文昌鱼线粒体基因组特征分析及分子标记探讨

申　欣[1, 2], 田　美[1], 孟学平[1], 程汉良[1]

(1. 淮海工学院 海洋学院, 江苏 连云港 222005; 2. 中国科学院 海洋研究所, 山东 青岛 266071)

摘要: 线粒体基因组已被广泛应用于后生动物分子系统发育和群体遗传的研究。文昌鱼(Amphioxus)作为研究脊椎动物起源和进化的模式动物, 在脊椎动物起源和进化研究中占据极为重要的位置。作者综合分析文昌鱼2科7个种的51条线粒体基因组全序列, 全面揭示了文昌鱼线粒体基因组的基本特征。文昌鱼线粒体基因组均编码后生动物标准的 37 个基因, 文昌鱼线粒体基因组共有 3 种基因排列方式, 其中文昌鱼属和侧殖文昌鱼属共有的基因排列与脊椎动物线粒体基因组的典型排列相比最为接近。因此, 文昌鱼属和侧殖文昌鱼属线粒体基因组的基因排列方式代表了文昌鱼原始的基因排列方式。与此相比, 偏文昌鱼属的 3 个物种发生了基因重排。文昌鱼线粒体基因组 13 个蛋白编码基因的 *Ka*/*Ks* 值均远远低于 1(0~0.3363), 显示出较强的纯化(负)选择。文昌鱼线粒体基因组种间和种内基因变异分析表明, 在文昌鱼群体遗传的研究中, *nad5*、*nad4* 和 *nad2* 基因是理想的分子标记, 可以作为 *cox1* 基因辅助的分子标记, 用于分析文昌鱼不同群体之间的遗传多样性, 为其生物多样性的保护及合理利用其生物资源提供更多基础资料。

关键词: 文昌鱼(amphioxus); 线粒体基因组; 差异位点; 分子标记

中图分类号: Q959.287　　**文献标识码**: A　　**文章编号**: 1000-3096(2011)07-0007-07

线粒体基因组作为核外遗传物质具有结构简单、基因排列紧凑和严格的母系遗传等特点, 在过去的十几年中已经被广泛应用于后生动物分子系统发育和群体遗传的研究[1-4]。文昌鱼(Amphioxus 或者 Lancelet)隶属于脊索动物门(Chordata)头索动物亚门(Cephalochordata), 一直是研究脊椎动物起源和进化的模式动物, 在脊椎动物起源和进化研究中占据极为重要的位置, 有重要的科研价值和一定的经济价值[5-7]。在 GenBank 线粒体基因组的数据库中, 目前有 54 条文昌鱼的线粒体基因组序列。经相似性比较后, 剔除了在物种鉴定上可能存在争议的 3 条线粒体基因组序列。作者在 51 条文昌鱼线粒体基因组序列(分别代表了文昌鱼 2 科 3 属 7 种, 图 1)研究的基础上, 全面分析了文昌鱼线粒体基因组的基本特征、基因排列、蛋白质编码基因和差异位点等, 对于文昌鱼的生物多样性的保护及其生物资源合理利用等诸多方面提供遗传信息, 为寻找合适的分子标记提供参考。

1　材料和方法

1.1　数据获取

从 GenBank 线粒体基因组数据库中检索、下载得到 54 条文昌鱼线粒体基因组全序列。经 BLAST 和 Clustal[8]相似性比较后, 3 条线粒体基因组(NC_001912 矛形文昌鱼(*Branchiostoma lanceolatum*)、NC_008069 日本文昌鱼(*B. japonicum*)和 AY932825 白氏文昌鱼(*B. belcheri*))在物种鉴定上可能会存在争议(文昌鱼物种鉴定上可能会出现错误, 在以前的文献中也有报道[9]), 剔除后剩余的 51 条线粒体基因组序列, 分别是隶属于文昌鱼科(Branchiostomidae)的白氏文昌鱼、佛罗里达文昌鱼(*B. floridae*)和矛形文昌鱼; 以及隶属于侧殖文昌鱼科(Epigonichthyidae)的马尔代夫侧殖文昌鱼(*Epigonichthys maldivensis*)、尖刀偏文昌鱼(*Asymmetron lucayanum*)、偏文昌鱼(*A. inferum*)和偏文昌鱼 *A.* sp. *A TK-2007*[9-13](图 1)。其中白氏文昌鱼、佛罗里达文昌鱼、矛形文昌鱼和尖刀偏文昌鱼分别有 14 条、21 条、11 条和 2 条线粒体基因组全序列(表 1)。

收稿日期: 2010-04-20; 修回日期: 2010-11-18

基金项目: 国家自然科学基金资助项目(40906067); 江苏省青蓝工程人才基金资助项目(苏教师[2010]27 号); 江苏省海洋生物技术重点建设实验室研究基金资助项目(2009HS13); 淮海工学院自然科学基金资助项目(Z2009048)

作者简介: 申欣(1981-), 男, 讲师, 博士, 主要从事海洋动物基因组学和蛋白质组学研究, E-mail: shenthin@163.com

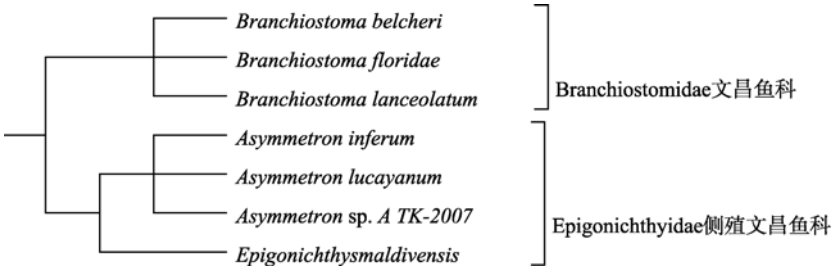

图 1　目前已测定线粒体基因组全序列的文昌鱼物种

Fig. 1　Amphioxus species with complete determination of mitochondrial genome

表 1　文昌鱼线粒体基因组的基本特征

Tab. 1　Basic features of amphioxus mitochondrial genomes

物种名称	线粒体基因组数目	线粒体基因组长度(bp)	基因数目	A+T 含量(%)
白氏文昌鱼	14	15074~15076	37	64.7~64.8
佛罗里达文昌鱼	21	15075~15083	37	62.6~63.1
矛形文昌鱼	11	15139~15146	37	61.7~61.8
偏文昌鱼	1	15084	37	62.7
尖刀偏文昌鱼	2	15100~15102	37	61.5
偏文昌鱼	1	15050	37	59.1
马尔代夫侧殖文昌鱼	1	14975	37	61.1

1.2　同义替换率(*Ks*)和非同义替换率(*Ka*)分析

进化速率是受稳定性(负)选择、突变和定向(正)选择控制的，为了检验选择压力对于文昌鱼(以白氏文昌鱼、佛罗里达文昌鱼、矛形文昌鱼和尖刀偏文昌鱼为代表)线粒体基因组的影响，线粒体基因组中 13 个蛋白质编码基因的核酸序列通过 Clustal X[8]进行多重序列比对。同义替换率(*Ks*)和非同义替换率(*Ka*)通过 PAML[14]和 DnaSP 4.10.7[15]进行了估算。

1.3　基因差异性分析

以 7 种文昌鱼线粒体基因组数据作为数据群进行了种间基因差异位点和基因变异分析，同时以 14 条白氏文昌鱼、21 条佛罗里达文昌鱼、11 条矛形文昌鱼和 2 条尖刀偏文昌鱼作为数据群进行了种内基因差异位点和基因变异分析。13 个蛋白质编码基因和 2 个核糖体 RNA 基因的核酸序列通过 Clustal X[8]进行多重序列比对，然后通过 DnaSP 4.10.7[15]分析了线粒体基因组主编码基因的差异位点和基因变异特征。

2　结果和讨论

2.1　文昌鱼线粒体基因组的特征

文昌鱼 7 个物种线粒体基因组长度介于 14 975 bp(马尔代夫侧殖文昌鱼)和 15 146 bp(矛形文昌鱼)之间。7 个物种的线粒体基因组均编码后生动物标准的 37 个基因(13 个蛋白质编码基因、22 个转运 RNA 基因和 2 个核糖体 RNA 基因)(表 1，图 2)。线粒体基因组主编码链的 A＋T 含量最高的为 64.8%(白氏文昌鱼)，最低的为 59.1%(偏文昌鱼 *A.* sp. *A TK-2007*)。文昌鱼 7 个物种线粒体基因组的基本特征见表 1。

2.2　文昌鱼线粒体基因组的基因排列

在文昌鱼 7 个物种的线粒体基因组中，白氏文昌鱼、佛罗里达文昌鱼、矛形文昌鱼和马尔代夫侧殖文昌鱼 4 个物种的基因排列完全相同；尖刀偏文昌鱼和偏文昌鱼 *A TK-2007* 的基因排列相同；而偏文昌鱼线粒体基因组的基因排列独具一格。因此，从目前的线粒体基因组数据分析，文昌鱼线粒体基因组共有 3 种基因排列方式(图 2)。在 3 种基因排列方式中，第一种基因排列方式(文昌鱼属和侧殖文昌鱼)与脊椎动物线粒体基因组的典型排列相比最为接近，仅有 4 个转运 RNA 基因的重排[10]。因此，文昌鱼属和侧殖文昌鱼线粒体基因组的基因排列方式代表了文昌鱼原始的基因排列方式[11]。与此相比，尖刀偏文昌鱼和偏文昌鱼 *A TK-20072* 的线粒体基因组发生 3 个基因或者基因区块(*cox3-nad3*，*W-A-C-Y* 和 *Q*)的易位以及 1 个基因区块(*L(UAG)-nad5-G-nad6*)的倒位；

偏文昌鱼 *A. inferum* 线粒体基因组发生 1 个基因区块(*cox3-nad3*)的易位、1 个基因区块(*L(UAG)-nad5-G-nad6*)的倒位和 1 个转运 RNA(*Q*)的易位并倒位。偏文昌鱼属 3 个物种的线粒体基因组，共有的“*cox3-nad3*”的易位和“*L(UAG)-nad5-G-nad6*”的倒位，可以作为偏文昌鱼线粒体基因组的共同衍征。

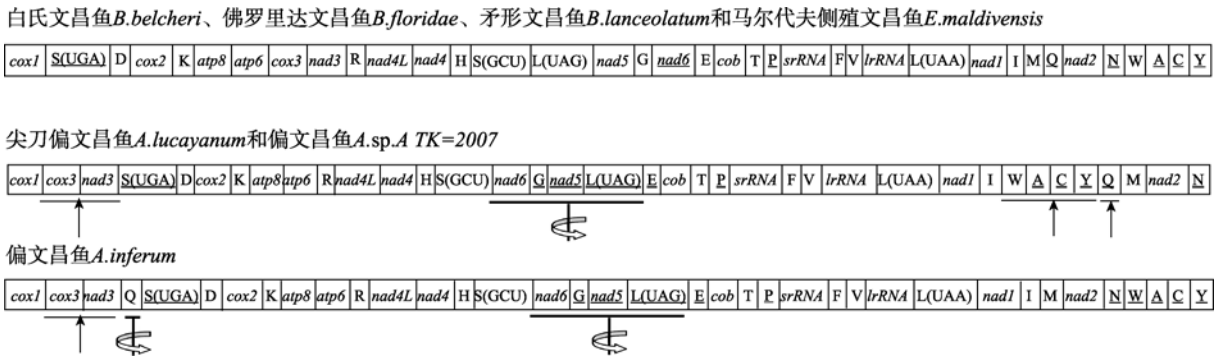

图 2　文昌鱼线粒体基因组图谱

Fig. 2　Gene map of amphioxus mitochondrial genomes

直线箭头表示发生重排的基因或者基因区块，弧形箭头表示基因发生倒置，下划线标示的基因在轻链上编码

Arrows show the rearranged genes or gene blocks. The circling arrows indicate inversions. Genes encoded in the light strand are underlined

2.3　文昌鱼线粒体基因组的蛋白质编码基因

在文昌鱼线粒体基因组的蛋白质编码基因中，对于 *atp6*、*cob*、*cox2*、*cox3*、*nad2*、*nad3*、*nad4* 和 *nad6* 基因，7 个文昌鱼线粒体基因组均以 ATG 作为起始密码子；对于 *atp8*、*cox1*、*nad1*、*nad4L* 和 *nad5* 基因，存在 GTG 和 ATG 两种起始密码子。在文昌鱼线粒体基因组中，除了 *atp8*、*cox1*、*nad2* 和 *nad6* 基因的所有终止密码子为完全终止密码子外(TAA 或者 TAG)，余下的 9 个基因(*atp6*、*cob*、*cox2*、*cox3*、*nad1*、*nad3*、*nad4*、*nad4L* 和 *nad5*)均存在不完全的终止密码子(TA-或者 T-)；而在 7 个文昌鱼线粒体基因组中，*cox2* 基因的终止子全部为不完全终止子“T-”(表 2)。文昌鱼线粒体基因组中的 7 个蛋白质编码基因(*cob*、*cox1*、*cox2*、*cox3*、*nad3*、*nad4* 和 *nad4L*)编码氨基酸数目相同，分别为 380、515、230、262、117、452 和 91 个；其余 6 个蛋白质编码基因编码的氨基酸数目稍有差异(表 2)。文昌鱼 7 个线粒体基因组蛋白质编码基因所编码氨基酸的数目及起始、终止密码子见表 2。

2.4　同义替代和非同义替代

为了分析文昌鱼线粒体基因组蛋白编码基因的选择压力，统计了白氏文昌鱼、佛罗里达文昌鱼、矛形文昌鱼和尖刀偏文昌鱼的 13 个蛋白编码基因的同义替换率(*Ks*)和非同义替换率(*Ka*)。4 个文昌鱼线粒体基因组的 13 个蛋白质编码基因的 *Ka*/*Ks* 比值都远远低于 1(介于 0 和 0.3363 之间)，显示出较强的纯化(负)选择(图 3)。其中，*nad4L* 基因的 *Ka*/*Ks* 均值最低(0.0019)，其次为 *cox1* 和 *cox2* 基因，其 *Ka*/*Ks* 均值分别为 0.0103 和 0.0160，说明这些基因承受较强的选择压力和功能束缚；而 *nad4* 基因的 *Ka*/*Ks* 均值最高(0.1653)，其次为 *atp8* 和 *nad2* 基因，其 *Ka*/*Ks* 均值分别达到了 0.1237 和 0.0933，说明这些基因的选择压力较弱。

2.5　文昌鱼线粒体基因组的差异位点分析

在种群遗传的研究中，分子标记的选择是至关重要的，以 7 种文昌鱼线粒体基因组数据作为数据群进行了种间基因差异位点分析，同时以 14 条白氏文昌鱼、21 条佛罗里达文昌鱼、11 条矛形文昌鱼和 2 条尖刀偏文昌鱼作为数据群进行了种内基因差异位点分析。7 种文昌鱼线粒体基因组 13 个蛋白质编码基因和 2 个核糖体 RNA 基因的差异位点分析详见表 3。*cox1* 基因通常被用于 barcoding 分析和海洋生物类群的群体遗传研究[16]。然而，从分析的结果可以看出，在 15 个主编码基因中，*cox1* 基因最为保守，差异位点的比例为 37.86%，这在群体遗传研究中可能是不够的。6 个基因(*cob*、*cox2*、*cox3*、*nad1*、*srRNA* 和 *lrRNA*)差异位点的比例介于 40%~50%；4 个基因(*atp6*、*nad3*、*nad4L* 和 *nad5*)差异位点的比例介于 50%~60%；4 个基因(*atp8*、*nad2*、*nad4* 和 *nad6*)差异位点的比例介于 60%~70%(表 3)。差异位点数最多的

基因为 *nad5* 基因(1025 个)，其次为 *nad4* 和 *nad2* 基因，差异位点数分别达到 814 和 646 个。因此，在文昌鱼群体遗传的研究中，*nad5*、*nad4* 和 *nad2* 基因可以作为 *cox1* 基因辅助的分子标记。

表 2　文昌鱼线粒体基因组 13 个蛋白质编码基因的氨基酸长度及起始、终止密码子

Tab. 2　Amino acids number, initiation and termination codons in the 13 protein-coding genes of amphioxus mitochondrial genomes

物种名称	接收号		*atp6*	*atp8*	*cob*	*cox1*	*cox2*	*cox3*	*nad1*	*nad2*	*nad3*	*nad4*	*nad4L*	*nad5*	*nad6*
白氏文昌鱼	NC_004537	NO.	227	54	380	515	230	262	314	346	117	452	91	598	167
		I codon	ATG	ATG	ATG	ATG	ATG	ATG	GTG	ATG	ATG	ATG	GTG	ATG	ATG
		T codon	TAA	TAG	T-	TAG	T-	TAA	TAG	TAG	TAA	TAG	TAA	TAA	TAG
佛罗里达文昌鱼	NC_000834	NO.	227	54	380	515	230	262	313	346	117	452	91	599	167
		I codon	ATG	ATG	ATG	GTG	ATG	ATG	ATA	ATG	ATG	ATG	ATG	ATG	ATG
		T codon	TA-	TAG	TAG	TAG	T-	TAA	T-	TAG	T-	TA-	TA-	TAG	TAG
矛形文昌鱼	AB194383	NO.	227	54	380	515	230	262	314	346	117	452	91	598	167
		I codon	ATG	ATG	ATG	GTG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG
		T codon	TA-	TAG	TAG	TAG	T-	TAA	T-	TAG	TAG	TA-	TA-	TAA	TAG
偏文昌鱼	NC_009774	NO.	228	57	380	515	230	262	314	346	117	452	91	597	166
		I codon	ATG	GTG	ATG	GTG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	GTG	ATG
		T codon	T-	TAA	TAA	TAA	T-	TA-	T-	TAA	TAA	TAA	TA-	T-	TAA
尖刀偏文昌鱼	NC_006464	NO.	228	57	380	515	230	262	314	346	117	452	91	596	167
		I codon	ATG	GTG	ATG	GTG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	GTG	ATG
		T codon	T-	TAA	TAG	TAA	T-	TAA	T-	TAA	TAG	TAA	TA-	T-	TAA
偏文昌鱼	NC_009775	NO.	228	57	380	515	230	262	314	346	117	452	91	596	167
		I codon	ATG	GTG	ATG	GTG	ATG	ATG	GTG	ATG	ATG	ATG	ATG	ATG	ATG
		T codon	T-	TAA	TAA	TAA	T-	TAG	T-	TAA	TAG	TAA	TA-	T-	TAA
马尔代夫侧殖文昌鱼	NC_006465	NO.	227	54	380	515	230	262	312	344	117	452	91	597	168
		I codon	ATG	ATG	ATG	GTG	ATG	ATG	GTG	ATG	ATG	ATG	ATG	ATG	ATG
		T codon	TA-	TAG	T-	TAG	T-	TA-	T-	TAA	T-	TAA	TA-	TAA	TAA

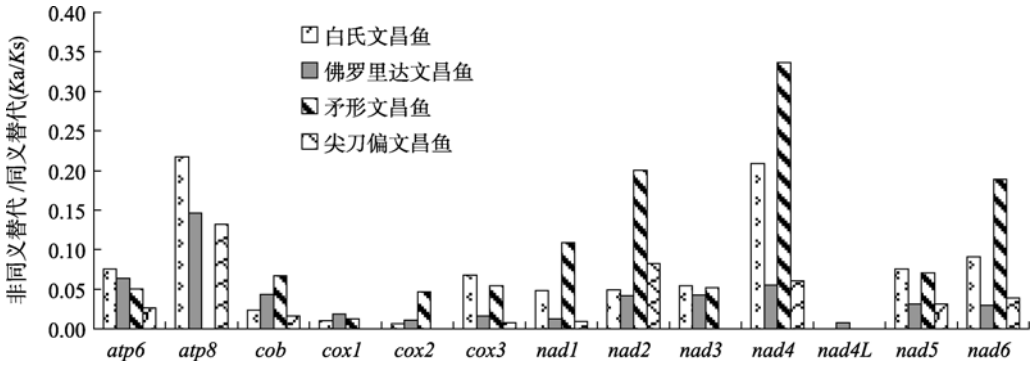

图 3　白氏文昌鱼、佛罗里达文昌鱼、矛形文昌鱼和尖刀偏文昌鱼线粒体基因组 13 个蛋白质编码基因的 *Ka*/*Ks* 分析

Fig. 3　The ratios of Nonsynonymous and synonymous substitutions rates (*Ka*/*Ks*) were estimated in all 13 protein-coding genes of *B. belcheri*, *B. floridae*, *B. lanceolatum*, and *A. lucayanum* mitochondrial genomes

从白氏文昌鱼线粒体基因组主编码基因的差异位点分析，可以看出，*srRNA* 和 *lrRNA* 基因最为保守，差异位点的比例分别为 2.13%和 2.24%；差异位点数最多的基因为 *nad5* 基因(90 个)，其次为 *cox1* 和 *nad2* 基因，差异位点数分别为 73 和 69 个(表 4)。因此，在白氏文昌鱼群体遗传的研究中，*nad5* 和 *nad2* 基因可以作为 *cox1* 基因辅助的分子标记。与白氏文昌鱼线粒体基因组类似，在佛罗里达文昌鱼线粒体基因组

主编码基因中，*srRNA* 和 *lrRNA* 基因最为保守，差异位点的比例分别为 6.74%和 5.79%；差异位点数最多的基因为 *nad5* 基因(164 个)，其次为 *nad4*、*nad2* 和 *cox1* 基因，差异位点数分别为 162、151 和 148 个(表 5)。因此，在佛罗里达文昌鱼群体遗传的研究中，*nad5*、*nad4* 和 *nad2* 基因可以作为 *cox1* 基因辅助的分子标记。从矛形文昌鱼线粒体基因组主编码基因的差异位点分析，可以看出，*nad4* 基因的差异位点数最多(106 个)(表 6)。因此，在矛形文昌鱼群体遗传的研究中，*nad4* 基因是理想的分子标记。而在尖刀偏文昌鱼线粒体基因组中，差异位点数最多的基因为 *nad5* 基因(174 个)，其次为 *nad4* 和 *cox1* 基因，差异位点数分别为 104 和 87 个(表 7)。因此，在尖刀偏文昌鱼群体遗传的研究中，*nad5* 和 *nad4* 基因可以作为 *cox1* 基因辅助的分子标记。

综上所述，在文昌鱼群体遗传的研究中，*nad5*、*nad4* 和 *nad2* 基因是理想的分子标记，可以作为 *cox1* 基因辅助的分子标记，用于分析文昌鱼不同群体之间的遗传多样性，为其生物多样性的保护及合理利用其生物资源提供更多保障。

表 3　文昌鱼线粒体基因组 13 个蛋白质编码基因和 2 个核糖体 RNA 基因的差异位点分析

Tab. 3　Analysis of different loci of 13 protein-coding genes and two ribosomal RNA genes in amphioxus mitochondrial genomes

差异位点	基因名称														
	atp6	*atp8*	*cob*	*cox1*	*cox2*	*cox3*	*nad1*	*nad2*	*nad3*	*nad4*	*nad4L*	*nad5*	*nad6*	*srRNA*	*lrRNA*
总位点数*	678	163	1141	1548	691	788	937	1034	352	1355	275	1786	500	805	1287
不变位点数	315	49	624	962	402	469	470	388	164	541	120	761	185	413	670
差异位点数	363	114	517	586	289	319	467	646	188	814	155	1025	315	392	617
单现突变数	86	36	118	153	83	69	123	180	48	205	41	286	85	127	212
简约位点数	277	78	399	433	206	250	344	466	140	609	114	739	230	265	405
差异位点比例 (%)	53.54	69.94	45.31	37.86	41.82	40.48	49.84	62.48	53.41	60.07	56.36	57.39	63.00	48.70	47.94

*表示总位点数不包含插入和缺失位点(表 4~表 7 同)

表 4　白氏文昌鱼线粒体基因组 13 个蛋白质编码基因和 2 个核糖体 RNA 基因的差异位点分析

Tab. 4　Analysis of different loci of 13 protein-coding genes and two ribosomal RNA genes in *B. belcheri* mitochondrial genomes

差异位点	基因名称														
	atp6	*atp8*	*cob*	*cox1*	*cox2*	*cox3*	*nad1*	*nad2*	*nad3*	*nad4*	*nad4L*	*nad5*	*nad6*	*srRNA*	*lrRNA*
总位点数*	681	162	1140	1545	690	786	945	1038	351	1356	273	1794	501	847	1382
不变位点数	629	154	1086	1472	658	748	891	969	337	1305	264	1704	476	829	1351
差异位点数	52	8	54	73	32	38	54	69	14	51	9	90	25	18	31
单现突变数	34	8	45	54	23	31	41	54	14	37	8	73	21	17	26
简约位点数	18	0	9	19	9	7	13	15	0	14	1	17	4	1	5
差异位点比例(%)	7.64	4.94	4.74	4.72	4.64	4.83	5.71	6.65	3.99	3.76	3.30	5.02	4.99	2.13	2.24

表 5　佛罗里达文昌鱼线粒体基因组 13 个蛋白质编码基因和 2 个核糖体 RNA 基因的差异位点分析

Tab. 5　Analysis of different loci of 13 protein-coding genes and two ribosomal RNA genes in *B. floridae* mitochondrial genomes

差异位点	基因名称														
	atp6	*atp8*	*cob*	*cox1*	*cox2*	*cox3*	*nad1*	*nad2*	*nad3*	*nad4*	*nad4L*	*nad5*	*nad6*	*srRNA*	*lrRNA*
总位点数*	681	162	1140	1545	690	786	939	1038	351	1356	273	1794	501	846	1364
不变位点数	587	148	993	1397	621	710	843	887	323	1194	246	1630	445	789	1285
差异位点数	94	14	147	148	69	76	96	151	28	162	27	164	56	57	79
单现突变数	56	11	89	97	37	45	59	97	15	97	18	114	37	35	55
简约位点数	38	3	58	51	32	31	37	54	13	65	9	50	19	22	24
差异位点比例(%)	13.80	8.64	12.89	9.58	10.00	9.67	10.22	14.55	7.98	11.95	9.89	9.14	11.18	6.74	5.79

表 6　矛形文昌鱼线粒体基因组 13 个蛋白质编码基因和 2 个核糖体 RNA 基因的差异位点分析
Tab. 6　Analysis of different loci of 13 protein-coding genes and two ribosomal RNA genes in *B. lanceolatum* mitochondrial genomes

差异位点	基因名称														
	atp6	*atp8*	*cob*	*cox1*	*cox2*	*cox3*	*nad1*	*nad2*	*nad3*	*nad4*	*nad4L*	*nad5*	*nad6*	*srRNA*	*lrRNA*
总位点数*	681	162	1140	1545	690	786	939	1038	351	1356	273	1794	501	843	1376
不变位点数	662	159	1119	1514	675	769	918	1013	342	1250	268	1776	491	825	1365
差异位点数	19	3	21	31	15	17	21	25	9	106	5	18	10	18	11
单现突变数	6	1	7	12	9	5	10	10	4	94	3	7	6	12	6
简约位点数	13	2	14	19	6	12	11	15	5	12	2	11	4	6	5
差异位点比例(%)	2.79	1.85	1.84	2.01	2.17	2.16	2.24	2.41	2.56	7.82	1.83	1.00	2.00	2.14	0.80

表 7　尖刀偏文昌鱼线粒体基因组 13 个蛋白质编码基因和 2 个核糖体 RNA 基因的差异位点分析
Tab. 7　Analysis of different loci of 13 protein-coding genes and two ribosomal RNA genes in *A. lucayanum* mitochondrial genomes

差异位点	基因名称														
	atp6	*atp8*	*cob*	*cox1*	*cox2*	*cox3*	*nad1*	*nad2*	*nad3*	*nad4*	*nad4L*	*nad5*	*nad6*	*srRNA*	*lrRNA*
总位点数*	685	174	1143	1548	691	789	943	1041	354	1359	275	1789	504	848	1366
不变位点数	638	163	1057	1461	662	751	860	964	322	1255	251	1615	452	828	1295
差异位点数	47	11	86	87	29	38	83	77	32	104	24	174	52	20	71
差异位点比例(%)	6.86	6.32	7.52	5.62	4.20	4.82	8.80	7.40	9.04	7.65	8.73	9.73	10.32	2.36	5.20

3　结论

作者在 51 个文昌鱼线粒体基因组全序列的基础上，全面分析了文昌鱼线粒体基因组的基本特征、基因排列、蛋白质编码基因和差异位点等。文昌鱼线粒体基因组长度介于 14, 975bp 和 15, 146bp 之间。7 个文昌鱼物种的线粒体基因组均编码后生动物标准的 37 个基因。从目前的线粒体基因组数据分析，文昌鱼线粒体基因组共有 3 种基因排列方式，其中文昌鱼属和侧殖文昌鱼共有的基因排列，与脊椎动物线粒体基因组的原始排列相比最为接近。因此，文昌鱼属和侧殖文昌鱼属线粒体基因组的基因排列方式代表了文昌鱼原始的基因排列方式。与此相比，尖刀偏文昌鱼和偏文昌鱼 *A.* sp. *A TK-20072* 的线粒体基因组发生 3 个基因或者基因区块的易位以及 1 个基因区块的倒位；偏文昌鱼 *A. inferum* 线粒体基因组发生 1 个基因区块的易位、1 个基因区块的倒位和 1 个转运 RNA 的易位并倒位。白氏文昌鱼、佛罗里达文昌鱼、矛形文昌鱼和尖刀偏文昌鱼的 13 个蛋白编码基因的同义替换率(*Ks*)和非同义替换率(*Ka*)。4 个文昌鱼线粒体基因组的 13 个蛋白质编码基因的 *Ka*/*Ks* 比值都远远低于 1(介于 0 和 0.3363 之间)，显示出较强的纯化(负)选择。文昌鱼线粒体基因组种间和种内基因差异位点和基因变异分析表明，在文昌鱼群体遗传的研究中，*nad5*、*nad4* 和 *nad2* 基因是理想的分子标记，可以作为 *cox1* 基因辅助的分子标记，用于分析文昌鱼不同群体之间的遗传多样性，为其生物多样性的保护及合理利用其生物资源提供更多保障。

参考文献:

[1] Dellaporta S L, Xu A, Sagasser S, et al. Mitochondrial genome of *Trichoplax adhaerens* supports placozoa as the basal lower metazoan phylum[J]. Proceedings of the National Academy of Sciences of the United States of America, 2006, 103(23): 8751-8756.

[2] Helfenbein K G, Fourcade H M, Vanjani R G, et al. The mitochondrial genome of *Paraspadella gotoi* is highly reduced and reveals that chaetognaths are a sister group to protostomes[J]. Proceedings of the National Academy of Sciences of the United States of America, 2004, 101(29): 10639-10643.

[3] Boore J L. Animal mitochondrial genomes[J]. Nucleic Acids Res, 1999, 27(8): 1767-1780.

[4] Boore J L, Brown W M. Big trees from little genomes: mitochondrial gene order as a phylogenetic tool[J]. Current Opinion In Genetics & Development, 1998, 8(6): 668-674.

[5] 王义权, 方少华. 文昌鱼分类学研究及展望[J]. 动物学研究, 2005, 26(6): 666-672.
[6] 王奕华, 张士璀, 刘振辉. 青岛文昌鱼遗传多样性的RAPD分析[J]. 海洋水产研究, 2004, 25(3): 21-27.
[7] 张士璀, 袁金铎, 李红岩. 文昌鱼—研究脊椎动物起源和进化的模式动物[J]. 生命科学, 2001, 13(5): 214-218.
[8] Thompson J D, Gibson T J, Plewniak F, et al. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools[J]. Nucleic Acids Res, 1997, 25(24): 4876-4882.
[9] Nohara M, Nishida M, Nishikawa T. New complete mitochondrial DNA sequence of the lancelet *Branchiostoma lanceolatum* (Cephalochordata) and the identity of this species' sequences[J]. Zoological Science, 2005, 22(6): 671-674.
[10] Boore J L, Daehler L L, Brown W M. Complete sequence, gene arrangement, and genetic code of mitochondrial DNA of the cephalochordate *Branchiostoma floridae* (Amphioxus) [J]. Molecular Biology and Evolution, 1999, 16(3): 410-418.
[11] Kon T, Nohara M, Yamanoue Y, et al. Phylogenetic position of a whale-fall lancelet (Cephalochordata) inferred from whole mitochondrial genome sequences[J]. BMC Evolutionary Biology, 2007, 7: 127.
[12] Naylor G J, Brown W M. Amphioxus mitochondrial DNA, chordate phylogeny, and the limits of inference based on comparisons of sequences[J]. Systematic Biology, 1998, 47(1): 61-76.
[13] Nohara M, Nishida M, Miya M, et al. Evolution of the mitochondrial genome in cephalochordata as inferred from complete nucleotide sequences from two epigonichthys species[J]. Journal of Molecular Evolution, 2005, 60(4): 526-537.
[14] Yang Z. PAML 4: phylogenetic analysis by maximum likelihood[J]. Mol Biol Evol, 2007, 24(8): 1586-1591.
[15] Rozas J, Sanchez-DelBarrio J C, Messeguer X, et al. DnaSP, DNA polymorphism analyses by the coalescent and other methods[J]. Bioinformatics, 2003, 19(18): 2496-2497.
[16] Schindel D E, Miller S E. DNA barcoding a useful tool for taxonomists[J]. Nature, 2005, 435(7038): 17.

Analysis of the characteristics of mitochondrial genomes and exploration of molecular markers in amphioxus

SHEN Xin[1, 2], TIAN Mei[1], MENG Xue-ping[1], CHENG Han-liang[1]
(1. College of Marine Science, Huaihai Institute of Technology, Lianyungang 222005, China; 2. Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China

Received: Apr., 20, 2010
Key words: Amphioxus; mitochondrial genome; genetic different loci; molecular marker

Abstract: Mitochondrial genomes have been widely used in metazoan molecular phylogeny and genetic research. Amphioxus, as research model animals, occupies an important position in the origin and evolution of vertebrates. The basic characteristics of amphioxus mitochondrial genomes were fully revealed by our comprehensive analysis of 51 mitochondrial genome sequences from seven species. Amphioxus mitochondrial genomes contain 37 standard metazoan genes. Three types of gene arrangements were found in amphioxus mitochondrial genomes. Gene arrangements of *Branchiostoma* and *Epigonichthys* are most similar to those of typical vertebrate mitochondrial genomes. Therefore, gene arrangement shared by *Branchiostoma* and *Epigonichthys* represents the ancestral amphioxus gene order. In contrast, gene rearrangements were found in three species of the genus *Asymmetron*. The *Ka*/*Ks* ratios of the 13 amphioxus mitochondrial protein-coding genes were much lower than one (range between 0 and 0.3363), indicating a strong purifying selection (negative selection). The genetic variation analysis of main genes (13 protein coding genes and two ribosomal RNA genes) among and within amphioxus species shows that genes *nad5*, *nad4* and *nad2* are ideal molecular markers supplementary to the *cox1* gene, and will provide useful information on conservation of amphioxus biological diversity and the utilization of biological resources.

(本文编辑: 谭雪静)