

日本産樹木DNAバーコーディングの現状 (2009)

吉村研介・鈴木節子 (森林総研)・田中孝尚・鈴木三男 (東北大)・神保宇嗣・伊藤元己 (東大)・館田英典 (九大)・大谷雅人・勝木俊雄・津村義彦・藤井智之・能城修一・河原孝行・吉丸博志 (森林総研)

要旨: DNA バーコーディングとは、特定の遺伝子領域の短い塩基配列(DNA バーコード)で生物種の同定を行う方法である。2003年に動物分野でミトコンドリア DNA の遺伝子領域 *COI* を対象とし、植物では、当初、葉緑体 DNA の *rbcL* 部分配列と *trnH-psbA* 遺伝子間領域が提唱された。2009年11月に *rbcL* と *matK* 部分配列を植物における DNA バーコードの標準とすることが決められた。筆者らのグループは日本産樹木を対象とする DNA バーコードの開発を2008年に開始し、*rbcL* 部分配列、*trnH-psbA* 遺伝子間領域での解析を行ってきた。*rbcL* 部分配列は、変異性が低く植物の科や属レベルの分類に有効であった。*trnH-psbA* 遺伝子間領域は、変異性は高いが、DNA 配列に欠失挿入が多いため長さの変異が大きく、単一塩基反復が多くみられるなど塩基配列の解析が困難な場合があった。今後は *rbcL* と *matK* 部分配列を中心に解析を進める。将来的に「証拠標本」「DNA バーコード」の2つのデータベースの構築を進め、同定支援システムと併せて、公開する予定である。

キーワード: DNA バーコーディング, 日本産樹木, 葉緑体 DNA, *rbcL*, *trnH-psbA* 遺伝子間領域

I はじめに

DNA バーコーディングとは、特定の遺伝子領域の短い塩基配列(DNA バーコード)で生物種の同定を行う方法である。2003年に動物において、ミトコンドリア DNA のチトクローム c オキシダーゼサブユニット I 遺伝子(*COI*)の部分配列(648bp)を使うことが、Hebert らによって提唱された(4)。短い DNA 領域ではあるが、全ての種で単一の領域についての情報を明らかにすることにより、これまで形態が似ていて同一種と同定されていた種が、遺伝学的には別種(隠ぺい種)であることが判明したりしている。また、種同定の特徴となる形態が発現していない時期の個体や種同定不能な部分からでも、種同定が可能になる等の利点がある。全ての生物種のバーコーディングを行うことを目的に、2004年に Consortium for the Barcode of Life (CBOL)が結成され各種のプロジェクトが世界的に進められている。

植物においては、2005年に Kress らが、ミトコンドリア DNA の *COI* 領域は保存性が高く変異が少ないことか

ら、葉緑体の *rbcL* の部分配列(約700bp)、*trnH-psbA* 遺伝子間領域を使うことを提唱した(6)。2007年第2回国際 Barcode of Life (BOL)会議では、情報量不足を補うためさらに複数の領域を使う必要があることが複数報告された(5)。その後、12か所の葉緑体遺伝子領域を比較した結果、*matK*, *rpoB*, *rpoC1*, *ndhJ*, *ycf5*, *accD* の6つが、バーコード領域候補として有力であるという報告がなされた(3)。CBOL Plant Working Group(2009)は、葉緑体の4つの遺伝子領域と3つの遺伝子間領域を比較検討した結果、植物では *rbcL* 部分配列と *matK* 部分配列を第一のターゲットとして使うことを提唱し(1)、2009年11月第3回国際 Barcode of Life (BOL)会議で合意された。一方、日本においては、植物全体について組織だった DNA バーコーディングは行われておらず、個々の研究のため、筆者らのグループは日本産樹木を対象に DNA バーコーディングを2008年より進めている。

II 材料と方法

筆者らの研究グループの内、東北大学と森林総研木材

Kensuke YOSHIMURA, Suzuki SETSUKO (For. and Forest Prod. Res. Ins., Ibaraki 305-8687), Takahisa TANAKA, Mitsuo SUZUKI (Tohoku Univ., Sendai 980-0862), Utsugi JINBO, Motomi ITO (Univ. of Tokyo, Meguro-ku 153-8902), Hidenori TACHIDA (Kyushu Univ., Fukuoka 810-8560), Masato OHTANI, Toshio KATSUKI, Yoshihiko TSUMURA, Tomoyuki FUJII, Shuichi NOSHIRO, Takayuki KAWAHARA, Hiroshi YOSHIMARU (FFPRI, Ibaraki 305-8687)

DNA barcoding on Japanese woody plants (2009)

特性研究領域は、1998年より国内樹種の木材標本・押し葉試料を、全国各地で26か所で採集をしており、これをメインの試料とした(表-1)。2008年より、これらのDNA抽出を開始し、その他の収集した試料も併せて、DNA解析を進めている。これまでのところ102科、294属、721種、826種・亜種・変種、4264個体のDNA抽出を行った。過去に収集したDNA用試料は、シリカゲルで乾燥し常温保存していた。そのためDNAが劣化してDNA解析に使えない試料もあった。2008年以降は、新鮮なサンプルを冷凍保存し随時DNA抽出するようにしている。DNA試料は、2007年の情報に基づきバーコードDNA領域の塩基配列を*rbcL*部分配列、*trnH-psbA*遺伝子間領域について解析を進めた。先行研究に示されたプライマーの設計が悪く増幅しない種が多くみられたため、ユニバーサルなプライマーを設計し直して解析を行った。シーケンスは、両鎖を読んで両側とも良好な物のみを解析対象とした。

III 結果と考察

一部のサンプルがシーケンスまで至らず欠測値となったため、それらのサンプルを除いた集計値を表-2にまとめた。*rbcL*部分配列は、98科、280属、675種、3624個体シーケンスが明らかになった。供試個体に対するシーケンスの成功率は全体で85%であった。2006年、2007年の乾燥葉から抽出したものが、95%以上の成功率で、生葉から抽出した2008年以降と変わらない。2000年から2002年にかけてのDNAサンプルが、38-82%と成功率が低かった。1998年と1999年のサンプルが76-92%なので、この期間のサンプルは、乾燥前の処理、保存もしくはDNA抽出に何か問題があったと考えている。失敗例のほとんどが、DNAが希薄すぎたか、劣化しているか、PCR実験の際のサンプル注入ミスやコンタミネーションのためと推定しているが、ナツツバキ属のヒメシャラのみは、プライマーが適合していない可能性がある。

*trnH-psbA*遺伝子間領域は、92科、255属、613種、3035個体のシーケンスを明らかにした。シーケンスの成功率は、全体で73%と*rbcL*と比べて低かった。*rbcL*で好成績だった2006年以降のサンプルでも、72-94%と低かった。プライマーのユニバーサル性が*rbcL*に比べて若干低いようで、裸子植物のいくつかの属ではPCRの増幅が見られなかった。この遺伝子間領域は、polyAやpolyTの単塩基反復配列が多く含まれ、PCRの増幅が見られても、スリッピングなどのため良好なシーケンスが得られない場合がある。特にバラ科のいくつかの属でその傾向が強かった。

表-3に種同定可能かどうかを示した。*rbcL*部分配列では、675種のうち347種(51%)が、現在のデータでは種同定に成功していたが、半数近くの種が近縁種と同一シーケンスのものがあり同定不能であった。これは、*rbcL*は比較的保存性が高いことと、対象とした塩基配列が599bpと短いため種間の差を検知できないためである。*rbcL*部分配列はPCR増幅もよく、良好なシーケンスが容易に得られるため、属レベルの同定には非常に有効であるが、分類群によっては種レベルの同定が不可能な場合がある。

*trnH-psbA*遺伝子間領域は、変異性は高いが、塩基配列の挿入・欠失が多く、長さが145bpから700bp以上のものまで多岐にわたっていた。長さが700bp以上だと1回のシーケンス操作で全て読み取れないためデータを取るのが困難になる。613種のうち454種(74%)が、種同定可能であった。変異性が高すぎるため、属内の種間のアラインメントが困難な分類群もあり系統関係の研究等には支障がある。種内多型も数多く見受けられ、DNAバーコーディングによる種同定を目的とした場合、誤同定の可能性が高くなる。一方、種内多型の情報を、地域情報とともにデータベース化しておけば、種同定の精度を高くすることもできる。また良好なシーケンスを得られにくい場合があるのも問題である。

2つの配列領域を併せて解析した結果、600種のうち462種(77%)が同定可能であった。昨年(2008年)の結果では、383種のうち86%が同定可能であった(6)。今後、解析する種数が増えるに従って同定不能の種の割合が多くなると予想される。

2009年11月にCBOL Plant Working Groupが、植物におけるDNAバーコード領域として*rbcL*部分配列と*matK*部分配列の2つの領域を標準とし、それで足りない部分を、*trnH-psbA*遺伝子間領域などで補うことを決定したので、今後は上記2つの領域を中心にDNAバーコーディングを進めていく。*matK*領域は遺伝子領域ではあるが、変異性が高く、*trnH-psbA*遺伝子間領域と同じぐらいの種同定能力があるとされている(1)。しかしながら、*matK*領域は変異性が高く、単一のプライマー組では、増幅できないため、分類群によって別のプライマー組を変えて解析を行うことが推奨されている(2)。このことはDNAバーコードの応用に際し問題が生じる恐れがある。

葉緑体は、被子植物では母親から、裸子植物では父親から遺伝するため、浸透交雑が進んでいる種群では、葉緑体DNAだけでは、同定できない場合が出てくる。その場合、核DNAのITSなどの領域でバーコーディングを進める必要が出てくると考えられる。

IV おわりに

DNA バーコードは、証拠標本と DNA 試料がセットとして保存され、データベース化されることにより、同定等に疑問が生じたときに、いつでも遡って検証し直すことができる。実際、解析を行っている際に、数個体ではあるが、DNA シークエンスにより同定に疑問が生じ、標本を見直した結果、同定が誤っていた例があった。同定

者の熟練度や、サンプルの採取時期、採取部位等の問題で、同定が困難であったり、同定を間違えることは起こりがちであり、植物の分類に携わる研究者が減少している現在の状況下において、DNA バーコードは有効なツールになりうる。

表-1. サンプルの採取年度とシークエンスの得られた数

採取年度	コレクションネーム	DNA 試料数	PCR 成功試料数		シークエンス成功試料数		シークエンス成功試料数 (%)	
			<i>rbcL</i>	<i>trnH-psbA</i>	<i>rbcL</i>	<i>trnH-psbA</i>	<i>rbcL</i> (%)	<i>trnH-psbA</i> (%)
1998	98 西表島	124	103	75	95	69	76.6	55.6
1998	98 静岡	81	76	60	72	58	88.9	71.6
1999	99 北海道	92	90	82	85	75	92.4	81.5
2000	2000 岡山	162	108	152	105	120	64.8	74.1
2001	2001 三重	190	80	55	72	29	37.9	15.3
2001	2001 国頭	152	132	70	93	61	61.2	40.1
2002	2002 奄美	131	127	59	107	52	81.7	39.7
2002	2002 対馬	148	130	69	100	60	67.6	40.5
2002	2002 東北	151	103	92	91	26	60.3	17.2
2003	2003 吾妻	148	145	141	140	132	94.6	89.2
2004	2004 内之浦	115	112	107	112	97	97.4	84.3
2004	2004 宮崎北部	175	161	160	157	155	89.7	88.6
2004	2004 富山	129	105	113	96	62	74.4	48.1
2005	2005 島根	190	169	173	160	110	84.2	57.9
2005	2005 岐阜	169	152	160	140	140	82.8	82.8
2006	2006 北薩	188	187	184	183	179	97.3	95.2
2006	2006 芦生	31	30	30	30	28	96.8	90.3
2007	2007 徳之島	97	96	95	95	92	97.9	94.8
2007	2007 四万十	155	150	149	149	146	96.1	94.2
2007	2007 岐阜三重	162	157	154	155	152	95.7	93.8
2008	2008 佐賀	145	145	135	143	135	98.6	93.1
2008	2008 木曾	215	208	199	203	182	94.4	84.7
2008	2008 滋賀	153	150	144	149	125	97.4	81.7
2009	2009 茨城	184	180	177	178	154	96.7	83.7
2009	2009 岩手	197	191	174	189	142	95.9	72.1
2009	2009 北海道	90	88	81	87	67	96.7	74.4
	その他	490	476	399	438	387	89.4*	79.0*
	総計	4264	3851	3489	3624	3035	85.0	71.2

*:一部のサンプルがシークエンスされていないため、シークエンス成功率が若干過小になっている。

配列領域	供試個体数	PCR 増幅個体数	(%)	シーケンス	(%)
<i>rbcL</i>	4252	3840	90.3	3624	85.2
<i>trnH-psbA</i>	4149	3489	84.1	3035	73.2
	供試種数				
<i>rbcL</i>	721	697	96.7	675	93.6
<i>trnH-psbA</i>	707	657	92.9	613	86.7

	種同定可能	種同定不能	同定率(%)
<i>rbcL</i> + <i>trnH-psbA</i>	462	138	77.00
<i>trnH-psbA</i>	454	159	74.06
<i>rbcL</i>	347	328	51.41

引用文献

- (1) CBOL PLANT WORKING GROUP (2009) A DNA barcode for land plants. PNAS **106**:12794-12797.
- (2) DUNNING, L.T., SAVOLAINEN, V. (2010) Broad-scale amplification of matK for DNA barcoding plants, a technical note. Bot. J. Linn. Soc. **164**:1-9.
- (3) FORD, C.S., AYRES, K.L., TOOMEY, N., HAIDER, N., STAHL, J.V.A., KELLY, L.J., WIKSTROM, N., HOLLINGSWORTH, P.M., DUFF, R.J., HOOT, S.B., COWAN, O.S., CHASE, M. W., WILKINSON, M.J.(2009) Selection of candidate coding DNA barcoding regions for use on land plants. Bot. J. Linn. Soc. **159**:1-11.
- (4) HEBERT, P.D.N., CYWINSKA, A., BALL, S.L., DEWAARD, J.R. (2003) Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B. Biol. Sci. **270**: 313-321.
- (5) HOLLINGSWORTH, P.M.(2008) DNA barcoding plants in biodiversity hot spots: Progress and outstanding questions. Heredity **101**:1-2.
- (6) 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. **2**: e508.
- (7) 吉村研介・大谷雅人・田中孝尚・鈴木三男・神保宇嗣・伊藤元己・永居寿子・館田英典・津村義彦・藤井智之・能城修一・河原孝行・吉丸博志(2009) 日本産樹木DNAバーコーディングの現状 (2008) :関東森林研究 **61** : 79-82.