

中华绒螯蟹线粒体 CO II 基因全序列测定

孙红英,周开亚,龚美蓉,童宗中

(南京师范大学生命科学学院遗传资源研究所,南京 210097)

[摘要] 通过克隆测序方法,报道了中华绒螯蟹(*Eriocheir japonica sinensis*)线粒体细胞色素氧化酶亚基 II(CO II)基因全序列.与已知甲壳动物线粒体同源序列的比较显示,绒螯蟹 CO II 基因的核苷酸组成及其编码的氨基酸组成与软甲类甲壳动物最相近似.其序列组成与密码子使用对 A、T 核苷酸没有明显的偏倚.对中华绒螯蟹 CO II 全序列及其序列特征的研究,为进一步研究短尾类的系统发生和绒螯蟹属的分子进化提供了分子标记.

[关键词] 中华绒螯蟹;CO II 基因;全序列

[中图分类号] Q959.223.03; [文献标识码] A; [文章编号] 1001-4616(2001)04-0088-05

线粒体细胞色素氧化酶亚基 II(cytochrome oxidase subunit II, CO II)基因是动物线粒体 13 种蛋白编码基因之一.近年来,对 CO II 序列的研究,为有关节肢动物分子系统学与分子进化等问题的探讨提供了依据^[1].迄今已测定了 6 种甲壳动物 CO II 基因的全序列^[2~6],包括鳃足纲(Branchiopoda):蚤状溞(*Daphnia pulex*)和卤虫(*Artemia franciscana*);软甲纲(Malacostraca):十足目(Dacapoda):斑节对虾(*Penaeus monodon*)、美菲对虾(*Penaeus notialis*)、美洲龙虾(*Panulirus argus*)和长手寄居蟹(*Pagurus longicarpus*).但对十足目短尾类(Brachyura)尚无 CO II 全序列的研究报道.本文通过克隆测序方法,测定了中华绒螯蟹(*Eriocheir japonica sinensis*)线粒体 CO II 基因的全序列.通过与已知甲壳动物以及昆虫同源序列的比较,研究中华绒螯蟹 CO II 基因核苷酸组成及其编码的氨基酸组成等特征.为进一步探讨短尾类的系统发生、研究绒螯蟹属的分子进化提供分子标记.

1 材料与方法

1.1 材料

中华绒螯蟹由安徽河蟹原种场提供.

1.2 线粒体 DNA 提取、克隆与测序

取雌性个体卵巢组织,使用碱变性法提取线粒体基因组 DNA.选择 DNA 限制性内切酶 *Hind* III 对线粒体 DNA 进行酶切,得到长度分别约为 0.7、0.8、1、2、4 kp 和 7 kp 的 6 个酶切片段.对其中约为 4kb 的片段进行纯化,与预先制备的 P^{GEM}3Zf 载体进行连接,使用 JM109 感受态细胞进行转化,筛选重组质粒.用质粒小量纯化试剂盒(GIBCO 公司)进行质粒 DNA 纯化.对重组质粒进行直接测序和亚克隆测序.测序反应使用载体通用引物, DNA 序列测定在 ABI 310 型

收稿日期 2001-07-20

基金项目 江苏省应用基础研究项目(BJ95109)

作者简介 孙红英,1964—,女,南京师范大学博士研究生,南京师范大学生命科学学院副教授,从事动物分子生物学研究.

联系人 周开亚,1932—,南京师范大学生命科学学院教授,博士生导师,主要从事动物学研究.(E-mail:kyzhounj@jlonline.com)

遗传分析仪(Perkin Elmers)进行。

1.3 序列比对与数据分析

将测得的 DNA 序列在 GenBank 中进行核苷酸序列检索,从而对其中的 CO II 序列进行初步定位。应用 Clustal X (version 1.8) 软件,将经过初步定位的 CO II 序列(包括其上、下游的核苷酸序列在内)与节肢动物同源序列进行比对,确定 CO II 基因的起始与终止信号。在 DNA-Club 软件中,使用果蝇(*Drosophila*)通用密码子,对 CO II 核苷酸序列进行翻译。在 MEGA (version 2.0) 软件^[7]中,分别对核苷酸序列与氨基酸序列的组成进行统计。参照 Crozier 和 Crozier^[8]提出的富 G + C 密码子(G + C - rich codons)与富 A + T 密码子(A + T - rich codons)比值,基于推导的各物种 CO II 多肽链的氨基酸组成,估计密码子使用模式的偏倚情况。其中,富 G + C 密码子的氨基酸包括 Pro、Ala、Arg 和 Gly;富 A + T 密码子的氨基酸包括 Phe、Ile、Met、Tyr 和 Asn。文中引用的 6 种甲壳动物与 1 种昆虫的同源序列在 GenBank 中的检索号如下:*D. pulex* (AF117817) *A. franciscana* (X69067) *P. monodon* (X69067) *P. notialis* (X84350) *P. argus* (AJ133050) *P. longicarpus* (AF150756) *D. yacuba* (X03240)。

2 结果

2.1 中华绒螯蟹 CO II 基因全序列

经测定,中华绒螯蟹线粒体基因 CO II 全序列长度为 696 bp,推导的多肽链由 231 个氨基酸组成。其核苷酸序列与氨基酸序列如图 1 所示。

1	ATG	GCA	ACA	TGA	GCG	TAT	TJA	AAT	TTT	CAA	GAC	AGA	GCT	TCT	CCT	TJA	ATA	GAA	CAA	TTA	60
1	M	A	T	W	A	Y	L	N	F	Q	D	S	A	S	P	L	M	E	Q	L	20
61	ATT	TTT	TTT	CAC	GAT	CAT	ATT	ATA	TTA	GTG	ATT	GTG	TTA	ATT	GTA	ACA	TTT	GTT	GGT	TAT	120
21	I	F	F	H	D	H	I	M	L	V	I	V	L	I	V	T	F	V	G	Y	40
121	ATA	ATA	GCA	ACA	TTG	TTT	TTC	AAC	TCC	CTC	ATT	AAC	CGC	TAT	ATA	TTA	GAA	AAT	CAA	CCA	180
41	M	M	A	T	L	F	F	N	S	L	I	N	R	Y	M	L	E	N	Q	P	60
181	ATC	GAA	GTA	ATT	TGA	ACA	TCA	GTT	CCA	GCT	TTT	ATT	TTA	ATC	TTT	ATT	GCT	TTA	CCA	TCC	240
61	I	E	V	I	W	T	S	V	P	A	F	I	L	I	F	I	A	L	P	S	80
241	CTT	CGC	TTA	CTT	TAT	CTT	TTA	GAT	GAA	GTA	AAC	AAT	CCT	TCT	ATA	ACT	CTI	AAA	ACA	ATT	300
81	L	R	L	L	Y	L	L	D	E	V	N	N	P	S	M	T	L	K	T	I	100
301	GGC	CAC	CAA	TGA	TAT	TGA	AGG	TAT	GAA	TAT	TCA	GAT	TTC	CTT	CAA	GTA	GAA	TTT	GAC	TCA	360
101	G	H	Q	W	Y	W	S	Y	E	Y	S	D	F	L	Q	V	E	F	D	S	120
361	TAT	ATA	GTA	CCA	TCT	AAT	GAG	CTG	GAA	GTA	TCA	GGG	TTC	CGG	CTA	CTA	GAC	GTA	GAC	AAT	420
121	Y	M	V	P	S	N	E	L	E	V	S	G	F	R	L	L	D	V	D	N	140
421	CGT	ACC	GTT	CTT	CCA	ATA	AAT	ACT	CAA	ATC	CGA	GTT	CTA	ATT	ACA	GCT	GCA	GAT	GTC	ATC	480
141	R	T	V	L	P	M	N	T	Q	I	R	V	L	I	T	A	A	D	V	I	160
481	CAT	TCC	TGG	ACA	GTC	CCA	TCC	TTA	GGT	ATT	AAA	GCC	GAT	GCA	ATT	CCC	GGG	CGG	CTA	AAT	540
161	H	S	W	T	V	P	S	L	G	I	K	A	D	A	I	P	G	R	L	N	180
541	CAA	TCA	AGG	TTT	TTA	ATC	AAC	CGA	CCT	GGA	CTA	TTT	TAC	GGC	CAA	TGT	TCA	GAA	ATT	TGC	600
181	Q	S	S	F	L	I	N	R	P	G	L	F	Y	G	Q	C	S	E	I	C	200
601	GGG	GCA	AAC	CAC	AGA	TTT	ATG	CCC	ATT	GTT	ATT	GAA	AGA	ACT	TCT	ATT	AAC	TCA	TTT	CTA	660
201	G	A	N	H	S	F	M	P	I	V	I	E	S	T	S	I	N	S	F	L	220
661	AAT	TGA	ATT	TCT	CTA	TCA	ATT	GAA	GCG	TCA	CCC	TAA	696								
221	N	W	I	S	L	S	I	E	A	S	P	*									

图 1 中华绒螯蟹(*E. j. sinensis*)线粒体 CO II 基因全序列

2.2 中华绒螯蟹 CO II 基因序列特征及其与其它甲壳动物同源序列的比较

中华绒螯蟹 CO II 基因的核苷酸组成中 ,A + T 含量为 65.5%(A ,31.2% ;T ,34.3% ;C ,20% ;G ,14.5%) ,其中 ,密码子第 3 位的 A + T 含量为 75.9% .由氨基酸组成估计的密码子使用模式 ,即富 G + C/富 A + T 密码子的比值为 0.53.

对中华绒螯蟹与已知 6 种甲壳动物 CO II 全序列特征进行比较 .结果显示(表 1.) ,中华绒螯蟹的 CO II 基因在已知的 7 种甲壳动物中最长 ,而 *D. pulex* 的最短.

中华绒螯蟹属于软甲类十足目 ,其核苷酸组成与甲壳动物另 4 种十足目物种间的相似性较高(75% ~ 68%) ,而与 2 种鳃足类间的相似性较低(62% ~ 61%) .与此类似 ,中华绒螯蟹与十足目另 4 种间 CO II 氨基酸组成的相似性较高(78% ~ 71%) ,而与鳃足类的相似性较低(69% ~ 61%) .

与果蝇线粒体基因的同源序列相比 ,7 种甲壳动物 CO II 序列的 A + T 含量(57.7% ~ 69.1%)和密码子第 3 位的 A + T 含量(60.3% ~ 88.2%)都比较低(果蝇分别为 73.9% 和 93.4%) ,而密码子第 3 位 G 和 C 的含量 ,分别达到 7.3% ~ 12.2% 和 14.8% ~ 29.2% ,但 *P. monodon* 和 *P. longicarpus* 例外 .*P. monodon* 密码子第 3 位 G 和 C 的含量均比较低(4.4% 和 7.4%) ,而 *P. longicarpus* 密码子第 3 位的 G 含量(3%)甚至低于果蝇(3.1%) .

甲壳动物(不包括 *P. longicarpus*)CO II 序列的富 G + C 与富 A + T 密码子比值(0.53 ~ 0.69)较果蝇同源序列的比值高 .其中 ,中华绒螯蟹的比值低于其它甲壳动物(0.62 ~ 0.69) ,但较 *P. longicarpus* 高 .

表 1 7 种甲壳类动物线粒体 CO II 基因全序列核苷酸组成及推导的氨基酸组成的比较

	<i>E. j. sinensis</i>	<i>P. longicarpus</i>	<i>P. monodon</i>	<i>P. notialis</i>	<i>P. argus</i>	<i>D. pulex</i>	<i>A. franciscana</i>	<i>D. yacuba</i>
A%	31.2	33.8	32.3	29.9	28.6	28.7	26.8	33.8
T%	34.3	35.1	36.8	34.4	29.1	30.3	35.7	40.1
C%	20.0	17.5	15.6	18.8	24.4	24.2	19.8	12.4
G%	14.5	13.6	15.4	16.9	17.9	16.8	17.8	13.0
总核苷酸数	696	690	688	688	688	679	687	685
核苷酸相似性(%)	100	75	75	74	68	62	61	71
总 A + T%	65.5	68.9	69.1	64.3	57.7	59	62.5	73.9
密码子第 3 位 A + T%	75.9	85.7	88.2	76.9	60.3	62.0	73.0	93.4
C%	16.8	11.3	7.4	14.8	28.4	29.2	14.8	3.5
G%	7.3	3.0	4.4	8.3	11.4	8.8	12.2	3.1
氨基酸总数	231	229	229	229	229	226	228	228
氨基酸相似性(%)	100	75	78	77	71	69	61	72
富 G + C/富 A + T 密码子	0.53	0.50	0.65	0.69	0.63	0.62	0.66	0.51

3 讨论

CO II 基因的长度在不同甲壳动物之间存在一定变化 ,这种蛋白编码基因长度的变化 ,在动物线粒体基因中比较常见 .对核苷酸序列的比对发现(图 1) ,7 种甲壳动物 CO II 的起始一致 ,除 *A. franciscana* 以 GTG 为起始密码子以外 ,其余均以 ATG 作为起始信号 .造成 CO II 基因

长度差异的原因是终止信号位点的后移.与核苷酸链最短的 *D. pulex* 相比,中华绒螯蟹核苷酸链的终止密码子向后移动了 5 个氨基酸位点.但在同一属内, *CO II* 核苷酸链的长度相对保守.

	起始信号↓		终止信号↓
<i>E.j.sinensis</i>	ATG GCA ACA TGA	ATT TCT CTA TCA ATT GAA GCG TCA CCC TAA	
	M A T W	I S L S I E A S P *	
<i>P.longicarpus</i>	ATG CCA ACA TGA	GTT TCA TCT TCT AAA GAC GAC TAA --- ---	
	M P T W	V S S S K D D *	
<i>P.monodon</i>	ATG CCA ACA TGA	ATT TCC TCA GCT GAT GAA GCT T -- --- ---	
	M P T W	I S S A S E A *	
<i>P.notialis</i>	ATG CCA ACA TGA	ATT TCT ACT TCC AGA GAA GCC T -- --- ---	
	M P T W	I S T S S E A *	
<i>P.pargus</i>	ATG GCA ACA TGA	GTA ATG GCT TCA AGA GAA GAT T -- --- ---	
	M A T A	V M A S S E D *	
<i>A.franciscana</i>	GTG TCT CAA TGA	CTT GAG TTA CAA ATT TCA TAG --- --- ---	
	V S Q W	L E L Q I S *	
<i>D.pulex</i>	ATG GCA ACA TGA	GTT AAA TCA TTC T -- --- --- --- ---	
	M A T W	V K S F *	
		终止信号↑	

图2 7种甲壳动物 CO II 基因的起始信号与终止信号

在 7 种甲壳动物 *CO II* 的核苷酸组成及其编码的氨基酸组成上,中华绒螯蟹与鳃足类之间的相似性(平均 61.5%和 65%)远低于它与其它软甲类之间的相似性(平均 73% ± 0.029%和 73% ± 0.026%〔*Panulirus* 例外〕),甚至低于中华绒螯蟹与果蝇间的相似性(71% ;72%).这一结果提示,中华绒螯蟹与其它软甲类之间的亲缘关系密切,而它与鳃足类之间的亲缘关系,较之与昆虫间的关系更远.这与 Garcia-Macado^[6]对甲壳类与昆虫间系统发生关系的研究结果一致.值得注意的是,中华绒螯蟹与 *Panulirus* 间 *CO II* 序列的相似性(包括核苷酸相似性和氨基酸相似性)低于它与其它 3 种十足类间的相似性(甚至低于它与果蝇间的相似性),提示中华绒螯蟹与 *Pagurus* 和 *Penaeus* 的亲缘关系较近,而 *Panulirus* 与它们的亲缘关系较远.

Flook^[9]对飞蝗 (*Locusta migratoria*) 线粒体 DNA 的研究分析认为,昆虫线粒体蛋白编码基因的核苷酸组成具有明显的 A + T 偏倚,密码子的使用也表现出对富 A + T 密码子的偏倚.与几种甲壳动物以及果蝇同源序列的比较表明,中华绒螯蟹 *CO II* 总的 A + T 含量和密码子第 3 位的 A + T 含量均高于甲壳动物的平均含量(63.7%和 74.6%),但远低于果蝇相应的 A + T 含量;其富 G + C 与富 A + T 密码子的比值较果蝇相应的比值高.可见,中华绒螯蟹 *CO II* 的序列组成对 A、T 核苷酸的偏倚程度比较低;同时,对使用富 A + T 密码子的偏倚也较昆虫低.

致谢:安徽河蟹原种场张小梅场长、李自保场长馈赠实验材料,特致衷心感谢.

[参考文献]

[1] 邵红光,张亚平,柯欣,等.南极隐虫兆线粒体 CO II 基因序列及无翅类昆虫的分子进化[J].中国科学(C 辑),2000,30,3:330—336.
[2] Crease T J. The complete sequence of the mitochondrial genome of *Daphnia pulex* (Cladocera : Crustacean) [J]. Gene ,1999 (233) 89—99.
[3] Valverde J R ,Batuecas B. The complete mitochondrial sequence of the crustacean *Artemia franciscana*[J]. J Mol 万方数据

- Evol ,1994 (39) :400—408.
- [4] Wilson K ,Cahill V ,Ballment E ,et al .The complete sequence of the mitochondrial genome of the crustacean *Penaeus monodon* :are malacostracan crustaceans more closely related to insects than to Branchiopods[J]. Mol Biol Evol , 2000 ,17(6) :863—874.
- [5] Hickerson M J ,Cunningham C W. Dramatic mitochondrial gene rearrangements in the Hermit crab *Pagurus longicarpus* (Crustacea ,Anomura) [J]. Mol Biol Evol 2000 ,17(4) :639—644.
- [6] Garcia-Machado E ,Pempera M. Mitochondrial genes collectively suggest the paraphyly of crustacea with respect to insect[J]. J Mol Evol ,1999 (9) :142—149.
- [7] Kumar S ,Tamura K ,Jakobsen I B ,et al . MEGA 2. 0 :Molecular evolutionary genetics analysis software[J]. Bioinformatics(Submitted) 2001.
- [8] Crozier R H ,Crozier Y C. The mitochondrial genome of the honeybee *Apis mellifera* : complete sequence and genome organization[J]. Genetics ,1993 ,133(1) :97—117.
- [9] Flook P K ,Rowell C H ,Gellissen G. The sequence ,organization and evolution of the *Locustamigratoria mitochondrial genome*[J]. J Mol Evol ,1995 ,41(6) :928—941.

The Complete Sequence of the Mitochondrial CO Ⅱ Gene of *Eriocheir japonica sinensis* (Crustacea : Malacostraca : Decapoda)

Sun Hongying ,Zhou Kaiya ,Gong Meirong ,Tong Zongzhong

(Institute of Genetic Resource ,College of Life Sciences ,Nanjing Normal University ,Nanjing 210097 ,PRC)

Abstract :The complete sequence of the mitochondrial cytochrome oxidase subunit Ⅱ (CO Ⅱ) gene of malacostracan crustacean *Eriocheir japonica sinensis* has been determined. Comparisons of this sequence with the homologous sequences from the other six crustaceans and one insect indicate that the nt composition and aa composition of *E. j. sinensis* are more similar to other malacostracans than they are to the branchiopod crustaceans. The sequence of *E. j. sinensis* is less A + T - rich than the counterpart of insect and the pattern of codon usage is less biased.

Key words :*Eriocheir japonica sinensis* ;CO Ⅱ gene ;complete sequence

[责任编辑 孙德泉]