICZNE

Nowa technika i związana z nią terminologia anglojęzyczna domaga się ustalenia także odnośnej terminologii w języku polskim. O ile sam *bar code* jako „kod paskowy” funkcjonuje już w języku polskim, o tyle tworzenie takich terminów czy wyrażań typu „kodowanie paskowe DNA” lub czegoś w tym rodzaju nie byłoby najlepszym rozwiązaniem. Stąd też proponujemy spolszczenie odnośnych terminów anglojęzycznych:

Barkoding DNA (ang. *DNA barcoding*) – na określenie techniki i całego procesu ustalania „kodu paskowego DNA” będącego *de facto* sekwencją DNA specyficzną dla danego gatunku. Kiedy mowa o *barkodingu DNA* (lub *barcodingu*) bez jakichkolwiek dookreśleń, zawsze należy mieć na myśli poziom gatunkowy.

Barkod DNA (ang. *DNA barcode*) – „kod paskowy DNA” – sekwencja DNA w wybranym, standardowym regionie genomu pozwalająca na identyfikację gatunku.

Tak ustaloną terminologią będziemy się posługiwać w dalszej części tekstu.

Barkoding DNA można zatem zdefiniować jako identyfikację gatunków za pomocą jednej lub kilku sekwencji DNA w określonym *locus* lub w kilku *loci*. Pomysł ten jest szeroko propa-

gowany od momentu opublikowania artykułów donoszących o udanym zastosowaniu mitochondrialnej sekwencji kodującej podjednostkę I genu oksydazy cytochromowej (*coxI*, COI) do rozróżniania gatunków zwierząt (Hebert et al. 2003a, b, 2004a, b). Stosowane obecnie metody identyfikacji gatunków: analizy morfologiczne czy innego typu badania molekularne często są skomplikowane, wymagają większej ilości czasu i dużej wiedzy specjalistycznej¹.

Z założenia system identyfikacji gatunków oparty o barkody DNA ma być:

- łatwy w obsłudze, tak by korzystać z niego mogli ludzie nie posiadający wiedzy specjalistycznej z zakresu taksonomii, genetyki molekularnej, itp.,
- szybki, aby użytkownicy otrzymali wyniki w jak najkrótszym czasie,
- standardowy w skali globalnej, co pozwoli korzystać z jednolitego zestawu danych użytkownikom na całym świecie,
- wiarygodny.

Umożliwienie łatwej i szybkiej identyfikacji gatunków dla szerokiego grona zainteresowanych, będące głównym celem barkodingu DNA, musi być poprzedzone pracą specjalistów reprezentujących wiele dziedzin nauki. Konieczne jest uzyskanie wiarygodnych sekwencji (barkodów DNA), oraz stworzenie bazy danych gromadzącej i udostępniającej otrzymane dane. Dla powszechnego zastosowania barkodingu DNA muszą zostać opracowane tanie i wygodne urządzenia do sekwencjonowania.

Sekwencję DNA można określić, korzystając z niewielkiej ilości materiału biologicznego. W związku z tym barkoding DNA pozwoli identyfikować organizmy na podstawie ich fragmentów, szczątków, a także rozmaitych form rozwojowych, często bardzo trudno iden-

¹ Sekwencjonowanie DNA, na którym to opiera się *barcoding DNA* jest również skomplikowaną i czasochłonną metodą. Jednym z elementów projektu opisu gatunków barkodami DNA jest więc m.in. stworzenie łatwych w obsłudze i szybkich urządzeń do sekwencjonowania, przy równoczesnej standaryzacji (używanie jednego lub paru zaledwie stałych regionów genomu do rozróżniania gatunków).

tyfikowalnych lub nieidentyfikowalnych w inny sposób.

Sekwencje rozmaitych genów i regionów niekodujących genomu wykorzystywane są już z powodzeniem w systematyce. Informacje uzyskiwane w ramach tworzenia bazy danych barkodów DNA staną się bardzo istotnym uzupełnieniem dotychczasowej wiedzy o zróżnicowaniu organizmów żywych, bowiem do tej pory nie prowadzono analizy zmienności jednej (lub kilku) sekwencji u wszystkich organizmów.

Znając już ideę barkodów DNA, można wymienić kilka potencjalnych jej zastosowań, takich jak:

- Opisywanie i ewidencjonowanie bioróżnorodności (na poziomie gatunkowym i nie tylko),
- Różne analizy ekologiczne i fitogeograficzne,
- Pomoc w wyróżnianiu i opisywaniu nowych gatunków,
- Analizy kryminalistyczne i ekspertyzy sądowe,
- Badanie składu żywności i leków,
- Kontrola „bezpieczeństwa biologicznego” i handlu gatunkami roślin czy zwierząt (m.in. objętymi ochroną).

CZYM NIE SĄ BARKODY DNA

Barkoding DNA posiada obecnie zarówno zwolenników i gorących propagatorów (np.: Blaxter 2004, Janzen 2004, Gregory 2005, Schindel, Miller, 2005), jak i krytyków (np.: Will, Rubinoff 2004, Ebach, Holdrege, 2005, Rubinoff et al. 2006). Aby uniknąć nieporozumień, należy podkreślić, że barkoding DNA nie jest próbą zastąpienia dotychczasowych metod taksonomii czy sposobu opisu gatunków lub innej rangi taksonów. Wręcz odwrotnie, wymaga współpracy specjalistów zajmujących się taksonomią różnych grup w celu choćby identyfikacji materiałów użytych do badań. Jak wspomnieliśmy we wstępie, barkody DNA mają ułatwić identyfikację gatunków, także tych jeszcze nie opisanych. Jednak barkody nie są i nie mają być, ani jedynym, ani głównym kryterium do

wyróżniania nowych gatunków. Nie można też twierdzić, że barkody DNA są konkurencją dla analiz filogenetycznych lub nawet powielaniem tychże. Wnioskowanie na temat pochodzenia gatunków na podstawie zróżnicowania tylko jednej lub kilku sekwencji, jak to nierzadko robiono, z pewnością nie przynosi wystarczająco wiarygodnych wyników. Dane, które daje barkoding DNA mogą stanowić jedynie uzupełnienie badań filogenetycznych (Stoeckle et al. 2003).

IDEALNY BARKOD DNA

Tylko nieliczne obszary genomu nadają się do wykorzystania jako barkody DNA. Dobitnie świadczą o tym wyniki poszukiwań standardu barkodu DNA do opisu gatunków roślin (piszemy o tym szerzej w dalszej części artykułu). Sekwencja przydatna w barkodingu DNA powinna spełniać następujące kryteria:

1. Niska zmienność w obrębie gatunków (ok. 2%), a równocześnie wysoka zmienność pomiędzy gatunkami. Cecha ta wynika wprost z definicji barkodingu DNA, która mówi o identyfikacji organizmów na poziomie gatunkowym, choć można wyobrazić sobie analogiczne postępowanie z użyciem innych sekwencji do wyróżniania jednostek np. wewnątrzgatunkowych.

2. Oflankowanie konserwatywnymi domenami. Specyfika sekwencjonowania DNA wymaga powielenia analizowanego fragmentu za pomocą techniki PCR (ang. *polymerase chain reaction*). Konserwatywne domeny flankujące sekwencję docelową pozwalają na zaprojektowanie startów PCR, dzięki którym możliwe jest pozyskanie tejże sekwencji z różnych gatunków. Pozwala to standaryzować barkody DNA.

3. Niewielka długość – do ok. 600 par zasad (pz). W swym założeniu identyfikacja gatunków za pomocą barkodów DNA ma być możliwie szybka i prosta. Sprzęt stosowany obecnie do sekwencjonowania DNA pozwala na stosunkowo szybką analizę sekwencji o długości do ok. 600 pz. Fragmenty o większej długości mogą być również sekwencjonowane, lecz wymaga to większych nakładów finansowych, oraz więcej czasu i pracy.

4. Niewielka liczba insercji/delekcji. Mutacje typu insercja i delekcja utrudniają porównywanie sekwencji (homologizację, ang. *alignment*) uzyskanych z różnych organizmów i w analizach powiązań genetycznych na ogół nie są traktowane jako cechy diagnostyczne. Pożądane są natomiast substytucje (podmiany) nukleotydów. Kress i Erickson (2007) odrzucają jednak ten wymóg, stwierdzając że barkod DNA ma służyć do identyfikacji gatunków, a nie do analiz filogenetycznych i każda mutacja, niezależnie od tego czy jest to insercja bądź delekcja czy też substytucja, jest cechą organizmu i jest równie dobra, jeśli tylko pozwala na ustalenie jego tożsamości.

5. Występowanie w genomie w dużej liczbie kopii. Cecha ta ma znaczenie ze względu na perspektywę zastosowania barkodów DNA do oznaczania organizmów na podstawie niewielkich fragmentów tkanek, czy też materiału kopalnego. Duża liczba kopii sekwencji docelowej ułatwia jej pozyskanie z niewielkiej ilości tkanki, czy też z materiału zdegradowanego, gdyż ze wzrostem liczby kopii danej sekwencji w genomie rośnie też szansa na zachowanie przynajmniej części z nich w stanie nieuszkodzonym.

CZY ISTNIEJE IDEALNY BARKOD DNA DLA ROŚLIN?

Obecnie jako standard „kodu paskowego DNA” dla królestwa zwierząt z powodzeniem stosuje wspomnianą wcześniej sekwencję *cox1*, pochodzącą z genomu mitochondrialnego. Jako standardy do barkodingu grzybów proponowane są: barkody oparte na *cox1* oraz na jądrowej sekwencji ITS (Internal Transcribed Spacer) – niekodującej domenie występującej w obrębie genów rybosomalnego DNA (rDNA) (por. All Fungi Barcoding, www.allfungi.org).

Opracowaniem standardu do opisu roślin barkodami DNA zajmuje się Grupa Robocza d.s. Roślin (Plant Working Group), działająca w ramach Consortium for the Barcode of Life (CBOL). Standardowa dla królestwa zwierząt sekwencja *cox1* u roślin lądowych wykazuje niewystarczającą zmienność (Kress et al.

2005) i może znaleźć zastosowanie do opisu barkodem DNA tylko niektórych grup glonów (Saunders 2005). Generalnie mitochondrialne DNA (mtDNA) roślin nie jest dobrym źródłem sekwencji przydatnych do barkodingu DNA, gdyż występują w nim częste rearanżacje oraz wymiana genów z genomem jądrowym, a konieczność standaryzacji wymaga wykorzystania sekwencji o stabilnej strukturze. Sekwencją najczęściej proponowaną jako standard w barkodingu DNA roślin był wspomniany już wcześniej ITS. Jednak ze względu na możliwość występowania różnych kopii tej domeny u jednego organizmu (tzw. paralogów) oraz niejasności w jej ewolucji, wspomniana wcześniej Grupa Robocza wykluczyła ITS z puli „kandydatów” na standard barkodów DNA dla świata roślin. Geny jądrowego DNA występują w zbyt małej liczbie kopii, aby można wykorzystać je do sekwencjonowania zdegradowanego DNA (*vide* Idealny barkod DNA, punkt 5). W ten sposób ograniczono obszar poszukiwań do DNA chloroplastowego (cpDNA). Występuje ono w komórkach roślinnych w wielu kopiach. Na cpDNA składają się zarówno sekwencje kodujące, jak i niekodujące. Te pierwsze wykazują niewielką zmienność, podczas gdy te drugie – stosunkowo dużą. Przy obecnym stanie wiedzy nie potrafimy wskazać idealnej sekwencji barkodu DNA w żadnym z genomów roślinnych (nDNA, cpDNA, mtDNA) (Kress et al. 2005, Newmaster et al. 2006). Konieczne jest zatem wykorzystanie kilku sekwencji o niższej zmienności, które wspólnie będą tworzyły barkod DNA. W lutym 2007 roku Grupa Robocza zaproponowała dwa zestawy sekwencji, z których każda obejmuje po trzy regiony cpDNA:

1. *rpoC1*, *matK*, *rpoB*.
2. *rpoC1*, *matK*, *trnH* – *psbA*.

Pierwszy zestaw sekwencji obejmuje tylko regiony kodujące, podczas gdy drugi zestaw składają się sekwencje dwóch genów i jednej domeny niekodującej (*trnH* – *psbA*). Prowadzone obecnie badania mają na celu określenie, która z przedstawionych grup sekwencji jest bardziej przydatna w tworzeniu barkodów DNA (Chase et al. 2007). Protokoły laboratoryjne i propozy-

cje sekwencji docelowych Grupy Roboczej d.s. Roślin zostały również opublikowane w Internecie (<http://www.kew.org/barcoding/index.html>). Obecnie trudno o rzetelną ocenę wartości proponowanych sekwencji docelowych, gdyż w przedstawionych materiałach brak opisu wyników sekwencjonowania i ich dyskusji. Wybór sekwencji uzasadniony jest tylko krótkim komentarzem.

Newmaster i współpracownicy (2006) zaproponowali, by identyfikację gatunków prowadzić w dwóch etapach:

1. sekwencja „pierwszorzędowa” pozwala zaszeregować obiekt do wyższej jednostki systematycznej (rząd, rodzina, rodzaj...),

2. sekwencja „drugorzędowa” pozwala określić przynależność gatunkową danego okazu w obrębie każdej z jednostek systematycznych, na jakie można różnicować obiekty za pomocą sekwencji „pierwszorzędowej”.

Pomysł ten rozwinęli Kress i Erickson (2007). Ocenili oni na podstawie dwóch kryteriów dziewięć sekwencji potencjalnie przydatnych w barkodingu DNA (*trnH-psbA*, *rbcL-a*, ITS1, *matK*, *rpoB2*, *rpoC1*, *accD*, *ycf5*, *ndhJ*). Po pierwsze analizowali uniwersalność amplifikacji, czyli możliwość powielania docelowych sekwencji za pomocą zestawu uniwersalnych starterów PCR. Drugim kryterium była zdolność do różnicowania gatunków. Wyniki powyższej analizy wskazują, że gen *matK*, choć bardzo dobrze różnicuje gatunki, nie może być amplifikowany uniwersalnymi starterami. Obecnie trwają testy starterów mających pozwolić na bardziej uniwersalną amplifikację fragmentu *matK* (Ki-Joong Kim, dane niepublikowane; CBOL Plant Working Group). Geny *rpoC1* i *rpoB* są natomiast dobrze powielane z użyciem uniwersalnych starterów, wykazują jednak niższy poziom zróżnicowania. Najłatwiejsze w powielaniu u różnych gatunków, a zarazem najlepiej je różnicujące, okazały się sekwencje *trnH-psbA* oraz *rbcL-a*. Według wspomnianych autorów druga z tych sekwencji pozwala z wysokim prawdopodobieństwem zaszeregować badany obiekt do rodzaju lub rodziny, podczas gdy pierwsza daje możliwość

identyfikacji gatunków w obrębie tychże grup systematycznych. Za pomocą takiego systemu Kress i Erickson (2007) byli w stanie rozróżnić ponad 85% z analizowanych par gatunków. Newmaster et al. (2007) przeprowadzili ocenę przydatności do barkodingu DNA siedmiu regionów (*Universal Plasmid Amplicon*, *rpoB*, *rpoC1*, *accD*, *rbcL*, *matK*, *trnH-psbA*), analizując ich sekwencje pochodzące z różnych gatunków rodzaju *Compsonura* (Myristicaceae). Ocenie podlegała zmienność wewnątrz- i międzygatunkowa. Spośród wymienionych sekwencji jedynie *matK* oraz *trnH-psbA* wykazują zróżnicowanie międzygatunkowe pozwalające na wykorzystanie ich jako barkodów DNA w rodzaju *Compsonura* (Myristicaceae).

Podsumowując tą część rozważań, podkreślić trzeba, że niemożliwe jest zastosowanie jednej uniwersalnej sekwencji DNA do stworzenia systemu barkodów DNA dla całego świata żywego. Należy się spodziewać, że barkoding DNA operował będzie różnymi standardami dla różnych jednostek taksonomicznych.

W samym świecie roślin prawdopodobnie nawet zastosowanie dwóch lub więcej sekwencji docelowych nie pozwoli na identyfikację wszystkich gatunków. Coraz większa komplikacja analiz może prowadzić do „konfliktu” pomiędzy komercyjnym a naukowym zastosowaniem barkodów DNA. Analizy komercyjne w założeniu mają być łatwe do wykonania, szybkie i tanie, podczas gdy dla celów naukowych liczy się przede wszystkim dokładność i pewność wyników.

PROBLEMY

Wykorzystanie sekwencji cpDNA do tworzenia barkodów DNA, pomimo znaczących zalet jest też obciążone istotną wadą. Chloroplasty, a więc również ich genom, są u roślin (z wyjątkiem drzew iglastych) przekazywane w linii żeńskiej. Z tego powodu cpDNA gatunku alloploidalnego jest uważane za siostrzane względem cpDNA gatunku będącego dawcą cytoplazmy, a co za tym idzie sekwencje cpDNA mogą nie pozwolić na odróżnienie gatun-

ków mieszańcowych i introgresantów (Chase et al. 2005, Cowan et al. 2006).

Problemem, który wiąże się z barkodami DNA, niezależnie od pochodzenia sekwencji docelowych, jest nierównowartościowość gatunków. Należy się spodziewać, że w wielu przypadkach barkody DNA pozwolą zaseregować badanego osobnika co najwyżej do pewnego kompleksu gatunków spokrewnionych. Przyszli użytkownicy barkodingu DNA nie będący taksonomami mogą być usatysfakcjonowani taką dokładnością oznaczenia. Dla wielu jednak specjalistów zajmujących się szczegółowo konkretnymi taksonami dokładność taka może dyskwalifikować barkoding DNA jako narzędzie badawcze (Chase et al. 2005, Cowan et al. 2006).

MIĘDZYNARODOWE KONSORCJUM BARKODINGU

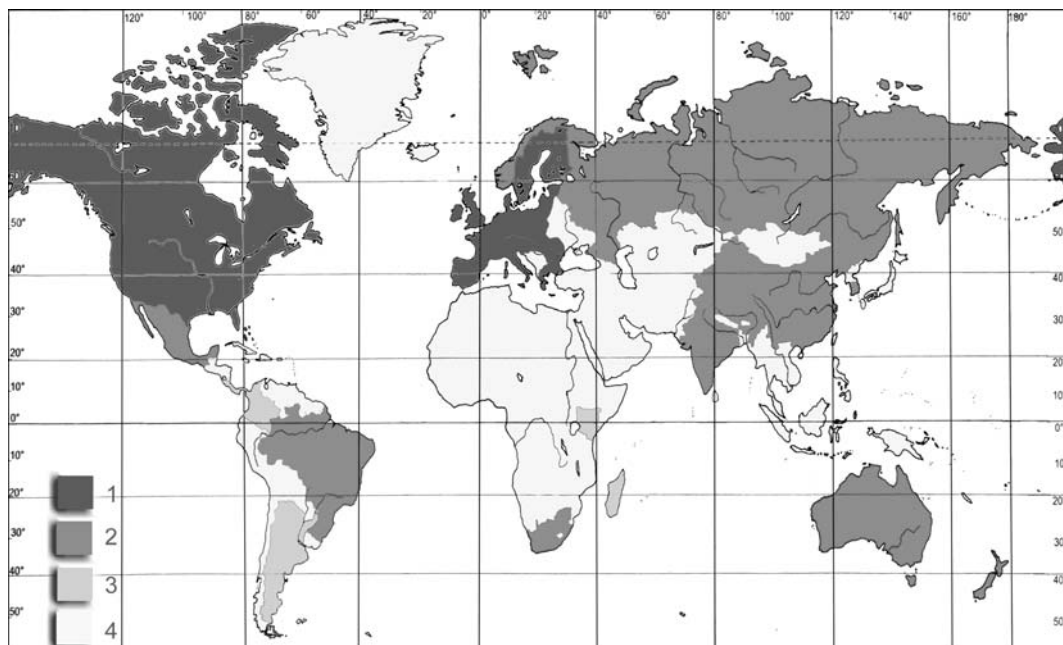
W krótkim czasie po opublikowaniu przez Dr Paula Heberta w 2002 roku propozycji barkodingu DNA (jak widać pomysł jest bardzo młody), odbyło się kilka roboczych warsztatów i spotkań, w efekcie których opracowano per-

spektywiczny projekt *Barcode of Life Initiative* (BOLI). Zgodnie z jego założeniami, barkoding DNA powinien funkcjonować jako technika standardowa w skali świata. Realizacja takiego postulatu wymaga koordynacji i planowania działań na szczeblu międzynarodowym, czym od 2004 roku zajmuje się *Consortium for the Barcode of Life* (CBOL). Jako inicjatywa o zasięgu ogólnoświatowym, CBOL skupia obecnie ponad 150 Organizacji Członkowskich z 50 krajów. Są to głównie muzea historii naturalnej i duże zielniki, placówki badawcze szkół wyższych, ogrody zoologiczne i botaniczne, instytucje zajmujące się ochroną przyrody, rządowe i prywatne agencje, których aktywność związana jest z badaniem i ochroną bioróżnorodności. Lokalizację wiodących placówek uczestniczących aktualnie w konsorcjum CBOL zaznaczono na załączonej mapce (Ryc. 1). Celem CBOL jest inicjowanie i ułatwianie współpracy między wspomnianymi instytucjami, a przez to wielopłaszczyznowe rozwijanie barkodingu DNA, oraz zwiększanie efektywności badań taksonomicznych zarówno w czysto naukowym wymiarze, jak i wykorzystywanie informacji z zakresu szeroko pojętej taksonomii do celów praktycznych. Konsorcjum



Ryc. 1. Lokalizacja organizacji i instytucji członkowskich Consortium for the Barcode of Life (CBOL).

Fig. 1. Location of Consortium for the Barcode of Life (CBOL) member organizations and institutions.



Ryc. 2. Stopień zaangażowania państw partycypujących w iBOL: 1 – komórki centralne, 2 – komórki regionalne, 3 – komórki rozwojowe, 4 – pozostałe kraje.

Fig. 2. Degree of commitment of iBOL participating countries: 1 – central nodes, 2 – regional nodes, 3 – developing nodes, 4 – other countries.

stymuluje i wspiera rozwój nowych technologii służących obsłudze barkodingu DNA od strony technicznej oraz poszukuje publicznych i prywatnych źródeł jego finansowania.

Przykładami największych dotychczas inicjatyw podejmowanych w ramach CBOL są: *All Birds Barcoding Initiative* (ABBI), *Fish Barcode of Life* (FISH-BOL) czy *International Network for Barcoding Invasive and Pest Species* (INBIPS). W obrębie wspomnianych inicjatyw o skali globalnej funkcjonują szczegółowe, inicjowane przez środowiska naukowe, projekty regionalne lub projekty ukierunkowane na konkretną jednostkę systematyczną.

Wielostronna aktywność CBOL jest nadzorowana przez Komitet Wykonawczy CBOL ustalający strategiczne kierunki działań, politykę i priorytety Konsorcjum. Przewodniczącym Komitetu Wykonawczego jest dr Scott Miller ze Smithsonian Institution, placówki która bezpośrednio przyczyniła się do powstania CBOL

i uformowania jego struktur organizacyjnych. Struktury te tworzy ponadto zarządzający finansami Sekretariat oraz Rada Naukowa zapewniająca istnienie płaszczyzny wymiany informacji w kwestiach naukowych i technicznych. Szczegółowe aspekty barkodingu DNA podejmowane są w ramach pięciu Grup Roboczych:

- *Data Analysis Working Group* (opracowywanie i rozwijanie sposobów i narzędzi do analizy danych);
- *Database Working Group* (tworzenie archiwum sekwencji DNA);
- *DNA Working Group* (doskonalenie poszczególnych etapów analiz genetycznych);
- *Technology Development Working Group* (opracowywanie i promowanie nowych systemów do uzyskiwania i wykorzystywania bar kodów DNA);
- *Plant Working Group* (identyfikacja rejonów DNA optymalnych dla barkodingu roślin).

Struktura CBOL w formie opisanych powyżej grup roboczych (sekcje problemowe) odzwierciedla podstawowe grupy potrzeb wyłaniające się na obecnym etapie rozwoju barkodingu DNA.

Jednym z najnowszych modeli szerszej współpracy międzynarodowej jest *International Barcode of Life Project* (iBOL), zainicjowany w 2006 roku przez Genom Canada. Kluczowym zamysłem tego przedsięwzięcia jest strukturyzacja regionalna, grupująca ośrodki naukowe na świecie w trzy „strefy aktywności” (Ryc. 2). Strefy te zostały wyodrębnione na podstawie aktualnego stopnia zaangażowania instytucji oraz ich potencjalnych możliwości uczestniczenia w barkodingu DNA. Stąd też priorytetowe obszary, gdzie ulokowane są główne centra barkodingu, stanowią Kanada, USA oraz kraje Unii Europejskiej. Rolę wspomagającą pełnią kraje będące tzw. komórkami regionalnymi: Australia, Brazylia, Chiny, Indie i Afryka Południowa. Trzecią grupę, tzw. komórek rozwojowych, stanowią obszary partycypujące w barkodingu na innych zasadach, np. poprzez dostarczanie materiałów do badań.

Oprócz jednostek naukowych zajmujących się badaniem i opisem bioróżnorodności, koncepcja iBOL kładzie także nacisk na współpracę z wszelkimi instytucjami, których aktywność powiązana jest z bioróżnorodnością i które zainteresowane są wykorzystywaniem barkodingu DNA jako narzędzia do opisu bioróżnorodności i ich wykorzystania w celach aplikacyjnych, takich jak ograniczanie ekspansji gatunków inwazyjnych, monitoring ekosystemów, ochrona bioróżnorodności, regulacje prawno-rynkowe, innowacje technologiczne.

Ponad dwa lata temu, w lutym 2005 roku w Londynie odbyła się pierwsza międzynarodowa konferencja poświęcona barkodingowi DNA. Od tego czasu nastąpił znaczący postęp w opracowywaniu kodów paskowych poszczególnych grup organizmów, zarówno w aspekcie teoretycznym, metodycznym jak i technologicznym. W dniach 17–21 września 2007 roku w Taipei na Tajwanie odbyło się kolejne przełomowe spotkanie, *The 2nd International Barcode of Life Conference*.

W gronie kilkuset przedstawicieli z ponad 45 państw podsumowano dotychczasowy przebieg barkodingu DNA, nakreślono kierunki dalszego rozwoju przedsięwzięcia oraz zwrócono uwagę na nierozwiązane dotąd problemy i potencjalne zagrożenia w harmonijnym funkcjonowaniu barkodingu DNA na świecie.

Jak wspomniano na wstępie, idea barkodingu DNA zrodziła się w efekcie analizowania wyników badań nad zmiennością sekwencji *COI*, a cechy tej sekwencji przyjęto jako cechy idealnego barkodu DNA dla większości zwierząt i grzybów. O wiele trudniejsze okazały się prace nad wyborem odpowiedniego rejonu czy rejonów w genomach roślin. W tym względzie wciąż brak rozstrzygającej decyzji. Nic więc dziwnego, że szczególne zainteresowanie i oczekiwania związane są dziś z pracami *Plant Working Group*, której powierzono misję stworzenia nowego systemu identyfikacji gatunków, opartego na innych niż *COI* regionach DNA. Badania w tym zakresie, trwające od 2004 roku koordynowane są przez Royal Botanic Gardens w Kew. Koncentrują się one na kilku kluczowych kwestiach, o których pisaliśmy powyżej.

BIBLIOTEKA BARKODÓW DNA

O zainteresowaniu ideą barkodingu DNA świadczy m.in. szybki wzrost liczby gatunków opatrzonych barkodem DNA. W specjalnie temu zadedykowanej bazie danych BOLD (*Barcode of Life Data Systems*) zarejestrowanych jest już ponad 39 tysięcy gatunków z barkodem DNA. Baza ta służy nie tylko do archiwizacji sekwencji DNA, ale zawiera także informacje taksonomiczne uzyskane na podstawie danych genetycznych (np. dendrogramy taksonów i dystanse genetyczne w obrębie poszczególnych jednostek systematycznych). Dane molekularne połączone są z opisem morfologicznym i anatomicznym gatunku oraz informacjami o jego rozmieszczeniu (poprzez linki do istniejącej literatury w tym zakresie), dzięki czemu baza BOLD stanowi swoisty pomost pomiędzy danymi molekularnymi i tradycyjną taksonomią.

BOLD funkcjonuje w oparciu o *COI* jako

standardowy region barkodujący, dlatego też stanowi bibliotekę sekwencji DNA głównie dla zwierząt. Jak dotąd nie istnieje baza przeznaczona dla barkodów roślinnych, ani jako część bazy BOLD, ani jako oddzielny system. Brak uniwersalnego sposobu rozróżniania gatunków roślin dzięki standardowym regionom DNA powoduje, iż informacje na temat gatunków, dla których zsekwencjonowano rejony potencjalnie użyteczne w barkodingu zamieszczane są w bazach o przeznaczeniu ogólnym, np. GenBank.

BARKODING W POLSCE

Koordinacja działań związanych z barkodingiem poprzez tworzenie struktur organizacyjnych typu CBOL skutkuje m.in. tendencją do tworzenia analogicznych struktur na poziomie poszczególnych kontynentów czy państw. Przykładem może być *Canadian Barcode of Life Network* (BOLNET), pierwsza krajowa sieć dedykowana barkodingowi DNA, zrzeszająca najważniejsze instytucje naukowo-badawcze w Kanadzie oraz aktywnie współpracująca z CBOL. Postulowana jest także konieczność koordynowania działań instytucji zainteresowanych barkodingiem w Europie oraz zaadaptowanie modelu specjalistycznych grup roboczych CBOL w skali kontynentu europejskiego. Zagadnieniom tym poświęcone było spotkanie *DNA Barcoding in Europe*, zorganizowane w Holandii w dniach 3–5 października 2007 roku w ramach sieci *European Distributed Institute of Taxonomy* (EDIT). Z Polski, w spotkaniu uczestniczyli przedstawiciele Instytutu Botaniki PAN oraz Muzeum i Instytutu Zoologii Polskiej Akademii Nauk w Warszawie placówek będących członkami EDIT. W efekcie prowadzonych w czasie spotkania dyskusji wyłoniono zasadnicze grupy zagadnień, z którymi konieczne jest zmierzenie się na gruncie europejskim:

- strategia barkodingu DNA w Europie,
- sieć europejskich banków DNA,
- identyfikacja priorytetów barkodingu DNA w Europie.

W międzynarodową współpracę związaną z barkodingiem DNA angażują się także polskie

instytucje naukowe. Od 2004 roku, a więc niemal od momentu pojawienia się idei barkodingu DNA, w Polsce istnieje tzw. sieć Barcodingu DNA, obejmująca 4 placówki aktywnie współdziałające ze sobą w celu prowadzenia tego przedsięwzięcia w naszym kraju. Poszczególne jednostki odpowiadają za barkoding DNA w grupach organizmów stanowiących główny przedmiot ich zainteresowań:

- Muzeum i Instytut Zoologii Polskiej Akademii Nauk w Warszawie – zwierzęta;
- Instytut Botaniki Polskiej Akademii Nauk w Krakowie – rośliny naczyniowe i grzyby;
- Leśny Bank Genów w Kostrzycy – drzewa leśne;
- Instytut Hodowli i Aklimatyzacji Roślin w Radzikowie – rośliny uprawne i ich dzikorośnące formy rodzicielskie.

Przedstawiciele dwu pierwszych placówek aktywnie uczestniczą również w pracach i spotkaniach międzynarodowej sieci barkodingu DNA. Aktywność krajowej sieci finansuje Ministerstwo Nauki i Szkolnictwa Wyższego.

LITERATURA

- BLAXTER M. L. 2004. The promise of a DNA taxonomy. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **359**(1444): 669–679.
- CHASE M. W., SALAMIN N., WILKINSON M., DUNWELL J. M., KESANAKURTH R. P., HAIDAR N., SAVOLAINEN V. 2005. Land plants and DNA barcodes: short-term and long-term goals. *Phil. Trans. R. Soc. B* **360**(1462): 1889–1895.
- CHASE M. W., COWAN R. S., HOLLINGSWORTH P. M., VAN DEN BERG C., MADRIÑÁN S., PETERSEN G., SEBERG O., JØRGENSEN T., CAMERON K., CARINE M., PEDERSEN N., HEDDERSON T. A. J., CONRAD F., SALAZAR G. A., RICHARDSON J. E., HOLLINGSWORTH M. L., BARRACLOUGH T. G., KELLY L., WILKINSON M. 2007. A proposal for a standardized protocol to barcode all land plants. *Taxon* **56**(2): 295–299.
- COWAN R. S., CHASE M. W., KRESS W. J., SAVOLAINEN V. 2006. 300,000 species to identify: problems, progress, and prospects in DNA barcoding of land plants. *Taxon* **55**(3): 611–616.
- EBACH M. C., HOLDREGE C. 2005. DNA barcoding is no substitute for taxonomy. *Nature* **434**: 697.

- GREGORY R.T. 2005. DNA barcoding does not compete with taxonomy. *Nature* **434**: 1067.
- HEBERT P. D. N., CYWINSKA A., BALL S. L., DEWAARD J. R. 2003a. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B. Biol. Sci.* **270**: 313–321.
- HEBERT P. D. N., RATNASINGHAM S., DEWAARD J. R. 2003b. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond. B. Biol. Sci.* **270**: 596–599.
- HEBERT P. D. N., STOECKLE M., ZEMLAK T., FRANCIS C. M. 2004a. Identification of Birds through DNA Barcodes. *PLoS Biology* **2**: 1657–1668.
- HEBERT P. D. N., PENTON E. H., BURNS J. M., JANZEN D. H., HALLWACHS W. 2004b. Ten Species in One: DNA Barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerater*. *Proc. Natl. Acad. Sci. U.S.A.* **101**(41): 12–17.
- JANZEN D. H. 2004. Now is the time. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **359**: 731–732.
- KRESS W. J., WURDACK K. J., ZIMMER E. A., WEIGT L. A., JANZEN D. H. 2005. Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci. U.S.A.* **102**(23): 8369–8374.
- KRESS W. J., ERICKSON D. L. 2007. A two-locus global DNA barcode for land plants: the coding *rbcl* gene complements the non-coding *trnH-psbA* spacer region. *PLOS One* **6**(e508): 1–10.
- NEWMASER S. G., FAZEKAS A. J., RAGUPATHY S. 2006. DNA barcoding in land plants: evaluation of *rbcl* in a multigenic tiered approach. *Can. J. Bot.* **84**: 335–341.
- NEWMASER S. G., FAZEKAS A. J., STEEVES R. A. D., JANOVEC J. 2007. Testing candidate plant barcode regions in the Myristicaceae. *Mol. Ecol. Notes*. OnlineEarly.
- RUBINOFF D., CAMERON S., WILL K. 2006. Are plant DNA barcodes a search for the Holy Grail? *Trends Ecol. Evol.* **21**: 1–2.
- SAUNDERS G. 2005. Applying DNA Barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Phil. Trans. R. Soc. B* **360**(1462): 1879–1888.
- SCHINDEL D. E., MILLER S. E. 2005. DNA barcoding a useful tool for taxonomists. *Nature* **435**: 17.
- STOECKLE M., JANZEN D., HALLWACHS W., HANKEN J., BAKER J. 2003. Taxonomy, DNA and the barcode of life. Meeting held at Banbury Center, Cold Spring Harbor Laboratory, New York, NY. Draft Conference Report.
- WILL K., RUBINOFF D. 2004. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics* **20**: 47–55.