

藻类 DNA 条形码研究进展*

崔翠菊** 张立楠 王 娜 李晓捷 刘延岭 江 鑫

(山东东方海洋科技股份有限公司 国家海藻工程技术研究中心 烟台 264003)

摘要 DNA barcode, 又称为 DNA 条形码, 是指利用短的标准 DNA 序列的核苷酸多样性进行物种的鉴定和快速识别。目前该方法在动物分类研究中应用广泛, 其中线粒体的细胞色素 c 氧化酶亚基 1 (cytochrome c oxidase subunit 1, COI 或 *cox 1*) 基因中的约 700bp 长度的一段被用来作为标准 DNA 片段。在陆地植物条形码研究中, 生命-植物条形码联盟会 (Consortium for the Barcode of Life-Plant Working Group, CBOL-Plant Working Group) 近期推荐将植物叶绿体中的两个基因片段 *rbcL* + *matK* 作为初步的陆生植物条形码, 此组合能在 70% 的程度上进行植物物种的鉴别。在海藻的分类研究中, DNA 条形码的应用较少, 已有的研究主要集中在硅藻、红藻和褐藻, 尚没有学者明确提出适合藻类的 DNA 条形码。总结了能够作为藻类 DNA 条形码的序列特点、应用流程及分析方法, 综述了 DNA 条形码在藻类中的研究现状和存在的问题, 展望了藻类 DNA 条形码的应用前景。

关键词 DNA 条形码 海藻 物种分类 COI

中图分类号 Q914.82

现有的生物物种鉴定和分类方法主要以形态为基础, 地球上有数亿个物种, 迄今为止, 也仅有不足 200 万个物种被认识, 物种鉴定和分类研究是一项任重而道远的任务。目前形态学分类方法有两方面的局限: 第一, 不同个体的形态特征受生物发育阶段和遗传可变的影响容易导致不正确的鉴定, 许多群体中存在的隐存种无法只通过形态观察准确辨别; 第二, 分类学家识别能力有限, 能够正确鉴定 1 000 个物种的分类学家极少, 对地球上 1 亿 ~ 10 亿个物种而言就需要 1 万 ~ 10 万个分类学家来支撑物种的识别体系^[1]。形态学鉴定的局限性和不断缩减的分类学家队伍, 使分类学的发展面临巨大的挑战。

分子生物学的发展和大量的分子数据为上述问题提供了一个解决方法。已有的研究表明, 通过分析特定基因片段能区别生物的多样性差异。DNA 条码(技术)的概念于 1993 年首次提出, 它是利用标准的 DNA 片段作为快速和准确进行物种鉴定和分类的标签, 但

在当时并没有引起学者们的广泛关注^[2]。直到 2003 年 Hebert 等^[3]将其应用到动物的物种鉴定研究, DNA 条形码技术才逐渐走入大家的视线。动物的 DNA 条形码研究通常使用线粒体的 *COI* 基因, 基因片段长度为 648bp, 侧翼有保守序列, 使其很容易扩增、测序和分析。更多的研究证明, *COI* 基因在种内的变异率仅为 1% ~ 2%, 但在相近物种中的变异率相对较高, 因此, 在用其进行物种鉴定时可信度高^[4]。在陆生植物中, Vijayan、CBOL Plant Working Group、闫化学、唐建阳等分别对植物 DNA 条形码研究进行了综述和应用前景预测^[5-8], 研究发现, *COI* 基因的替代率很低, 而且基因含量和结构改变很大, 不适合作为植物的 DNA 条形码^[5], 而 CBOL Plant Working Group 推荐将 *rbcL* + *matK* 两个基因片段的组合作为植物的 DNA 条形码^[6]。此外, DNA 条形码技术已在病毒、细菌和原生动物类群以及司法鉴定, 生物工程, 食品工业等领域得到应用^[9]。藻类的 DNA 条码研究起步较晚, 研究系统性不强, 通用性差, 发展速度缓慢。本文综述了关于藻类 DNA 条形码研究的标准和意义, 探讨了目前国内外的研究现状和存在问题, 希望能为相关研究的学者提供一些有用的信息。

收稿日期: 2012-01-31 修回日期: 2012-06-30

* 国家“863”计划资助项目(2012AA10A406)

** 电子信箱: cuicuiju@163.com

1 藻类 DNA 条形码概念、标准及分析方法

藻类的 DNA 条形码就是指用一个或几个标准的 DNA 片段进行藻类的物种鉴定和识别,不受个体形态、生长阶段的限制。若能建立藻类的 DNA 条形码数据库,便可以加快海洋藻类的物种鉴定速度和认识新的物种,还可以克服形态学鉴定的局限性和准确率不高的问题,减少对藻类分类的物力和人力要求,促进藻类分类研究进展。条形码是 DNA 序列,应用其进行物种鉴定时,与微卫星 DNA 标记、随机扩增多态性 DNA 等技术相比较,重复性高,省时省力。

藻类 DNA 条形码的几个重要特点:(1)要求不同种间有明显的遗传变异和分化,同时种内变异不大,不足以影响对种间的变异判断;(2)通用性好,存着于较保守的区域,便于设计在不同种间均可应用的通用引物;(3)长短适中,最好为 500 ~ 1 000bp,太短不能很好的反应遗传变异,太长不利于 PCR 扩增和测序反应。

DNA 条形码的应用流程:对所研究的样品进行采集、清洗和处理藻体表面的杂藻和菌类,提取基因组 DNA;选择所用的 DNA 条形码基因序列,设计和合成引物;对样品 DNA 进行 PCR 扩增;对 PCR 产物进行纯化处理和测序分析;对测序得到的结果进行序列编辑,并对结果进行分析。

首先进行序列比对和人工校正,再通过 MEGA 或 PAUP 计算种内和种间的 K2P 距离,比较种、属和科 3 个水平上的序列差异,根据结果建立 NJ 树,依据 DNA 条形码基因得出的遗传距离对待检样品进行分类鉴定^[10]。这种方法相对简单,对条形码数据库相对成熟的动物条形码的分析比较适用,但对正在进行条形码评估阶段的植物条形码和藻类条形码而言,还需要更进一步的分析,主要分以下三步:(1)测序得到的序列进行人工校正和序列比对,用 BLAST 方法在 GenBank 的数据库中查找同源性高的基因片段,进行校正;(2)进行种内和种间距离的计算, Saunders 等^[11]使用 PAUP 4.0b10 计算种间和种内的遗传距离,采用 General Time-Reversible model (GTR) 模型进行检验^[12];(3)系统学分析,通常采用 NJ、MEGA、ML、MP 等方法建立多种系统发育树,检验物种的聚类统一性。Lahaye 等^[10]对不同的方法进行了比较,发现 UPGMA 和 MP 树的物种识别正确率较高,可以对不同的条形码进行评估和验证。但因为不同的软件在进行数据处理时的复杂程度不同,使用条件不同,所用时间也不同,

所以可根据具体需要进行选择,以便可以达到条形码快速鉴别物种的目的。

2 藻类 DNA 条形码的研究现状

虽然 DNA 条形码技术在动物和植物中的研究颇多,对动植物的分类鉴定等相关研究也有一定的帮助。但在藻类的研究中,至今为止,还未有学者或组织明确的提出通用的藻类 DNA 条形码。

Evans 等^[13]选取了 22 个硅藻属的 34 个个体为样本进行硅藻的条形码研究,分别对 *cox1*、*rbcL*、*18S rDNA*、*ITSrDNA* 4 个基因序列的有效性进行了评价,结果表明 *cox1* 基因的种内变异为 0 ~ 5bp,种间至少为 18bp,比 *rbcL* 和 *18srDNA* 的变异率都高,*ITSrDNA* 的种内变异率太高,而且现有数据库中的硅藻的 *ITSrDNA* 序列信息很少,综合比较,*cox1* 基因更适合作为硅藻的 DNA 条形码。McDevit 等^[14]对褐藻纲的 5 目、9 科、20 属、29 种的 106 个个体进行了条形码的适用性研究,发现 *cox1* 能够明确的进行种间的区分,种内的变异率为 0.00% ~ 0.46%。Clarkston 等^[15]利用线粒体的细胞色素氧化酶 I 基因的 5' 端序列 (COI-5P) 和质体的 *23SrRNA* 基因 (UPA) 作为 DNA 条形码在红藻目、嵴膜藻科的 102 个个体中进行比较评价发现,COI-5P 可以很好的进行种间的区分,但是很多个体不能扩增到清晰的产物,该问题可以通过使用特异引物解决。UPA 虽然在红藻、褐藻和绿藻中引物的通用性强,但种间序列变异小。该研究通过 *COI-5P* 基因的系统学分析还发现了嵴膜藻科的一个新种。McDevit 等^[16]对在海带科中扩增得到的 194 条 COI-5P 序列进行了分析,确定在加拿大的海域有 12 个种。Lane 等^[17]利用 DNA 条形码对太平洋东北沿岸的褐藻纲、海带目、翅藻属进行了分子评价,他利用线粒体的 *COI-5P* 基因对翅藻目的群体进行了系统进化分析,并比较了细胞核中的基因 *ITSrDNA* 和质体中的基因 *rbcSp* 作为条形码辅助研究了优劣性。Saunders 等^[11]对在加拿大海区收集到的江篱目的海藻进行条形码研究时发现了 *COI5'* 基因的一个新序列,后经形态和解剖学辅助研究发现该物种是一个入侵物种。Saunders 等^[18]对加拿大海区的白蒙藻科的藻类进行条形码 COI 5' 的评价研究,发现了 *Weeksia* 属的 3 个新物种。Uwai 等^[19]用线粒体基因 *cox3* 和核基因组中的 *ITS1* 对日本周围海域 21 个不同位点采集到的裙带菜进行了种内遗传多样性分析,发现 *cox3* 在 106 个植株中有 9 个单倍型,可以对来自不同海域的样

本进行区分。

目前藻类研究中使用较多的 DNA 条形码主要是 *COI* 基因或其 5' 端。但因地域的差异,对物种的收集形成了障碍,目前研究者主要在就近海域取材进行鉴定。若验证 DNA 条形码的通用性,需考虑世界范围内不同

海区的同科或同属物种的适用性。在不能亲自采集到样品时,只能下载其他研究者发布在数据库中的相关序列信息进行分析,所以在时间和地域上也制约了藻类 DNA 条形码的快速发展。表 1 总结了目前应用在藻类 DNA 条形码研究中的一些基因及其引物序列和来源。

表 1 藻类中用到的 DNA 条形码引物

Table 1 Primers used for algae DNA barcoding analysis by research groups

Gene	Primer	Sequence 5'-3'	Reference
<i>COI</i>	GazF1f	TCAACAAATCATAAAGATATTGG	[20]
	GazR1r	ACTTCTGGATGTCCAAAAAYCA	[20]
	GazF2f	CAACCAYAAAGATATWG GTAC	[13]
	GazR2r	GGATGACCAARAACCAAAA	[13]
	KEintFf	GAGAGCAAAAAGTTTACCATTTC A	[13]
	KEintRr	CAAATAAAATTRATWGCWCCTAA	[13]
	KEint2Ff	GAAGCWGGWGTWGGTACWGGWTG	[13]
	KEdtmRr	AAACTTCWGGRTGACCAAAAA	[13]
	GWSFf	TCCCAGTCACGACGTCGTTC AACAAYCAYAAAGATATYGG	[11]
	GWSF5f	ACAAAYCAYAAIGATATYGG	[11]
	GWSRr	GGAAACAGCTATGACCATGGGRTGTCCRAARAAYCARAA	[11]
	GWSR3r	GGAAACAGCTATGACCATGGGRTGTCCAAAIAAYCARAA	[11]
	GWAR5r	TCAGGRTGNCCIAARAAYCA	[11]
	Cox1R1r	GTATACATATGATGHGCTCAA	[18]
	GHalFf	TCAACAAATCATAAAGATATYGG	[18]
	GHalRr	CTTCWGGATGRCCAAAAATCA	[21]
	DiamF3f	CCAACCAYAAAGATATWGGWAC	[16]
	DumR1r	AAAAAYCARAATAAATGTTGA	[18]
	GazR4r	AAAATCAAAATAAATGCTG	[18]
<i>ITS</i>	ITS1f	GGAAGGAGAAGTCGTAACAAGG	[22]
	ITSR1r	TTCAAAGATTTCGATGATTAC	[22]
	P5f	GCATCGATGAAGAACGCAG	[22]
	G4r	CTTTTCCTCCGCTTATTGATATG	[22]
<i>UPA</i>	P23SrV_flf	GGACAGAAAACCTATGAA	[23]
	P23SnewRr	TCAGCCTGTTATCCCTAGA	[21]

Note: f. forward primer; r. reverse primer.

3 藻类 DNA 条形码研究存在的问题及对策

目前在藻类应用较多的是 *COI*、*ITS*、*rbcL* 基因,在不同的物种中各有优势。很难找到像动物的 *COI* 序列可以在物种鉴定和新物种发现中广泛通用的条形码。动物和陆地植物的线粒体和叶绿体是母性遗传。在藻类中,母性遗传也占优势,但研究发现也存在同配生殖现象^[24]。褐藻目的萱藻就有同配生殖发生,其线粒体属于单性遗传,但叶绿体却是双性遗传^[25]。绿藻中的莱茵衣藻也有单配生殖,但其叶绿体和线粒体来自不同交配型的单亲^[26]。如果选择的 DNA 条形码基因内部含有内含子或被分成多个部分,就存在扩增和测序的问题。在硅藻 *T. pseudonana* 和 *T. nordenskioeldii* 还有几种绿藻的线粒体中均发现有内含子^[27-28]。

要解决上述问题,需要加大、加快藻类的测序规模和速度,特别是同属内不同近缘种的序列测定,在大规模数据的基础上进行分析,选择更合适的条形码序列或组合,同时也可以参考动物或植物中的方法,借鉴形态分类中涉及的门、纲、目、科属等筛选出不同进化级别上的藻类 DNA 条形码。只从形态特征对藻类进行精确鉴定很困难。分子辅助分类方法的出现很大程度上解决了这个问题。对很多不同种类的个体的同一段标准 DNA 标记进行测序,进行遗传聚类分析,然后再用形态学观察进行鉴定验证。形态和分子方法相结合是进行生物物种鉴定的一种发展趋势。

4 藻类 DNA 条形码应用前景

目前藻类的 DNA 条形码研究还不成熟,在数据库

中的序列信息可能存在序列错误、样品 DNA 污染或分类命名错误等问题,很多物种的序列还没有被及时上传到数据库,致使分析结果不可靠或找不到可参考的数据。早在 2003 年的“Taxonomy and DNA”会议上,研究者就提出要对全球的所有物种的 *COI* 基因进行测序^[4]。由加拿大国际条形码协会发起的“国际 DNA 条形码工程”计划用未来 5 年收集 10 万个物种的条形码序列,为每个物种注册登记各处的 DNA 条形码信息。而且国际上也成立了诸多关于 DNA 条形码的网站和组织,虽然现在涉及藻类的 DNA 条形码的组织还较少,但相信随着物种的发现和条形码研究的深入,会有越来越多的藻类的 DNA 条形码信息被录入数据库供研究者参考使用。藻类 DNA 条形码的研究也将会成为藻类物种分类鉴定和系统分化研究中的一种发展趋势,我们可以参照动物和植物的 DNA 条形码发展历程,逐步深入和广泛的研究藻类 DNA 条形码,还可用于对外来入侵物种的评估及外来物种的识别等研究^[29]。虽然藻类 DNA 条形码的研究工作进展缓慢,并存在一些问题,但随着测序技术的广泛应用,藻类的 DNA 条形码的研究也将日趋完善,为藻类的分类鉴定和新物种的发现、生物多样性的评价等研究提供准确、快速的信息,加快藻类学研究的进展。

参考文献

- [1] 肖金花, 肖晖, 黄大卫. 生物分类学的新动向-DNA 条形码. 动物学报, 2004, 50(5): 852-855.
- Xiao J H, Xiao H, Huang D W. DNA barcoding: new approach of biological taxonomy. Acta Zoologica Sinica, 2004, 50(5), 852-855.
- [2] Arnot D E, Roper C, Bayoumi R A L. Digital codes from *Hypervariable tandemly repeated DNA sequences in the Plasmodium falciparum circumsporozoite gene* can genetically barcode isolates. Molecular and biochemical parasitology, 1993, 61(1): 15-24.
- [3] Hebert P D N, Ratnasingham S, de Waard J R. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society of London. Series B: Biological Sciences, 2003, 270(Suppl 1): S96-S99.
- [4] Hebert P D N, Cywinska A, Ball S L, et al. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences, 2003, 270(1512): 313-321.
- [5] Vijayan K, Tsou C. DNA barcoding in plants: taxonomy in a new perspective. Current Science, 2010, 99(11): 1530-1541.
- [6] Hollingsworth P M, Forrest L L, Spouge J L, et al. A DNA barcode for land plants. Proceedings of the National Academy of Sciences, 2009, 106(31): 12794-12797.
- [7] 闫化学, 于杰. DNA 条形码技术在植物中的研究现状. 植物学报, 2010, 45(1): 102-108.
- Yan H X, Yu J. Current status of the study of DNA barcoding in plants. Chinese Bulletin of Botany, 2010, 45(1), 102-108.
- [8] 唐建阳, 周先治. 植物 DNA 条形码研究现状及应用前景. 中国农学通报, 2009. 25(24): 35-43.
- Tang J Y, Zhou X Z. Progress and outstanding of DNA barcoding in plant. Chinese Agriculture Science Bulletin, 2009, 25, 35-43.
- [9] Chase M W, Salamin N, Wilkinson M, et al. Land plants and DNA barcodes: short-term and long-term goals. Philosophical Transactions of the Royal Society B: Biological Sciences, 2005, 360(1462): 1889-1895.
- [10] Lahaye R, Van Der Bank M, Bogarin D, et al. DNA barcoding the floras of biodiversity hotspots. Proceedings of the National Academy of Sciences, 2008, 105(8): 2923-2928.
- [11] Saunders G W. Routine DNA barcoding of Canadian *Gracilariales* (Rhodophyta) reveals the invasive species *Gracilaria vermiculophylla* in British Columbia. Molecular Ecology Resources, 2009, 9: 140-150.
- [12] Swofford D L. {PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4.}. Sinauer&Associates, Sunderland, Massachusetts, 2003.
- [13] Evans K M, Wortley A H, Mann D G. An assessment of potential diatom “Barcode” Genes (*coxI*, *rbcL*, *18S* and *ITS rDNA*) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). Protist, 2007, 158(3): 349-364.
- [14] McDevit D C, Saunders G W. On the utility of DNA barcoding for species differentiation among brown macroalgae (Phaeophyceae) including a novel extraction protocol. Phycological Research, 2009, 57(2): 131-141.
- [15] Clarkston B E, Saunders G W. A comparison of two DNA barcode markers for species discrimination in the red algal family Kallymeniaceae (Gigartinales, Florideophyceae), with a description of *Euthora timburtonii* sp. nov. Botany, 2010, 88(2): 119-131.
- [16] McDevit D C, Saunders G W. A DNA barcode examination of the Laminariaceae (Phaeophyceae) in Canada reveals novel biogeographical and evolutionary insights. Phycologia, 2010, 49(3): 235-248.
- [17] Lane C E, Lindstrom S C, Saunders G W. A molecular assessment of northeast Pacific *Alaria* species (Laminariales, Phaeophyceae) with reference to the utility of DNA barcoding. Molecular Phylogenetics and Evolution, 2007, 44(2): 634-648.
- [18] Saunders G W. A DNA barcode examination of the red algal

- family Dumontiaceae in Canadian waters reveals substantial cryptic species diversity. 1. The foliose Dilsea-Neodilsea complex and Weeksia. Botany, 2008, 86(7): 773-789.
- [19] Uwai S, Yotsukura N, Serisawa Y, et al. Intraspecific genetic diversity of *Undaria pinnatifida* in Japan, based on the mitochondrial *cox3* gene and the ITS1 of nr DNA. Hydrobiologia, 2006, 553:345-356.
- [20] Saunders G W. Applying DNA barcoding to red macroalgae; a preliminary appraisal holds promise for future applications. Philosophical Transactions of the Royal Society B: Biological Sciences, 2005, 360(1462): 1879-1888.
- [21] Saunders G W, McDonald B. DNA barcoding reveals multiple overlooked Australian species of the red algal order Rhodymeniales (Florideophyceae), with resurrection of *Halopeltis* J. Agardh and description of *Pseudohalopeltis* gen. nov. Botany, 2010, 88(7): 639-667.
- [22] Tai V, Lindstrom S C, Saunders G W. Phylogeny of the Dumontiaceae (Gigartinales, Rhodophyta) and associated families based on SSU rDNA and internal transcribed spacer sequence data. Journal of Phycology, 2001, 37(1): 184-196.
- [23] Sherwood A R, Presting G G. Universal Primers Amplify a 23s rDNA Plastid Marker in Eukaryotic Algae and Cyanobacteria. Journal of Phycology, 2007, 43(3): 605-608.
- [24] Coyer J, Peters A, Hoarau G, et al. Inheritance patterns of ITS1, chloroplasts and mitochondria in artificial hybrids of the seaweeds *Fucus serratus* and *F. evanescens* (Phaeophyceae). European Journal of Phycology, 2002, 37(2): 173-178.
- [25] Kato Y, Kogame K, Nagasato C, et al. Inheritance of mitochondrial and chloroplast genomes in the isogamous brown alga *Scytosiphon lomentaria* (Phaeophyceae). Phycological Research, 2006, 54(1): 65-71.
- [26] Aoyama H, Hagiwara Y, Misumi O, et al. Complete elimination of maternal mitochondrial DNA during meiosis resulting in the paternal inheritance of the mitochondrial genome in *Chlamydomonas* species. Protoplasma, 2006, 228(4): 231-242.
- [27] Armbrust E, Berges J A, Bowler C, et al. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science, 2004, 306(5693): 79-86.
- [28] Ehara M, Watanabe K I, Ohama T. Distribution of cognates of group II introns detected in mitochondrial *cox1* genes of a diatom and a haptophyte. Gene, 2000, 256(1-2): 157-167.
- [29] Cross H B, Lowe A J, Gurgel F D. DNA barcoding of invasive species. Fifty Years of Invasion Ecology, 2010, 289-299.

Progress of DNA Barcoding in Algae

CUI Cui-ju ZHANG Li-nan WANG Na LI Xiao-jie LIU Yan-ling JIANG Xin

(Shandong Oriental Ocean Sci-Tech Co. Ltd. National Algae Project Technology Research Centre, Yantai 264003, China)

Abstract DNA barcode technology is a method of rapid and accurate species identification and recognition on the utility of the nucleotides diversity of some short and standardized DNA sequences. At present, this method is widely used in the classification of animals, the mitochondria cytochrome oxidase c subunit 1 (COI or *cox 1*) gene in 700 bp length is being used as a standard DNA fragment. In Plant barcoding study, Consortium for the Barcode of Life-Plant Working Group (CBOL-Plant Working Group) recently recommended *rbcL* + *matK*, two gene fragments in chloroplast genome as preliminary potential candidate for plant barcode, with 70% species discriminatory power. Few application of the DNA barcode is reported in the classification of algae, mainly in the red algae and brown algae. A well-characterized algal locus that meets the barcoding criteria is lacking. The DNA sequence of standard and barcoding application process were reviewed, methods and the advantages were analysed, then discusses the current status, problems and application prospect in algae barcode research.

Key words DNA barcode Alga Species identification COI