

Ocena przydatności loci barkodowych do identyfikacji gatunków roślin łąk i muraw kserotermicznych. Komunikat

Barcoding loci usefulness for the identification of Polish meadow and xerothermophilous sward plant species. Short communication

Maja Boczkowska^{1,2}, Anna Rucińska¹, Marcin Olszak¹, Arkadiusz Nowak^{1,3}

¹ Polska Akademia Nauk Ogród Botaniczny – Centrum Różnorodności Biologicznej w Powsinie, 02-976 Warszawa, Polska,

² Instytut Hodowli i Aklimatyzacji Roślin – Państwowy Instytut Badawczy, Radzików, 05-870 Błonie, Polska,

³ Instytut Biologii, Uniwersytet Opolski, 45-052 Opole, Polska,

✉ e-mail: m.boczkowska@ihar.edu.pl

Analiza sekwencji kilku *loci* nosząca nazwę barkodingu, została wykorzystana do identyfikacji gatunków roślin występujących w środkowoeuropejskich zbiorowisk traworoślowych. W toku analizy przebadano użyteczność 14 regionów chloroplastowych i jednego jądrowego do identyfikacji gatunków. Osiem spośród nich (*matK*, *rbcl*, *rpoCl*, *trnH-psbA*, *atpF-atpH*, *trnL*, *psbI-psbK* i *ITS*) wykorzystano do stworzenia kombinacji barkodów, które umożliwiają określenie tożsamości próbki tkanki roślinnej o pochodzeniu łąkowym, bez konieczności porównania z kluczem do oznaczania roślin.

Słowa kluczowe: barkoding, *matK*, *rbcl*, *rpoCl*, *trnH-psbA*, *atpF-atpH*, *trnL*, *psbI-psbK*, *ITS*, traworośla

Barcoding analysis of the sequence of several loci was used to identify plant species that compose the Central European graminoid communities. During the analysis, the usefulness of 14 chloroplast regions and one nuclear region for species identification was investigated. Eight of these (*matK*, *rbcl*, *rpoCl*, *trnH-psbA*, *atpF-atpH*, *trnL*, *psbI-psbK* and *ITS*) were used to construct a combination of barcodes that will allow a determination of the taxonomic identity of a sample of plant tissue of meadow origin, without the need of utilising standard determination keys based on morphological characters.

Key words: barkoding, *matK*, *rbcl*, *rpoCl*, *trnH-psbA*, *atpF-atpH*, *trnL*, *psbI-psbK*, *ITS*

Wstęp

Poprawna identyfikacja gatunków jest problemem, z którym naukowcy borykają się od wieków. Ustalenie tożsamości taksonomicznej na podstawie plastycznego fenotypu, który może ulegać zmianom m. in. w toku ontogenezy czy pod wpływem środowiska, wymaga szerokiej wiedzy, doświadczenia i czasu. W dzisiejszym świecie, niejednokrotnie to właśnie czas jest czynnikiem, który najbardziej limituje prace naukowe. Na początku XXI wieku opracowana została metoda oznaczania gatunku na podstawie sekwencji jednego lub kilku *loci* (Hebert i in., 2003). Metoda ta została nazwana barkodingiem (ang. *barcoding*) przez analogię do wszechobecnych kodów paskowych identyfikujących rozmaite obiekty (ang. *bar code*). Sekwencje wykorzystywane w tej metodzie muszą spełniać kilka warunków. Efektywność ich amplifikacji w reakcji PCR

(ang. *polymerase chain reaction*) musi być na tyle wysoka, aby dało się ją przeprowadzić z niewielkiej ilości tkanki. W związku z tym powinny być obecne w genomie w wielu kopiach. Ponadto *loci* diagnostyczne muszą być oflankowane silnie konserwatywnymi sekwencjami, tak aby możliwe było wykorzystanie uniwersalnych starterów w reakcji PCR niezależnie od badanego gatunku. *Locus* wykorzystywane w barkodingu musi jednocześnie wykazywać wysoką zmienność na poziomie międzygatunkowym i niską zmienność wśród osobników reprezentujących ten sam gatunek (Mirek i in., 2007).

Grupa robocza działająca w ramach Consortium for the Barcode of Life (CBOL, 2009), które zajmuje się opracowaniem standardu dla identyfikacji molekularnej w świecie roślin, zalecała aby wykorzystywać w tym celu sekwencje znajdujące się w obrębie genomu plastydowego.

Genom mitochondrialny, który wykorzystywany jest powszechnie w identyfikacji zwierząt i grzybów, został odrzucony ze względu na występowanie w jego obrębie częstych rearanżacji i horyzontalnego transferu genów między mitochondrium i jądrem komórkowym (Hebert i in. 2003, Kress i in. 2015).

Za sprawą rozwoju technologii sekwencjonowania następnej generacji (ang. *Next Generation Sequencing*), dzięki któremu możliwa jest równoległa i równoczesna analiza sekwencji wielu *loci*, powstała technika jednoczesnej identyfikacji wielu gatunków np. w próbkach środowiskowych. Początkowo, stosowana była ona jedynie do wykrywania mikroorganizmów (Sogin i in. 2006). W kolejnych latach technika ta została zaadaptowana do identyfikacji organizmów eukariotycznych (Porazinska i in. 2010; Kowalczyk i in. 2011; Raye i in. 2011; Brown i in. 2012; Fouts i in. 2012; Hiiesalu i in. 2012). Warunkiem koniecznym do identyfikacji gatunku w próbach środowiskowych jest jednak odpowiednie nasycenie publicznie dostępnych baz danych wiarygodnymi sekwencjami wielu *loci* barkodowych.

Celem prezentowanych badań była ocena przydatności *loci* barkodowych do identyfikacji gatunków występujących w środkowoeuropejskich zbiorowiskach traworoślowych. Wytypowane w toku analizy *loci* zostaną wykorzystane w dalszych badaniach do oceny podziemnej różnorodności fitocenozy traworoślowych w Polsce.

Material i Metody

Materiał badawczy stanowiło 176 gatunków roślin typowych dla zbiorowisk traworoślowych, z czego 118 odnotowano w czasie wykonywania spisów florystycznych na trzech stanowiskach: Nowa Wieś Królewska (*Calthion palustris*; łąka bagienna), Opole-Kolonia Gosławicka (*Arrhenatherion elatioris*, łąka świeża) i Gogolin (*Mesobromion*, murawa kserotermiczna). DNA izolowano z młodych liści pojedynczych osobników o potwierdzonej tożsamości taksonomicznej przy użyciu komercyjnego zestawu Genomic Mini AX Plant (A&A Biotechnology) lub metody CTAB (Murray i Thompson, 1980) z modyfikacją autorską.

Do wytypowania regionów diagnostycznych dla każdego z gatunków posłużono się bazami danych NCBI i CBOL (Coordinators NCBI, 2017; CBOL, 2009). Sekwencje starterów do wybranych regionów pochodziły z danych literaturowych (Tab. 1). Reakcję PCR przeprowadzono z użyciem wysoce specyficznej polimerazy KAPA HiFi PCR Kit (Roche). Skład reakcji i profil temperaturowy zoptymalizowano osobno dla każdej pary starterów

wykorzystując dane literaturowe zamieszczone w Tabeli 1. Reakcję sekwencjonowania metodą Sanger'a wykonywano z użyciem zestawu Bright-Dye® Terminator Cycle Sequencing Kit (MCLAB). Rozdział fragmentów przeprowadzono w 3130xl Genetic Analyzers (Applied Biosystem) przy użyciu kapilar o długości 36 cm wypełnionych polimerem NanoPOP-7™ (MCLAB). Wyniki opracowywano przy użyciu FinchTV 1.4 software (Geospiza, Inc., Seattle, WA) i BioEdit (Hall i in., 2011).

Wyniki i Dyskusja

Na podstawie danych literaturowych wytypowano 14 regionów chloroplastowych i jeden jądrowy (Tab. 1), dla których przetestowano łącznie 45 par starterów. Podstawowym kryterium oceny starterów była efektywność amplifikacji pojedynczego fragmentu w jak największej liczbie gatunków. Najwięcej par starterów przetestowano dla regionu genomowego tj. niekodującej domeny *ITS* (ang. *Internal Transcribed Spacer*) występującej w obrębie genów rybosomalnego DNA (rDNA) i chloroplastowego regionu *matK* kodującego enzym maturazę K (Tab. 1). W przypadku dwóch regionów: *trnK* intron i *atpF* intron, mimo wielu prób optymalizacji, nie udało się uzyskać satysfakcjonującej jakości amplikonów dla wszystkich analizowanych gatunków roślin. Spośród 45 kombinacji par starterów ponad 30% generowało złożone wzory prążkowe reakcji PCR, czyli amplifikowało fragmenty niespecyficzne. W takich przypadkach uzyskanie informacji na temat specyficznej gatunkowo sekwencji DNA wiązałoby się z koniecznością izolacji właściwego fragmentu po rozdziale elektroforetycznym z żelu agarozowego. Ze względu na niewspółmierny do uzyskiwanych efektów nakład pracy, te pary starterów zostały odrzucone z dalszej analizy. Kolejne 14 par nie wykazywało uniwersalności w stosunku do części analizowanych gatunków, przez co nie zostały zakwalifikowane do dalszych badań. Udało się jednak z sukcesem wytypować z pozostałej grupy sześć par starterów, gwarantujących uzyskanie amplikonów spełniających kryteria regionów barkodowych, amplifikujących regiony genomu chloroplastowego tj.: *matK*, *rbcl*, *rpoCl*, *trnH-psbA*, *atpF-atpH*, *trnL* i *psbI-psbK* oraz jądrowego *ITS* (Tab. 1). Dla tych regionów w obrębie wytypowanej grupy gatunków roślin amplifikowano produkty o długości od 300 do 900 par zasad. Poszczególne regiony charakteryzowały się różną efektywnością w weryfikacji przynależności taksonomicznej, przy czym regiony takie jak *trnH-psbA*, *trnL* czy *ITS* pozwoliły na identyfikację do poziomu gatunkowego. Najniższa zmienność

sekwencji DNA charakteryzowała region *rbcL*, leżący w obrębie sekwencji kodującej dużą podjednostkę karboksylazy/oksygenazy rybulozo-1,5-bisfosforanu (RuBisCO), enzymu związanego z ciemną fazą fotosyntezy (Parry i in., 2003). Zmienność tej sekwencji związana była niemalże wyłącznie z obecnością polimorfizmów pojedynczych nukleotydów (SNP) (Rys. 1a). Locus *trnL* wykazywał średni poziom zróżnicowania, a obserwowana zmienność między gatunkami miała zarówno charakter mutacji typu SNP jak i insercji i delecji (In/Del) (rys. 1b). Wyniki te są zgodne z danymi literaturowymi, w których ten niekodujący region (intron) chloroplastowy nie jest uważany za wystarczająco zmienny do identyfikacji taksonów (Shaw i in., 2005). Największą rozdzielczość spośród analizowanych loci chloroplastowych odnotowano w przypadku regionu zlokalizowanego pomiędzy genami *trnH* i *psbA*. W regionie tym obserwowano zarówno mutacje typu SNP jak i obecność dużej liczby In/Del, których położenie i długość wykazywała dużą zmienność (rys. 1c). Użyteczność tego regionu potwierdzają liczne wcześniejsze prace badawcze m. in. Kress i in. (2005) i Shapcott i in. (2015). Zestawienie wszystkich uzyskanych sekwencji DNA pozwoliło na poprawną identyfikację taksonomiczną analizowanych gatunków. Tym samym wytypowano dla każdego gatunków sekwencje DNA dla min. trzech regionów barkodowych. Nieskuteczność identyfikacji taksonomicznej roślin przy użyciu pojedynczego locus barkodowego zauważono już na wstępie przystosowywania tej metody do analizy świata roślin. Różne grupy badawcze sugerowały różnorodne, efektywne w identyfikacji zestawy loci diagnostycznych. Kress i in. (2005) wskazywali na dwa niekodujące regiony (jądrowy obszar *ITS* i chloroplastowy *trnH-psbA*). Natomiast w kolejnej pracy Kress i Erickson (2007) jako regiony odpowiednie do barkodingu u roślin rekomendowali *rbcL* i *trnH-psbA*. Shapcott i in. (2015, 2017) wykorzystywali w badaniach lasów deszczowych Australii trzy regiony (*rbcL*, *matK*, i *trnH-psbA*). Z kolei Braukmann i in. (2017) do identyfikacji roślin naczyniowych w Kanadzie wykorzystali *rbcL*, *matK* i *ITS2*. Obecnym standardem w aplikacji metody barkodignu dla gatunków roślin jest wykorzystanie sekwencji DNA kilku regionów jednocześnie. Wytypowanie tych regionów wymaga jednak przeprowadzenia wstępnej analizy wielu loci pod kątem ich użyteczności dla wybranej grupy obiektów. Uzyskane w ramach niniejszej pracy sekwencje DNA wytypowanej grupy gatunków roślin stanowią podstawę do weryfikacji składu gatunkowego zbiorowisk łąkowych.

Część wyników została zaprezentowana podczas Konferencji Dni Młodego Naukowca w 2019 r. w Radzikowie.

Literatura

- Braukmann TWA., Kuzmina ML., Sills J., Zakharov EV., Hebert PDN. 2017. Testing the efficacy of DNA barcodes for identifying the vascular plants of Canada. PLoS ONE; 12(1): e0169515.
- Brown DS., Jarman SN., Symondson WO. 2012. Pyrosequencing of prey DNA in reptile faeces: analysis of earthworm consumption by slow worms. Mol Ecol Res; 12: 259-266
- CBOL Plant Working Group (2009) A DNA barcode for land plants. Proc Natl Acad Sci; 106, 12794-12797.
- Coordinators NCBI Resource., 2017. Database resources of the national center for biotechnology information. Nucleic Acids Res; 45: D12-D17.
- Cox CJ., Goffinet B., Shaw AJ., Boles SB. 2004. Phylogenetic relationships among the mosses based on heterogeneous Bayesian analysis of multiple genes from multiple genomic compartments. Syst Bot; 29: 234-250.
- Cuénoud P., Savolainen V., Chatrou L. W., Powell M., Grayer R. J., Chase M. W. 2002. Molecular phylogenetics of *Caryophyllales* based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *matK* DNA sequences. Am J Bot; 89: 132-144
- Demaio PH., Barfuss MHJ., Kiesling R., Till W., Chiappella JO. 2011. Molecular phylogeny of *Gymnocalycium* (*Cactaceae*): assessment of alternative infrageneric systems, a new subgenus, and trends in the evolution of the genus. Am J Bot; 98: 1841-1854.
- Dobeš C., Mitchell-Olds T., Koch M. 2004. Extensive chloroplast haplotype variation indicates Pleistocene hybridization and radiation of North American *Arabis drummondii*, *A. divaricarpa*, and *A. holboellii* (*Brassicaceae*). Mol Ecol; 13: 349-370.
- Dong W., Xu C., Li C., Sun J., Zuo Y., Shi S. i in. 2015. *ycf1*, the most promising plastid DNA barcode of land plants. Sci Rep; 5, 8348.
- Fazekas AJ., Burgess KS., Kesanakurti PR., Graham SW., Newmaster SG., i in. 2008. Multiple Multilocus DNA Barcodes from the Plastid Genome Discriminate Plant Species Equally Well. PLoS ONE; 3(7): e2802
- Ford CS., Ayres KL., Toomey N., i in. 2009 Selection of candidate coding DNA barcoding regions for use on land plants. Bot J Linn Soc; 159:1-11
- Fouts DE., Szpakowski S., Purushe J., Torralba M., Waterman RC., MacNeil MD., Alexander LJ., Nelson KE. 2012. Next Generation Sequencing to Define Prokaryotic and Fungal Diversity in the Bovine Rumen. PLoS One; 7: e48289.
- Hall, T., Biosciences, I., Carlsbad, C., 2011. BioEdit: an important software for molecular biology. GEF Bull Biosci; 2: 60-61.

- Hamilton MB. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Mol Ecol*; 8: 521–523.
- Hebert PD., Cywinska A., Ball SL., 2003. Biological identifications through DNA barcodes. *Proc Royal Soc London B: Biol Sci*; 270, 313–321.
- Hiiesalu I., Opik M., Metsis M., Lilje L., Davison J., Vasar M., Moora M., Zobel M., Wilson SD., Partel M., 2012. Plant species richness belowground: higher richness and new patterns revealed by next-generation sequencing. *Mol Ecol*; 21: 2004–2016.
- Kowalczyk R., Taberlet P., Coissac E., Valentini A., Mique C., Kamiński T., Wójcik JM. 2011. Influence of management practices on large herbivore diet—Case of European bison in Białowieża Primeval Forest (Poland). *For Ecol Manag*; 261: 821–828.
- Kress WJ., Erickson DL. 2007. A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE*; 2(6): e508.
- Kress WJ., García-Robledo C., Uriarte, M., Erickson D. L., 2015. DNA barcodes for ecology, evolution, and conservation. *Trends Ecol Evol*; 30, 25–35.
- Kress WJ., Wurdack KJ., Zimmer EA., Weigt LA., Janzen DH. 2005. Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci USA*; 102: 8369–8374.
- Lahaye R., Van der Bank M., Bogarin D., Warner J., Pupulin F., Gigot G. i in. 2008. DNA barcoding the floras of biodiversity hotspots. *Proc Natl Acad Sci USA*; 105: 2923–2928.
- Moorhouse-Gann RJ., Dunn JC., Vere N., Goder M., Cole N., Hipperson H., Symondson WOC. 2018. New universal *ITS2* primers for high-resolution herbivory analyses using DNA metabarcoding in both tropical and temperate zones. *Sci Rep*; 8: 29414.
- Mirek, Z., Bieniek, W., Sztorc, A., 2007. Barkoding DNA—nowe narzędzie do opisu bioróżnorodności. *Wiad Bot*; 51: 41–50.
- Murray, MG., Thompson, WF., 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res*; 8: 4321–4326.
- Parry MA., Andralojc PJ., Mitchell RA., Madgwick PJ., Keys AJ. 2003. Manipulation of Rubisco: the amount, activity, function and regulation. *J Exp Bot*; 386: 1321–1333.
- Porazinska DL., Giblin-Davis RM., Esquivel A., Powers TO., Sung W., Thomas WK. 2010. Ecometagenetics confirms high tropical nematode diversity. *Mol Ecol*; 19: 5521–5530.
- Raye' G., Miquel C., Coissac E., Redjadj C., Loison A., Taberlet P. 2011. New insights on diet variability revealed by DNA barcoding and high-throughput pyrosequencing: chamois diet in autumn as a case study. *Eco Res*; 26: 265–276.
- Russell A., Samuel R., Rupp B., Barfuss MH., Šafran M., Besendorfer V., Chase MW. 2010. Phylogenetics and cytology of a pantropical orchid genus *Polystachya* (*Polystachyinae*; *Vandaeae*; *Orchidaceae*): Evidence from plastid DNA sequence data. *Taxon*; 59: 389–404.
- Samuel R., Kathirarachchi H., Hoffmann P., Barfuss MH., Wurdack KJ., Davis CC., Chase, MW. 2005. Molecular phylogenetics of *Phyllanthaceae*: Evidence from plastid *matK* and nuclear *PHYC* sequences. *Am J Bot*; 92: 132–141.
- Shapcott A., Forster PI., Guymer GP., McDonald WJF., Faith DP., Erickson DL., Kress WJ. 2015. Mapping biodiversity and setting conservation priorities for SE Queensland's rainforests using DNA barcoding. *PLoS ONE*; 10: e0122164.
- Shapcott A., Liu Y., Howard M., Forster PI., Kress WJ., Erickson DL., Faith DP., Shimizu Y., McDonald WJF. 2017. Comparing floristic diversity and conservation priorities across south east Queensland regional rain forest ecosystems using phylodiversity indexes. *Int J Plant Sci*; 178: 211–229.
- Shaw J., Lickey EB., Beck JT., Farmer SB., Liu W., Miller J., Siripun K.C., Winder CT., Schilling EE., Small RL., 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am J Bot*; 92: 142–166.
- Sogin ML., Morrison HG., Huber JA., Welch DM., Huse SM., Neal PR., Arrieta JM., Herndl GJ. 2006. Microbial diversity in the deep sea and the underexplored “rare biosphere.”. *Proc Natl Acad Sci USA*; 103: 12115–12120.
- Taberlet P., Gielly L., Pautou G., Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol Biol*; 17: 1105–1109.
- Weising K., Gardner RC. 1999. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome*; 42:9–19.
- Yu J., Xue JH., Zhou SL. 2011. New universal *matK* primers for DNA barcoding angiosperms. *J Syst Evol*; 49(3), 176–181.

Tabela 1

Table 1

Lista testowanych starterów [S- pojedynczy amplikon; M-amplikony niespecyficzne; pogrubieniem zaznaczono startery, które wykorzystano do pełnej analizy]

The list of tested primers [S- single amplicon, M- multiple amplicons; primers that were used for full analysis were marked in bold]

region	primer name	sequence (5'-3')	amplicons	range of amplicon size [bp]	present	references		
ITS	ITS-p5	CCTTATCAYTTAGAGGAAGGAG	S	800 - 900	all species	Cheng i in., 2016		
	ITS-u4	RGTTTCTTTTCCTCCGCTTA						
	ITS-u1	GGAAGKARAAGTCGTAACAAGG	M					
	ITS-u4	RGTTTCTTTTCCTCCGCTTA						
	ITS-p5	CCTTATCAYTTAGAGGAAGGAG	S	400-500	all species			
	ITS-u2	GCGTTCAAAGAYTCGATGRITC						
	ITS-u1	GGAAGKARAAGTCGTAACAAGG	S	400-500	all species			
	ITS-u2	GCGTTCAAAGAYTCGATGRITC						
	ITS-p3	YGACTCTCGGCAACGGATA	S	400-500	all species			
	ITS-u4	RGTTTCTTTTCCTCCGCTTA						
	ITS-u3	CAWCGATGAAGAACYGAGC	M					
	ITS-u4	RGTTTCTTTTCCTCCGCTTA						
	P17a	CTACCGATTGAATGGTCCGGTGAA	S	800-900	all species		Popp i Oxelman, 2001	
	26S-82Ra	TCCCGGTTCCGCTCGCCGTACTA						
ITS	ITS1	TCCGTAGGTGAACCTGCGG	M			White i in., 1990		
	ITS4	TCCTCCGCTTATTGATATGC						
	ITS5 (F)	GGAAGTAAAAGTCGTAACAAGG	M					
	ITS4 (R)	TCCTCCGCTTATTGATATGC						
	ITS1	TCCGTAGGTGAACCTGCGG	M					
	ITS2	GCTGCGTTCTTCATCGATGC						
	ITS5 (F)	GGAAGTAAAAGTCGTAACAAGG	M					
	ITS2	GCTGCGTTCTTCATCGATGC						
	ITS3	GCATCGATGAAGAACGCAGC	M					
	ITS4	TCCTCCGCTTATTGATATGC						
	UniPlantF	TGTGAATTGCARRATYCMG	S	300 - 550	all species		Moorhouse-Gann i in., 2018	
	UniplantR	CCCGHYTGAYYTRGGTDC						
	matK	matK-472F	CCCRTYCATCTGGAAATCTGGTTC	S	300-900		not all species	Yui in., 2011
		matK-1248R	GCTRTRATAATGAGAAAGATTCTGC					
matK-KIMIR		ACCCAGTCCATCTGGAAATCTGGTTC	S	800-900	all species	CBOL, 2009		
matK-KIM3F		CGTACAGTACTTTTGTTTACGAG						
matK2.1a-f		ATCCATCTGGAAATCTTAGTTC	M			Ford i in., 2009		
matK2.1ar		GTTCTAGCACAAAGAAAGTCG						
matK2.1F		CCTATCCATCTGGAAATCTTAG	M					
MatK.5r R		GTTCTAGCACAAAGAAAGTCG						
matK-390f		CGATCTATTCATTCAATATTC	S	800-900	all species	Cuénoud i in., 2002		
matK-1326r		TCTAGCACACGAAAGTCGAAGT						
matK2.1F		CCTATCCATCTGGAAATCTTAG	S	800-900	not all species	Ford i in., 2009		
matK3.2 r		CTTCTCTGTAAAGAATTC						
550Fdi		CTRGAAATCTGGTTCAARCTCTTCG	M			Demaio i in., 2011		
750Rca		TGATCGTAAATGAGAGGATTGGTTACAG						
trnK 570F	TCCAAAATCAAAGAGCGATTGG	M			Russell i in., 2010			
190R	ATTCGAGTAATTAACGTTTACAA				Samuel i in., 2005			
80F	CTATACCA CTATCTTTCGGGAG	no ampli-cons			Samuel i in., 2005			
530R	GTTCCAATTCCAATACTCGTGAAG							
matK-390f	CGATCTATTCATTCAATATTC	M			Cuénoud i in., 2002			
530R	GTTCCAATTCCAATACTCGTGAAG				Samuel i in., 2005			
80F	CTATACCA CTATCTTTCGGGAG	no ampli-cons			Samuel i in., 2005			
190R	ATTCGAGTAATTAACGTTTACAA							
matK-472F	CCCRTYCATCTGGAAATCTGGTTC	S	250-300	not all species				
750Rca	TGATCGTAAATGAGAGGATTGGTTACAG							

Lista testowanych starterów [S- pojedynczy amplikon; M-amplikony niespecyficzne; pogrubieniem zaznaczono startery, które wykorzystano do pełnej analizy]

The list of tested primers [S-single amplicon, M- multiple amplicons; primers that were used for full analysis were marked in bold]

region	primer name	sequence (5'-3')	ampli-cons	range of amplicon size [bp]	present	references
<i>rbcL</i>	rbcLa-F	ATGTCACCACAAACAGAGACTAAAGC	M			Kress i Erickson, 2007
	rbcLa-R	GTAAAATCAAGTCCACCRCG				
	rbcLajf634R rbcL-aaf F	GAAACGGTCTCTCCAACGCAT ATGTCACCACAAACAGAGACTAAAGC	S	500-600	all species	CBOL, 2009
<i>rpoC1</i>	rpoC1 1f	GTGGATACACTTCTTGATAATGG	S	600-700	not all species	Fazekas i in., 2008
	rpoC1 4r	CCATAAGCATATCTTGAGTTGG	S	500-600	all species	
	rpoC1 2f	GGCAAAGAGGGGAAGATTTTCG				
	rpoC1 4r	CCATAAGCATATCTTGAGTTGG				
<i>trnH-psbA</i>	psbAF trnH2	GTTATGCATGAACGTAATGCTC CGCGCATGGTGGATTACAATCC	S	400-700	all species	Kress i in., 2005
	psbA trnH(GUG)	CGAAGCTCCATCTACAAATGG ACTGCCTTGATCCACTTGCC	S	400-700	not all species	Hamilton, 1999
	psbA501f trnHR	TTTCTCAGACGGTATGCC GAACGACGGGAATTGAAC	S	400-700	not all species	Cox i in., 2004
	<i>yef</i>	yef1bF yef1bR yef1gF yef1gR	TCTCGACGAAAATCAGATTGTTGTAAT ATACATGTCAAAGTGATGGAAAA TGAAAGCTCTAAGCAATGGATCYCC ATACGACCAATATTTTTRGCTATTAT	S M	800-900	not all species
atpF atpH		ACTCGCACACACTCCCTTTC GCTTTTATGGAAGCTTTAACAAT	S	400-800	all species	Fazekas i in., 2008
psb1F psbKR		AGAGTTTGAGAGTAAGCAT TTAGCCTTTGTTTGCAAG	S	350-500	not all species	Lahaye i in., 2008
<i>trnK intron</i> <i>5' to trnS</i>		ccmp1 ccmp2	5'-CAGGTAAACTTCTCAACGGA-3' 5'-CCGAAAGTCAAAGAGCGATT-3' 5'-GATCCCGGACGTAATCCTG-3' 5'-ATCGTACCGAGGGTTCGAAT-3'	M S	100-400	all species
<i>trnG intron</i>	ccmp3	5'-CAGACCAAAGCTGACATAG-3' 5'-GTTTCATTTCGGCTCCTTTAT-3'	S	100-150	not all species	
<i>atpF intron</i>	ccmp4	5'-AATGCTGAATCGAYGACCTA-3' 5'-CCAAAATATTBGGAGGACTCT-3'	M			
<i>ORF 77-ORF82 intergenic</i>	ccmp5	5'-CGATGCATATGTAGAAAGCC-3' 5'-CATTACGTGCGACTATCTCC-3'	S	100-150	not all species	
<i>rpl20-rpl12 intergenic</i>	ccmp6	5'-TTGGCTACTCTAACCTTCCC-3' 5'-TTCTTTCTATTTCGAGDGA-3'	S	100-200	not all species	
<i>trnL</i>	trnL-FF trnL-FR	GGTTCAAGTCCCTCTATCCC GATTTTCAGTCTCTGCTCTA	S	400-500	not all species	Dobeš i in., 2004
	trnL-FF trnL-FRTab	GGTTCAAGTCCCTCTATCCC ATTTGAACTGGTGACACGAG	S	400-600	all species	Dobeš i in., 2004 Taberlet i in., 1991
	trnL-Ffor trnL-FR	GGTTCAAGTCCCTCTATCATCCC GATTTTCAGTCTCTGCTCTA	S	500-800	not all species	Taberlet i in., 1991 Dobeš i in., 2004

Rys.1 Schematyczny obraz zróżnicowania w obrębie trzech regionów dla 16 przykładowych gatunków: a) *rbcl*; b) *trnL*; c) *trnH-psbA* [górny schemat przedstawia obecność in/del – delekcje zaznaczone na szaro; dolny schemat przedstawia obecność SNP – nukleotydy zaznaczone kolorami]

Fig. 1 Schematic image of the diversity within three regions for 16 different species: a) *rbcl*; b) *trnL*; c) *trnH-psbA* [the upper diagram shows the presence of in/del - deletions marked grey; the lower diagram shows the presence of SNP - nucleotides marked with colours].

