

# 动物 mtDNA 控制区及保守与异质

苏瑛<sup>1,2\*</sup>

(1. 甘肃农业大学, 兰州 730070; 2. 广东海洋大学)

**摘要:** 本文通过文献综述, 对动物线粒体 DNA 控制区进行了阐述。从线粒体控制区 (control region) 基因组的研究出发, 重点介绍了动物线粒体控制区基因组结构特点。主要结论: 由于碱基替换、插入和缺失以及重复序列数目的变异致使 D-loop 成为 mtDNA 中变异最多的区域, 但突变和结构重排并不是发生在整个 D-loop 区域, 而是在高变区; 大多研究集中在 mtDNA D-loop 保守区和异质方面: 对 D-loop 序列分析, 能较好地阐明动物的起源, 在动物亲缘关系鉴定、系统进化和物种形成方式的研究等领域具有广阔的研究和应用前景。

**关键词:** 动物; mtDNA 控制区; D-loop; 保守; 异质

**中图分类号:** Q78 **文献标识码:** A **文章编号:** 1000-7083 (2005) 04-0669-04

## Conserved and Heteroplasmy on Mitochondrial DNA Control Region in Animal

SU Ying<sup>1,2</sup>

(1. Gansu Agricultural University, Lanzhou 730070; 2. Guangdong Ocean University)

**Abstract:** In this paper, the mitochondrial DNA control region (D-loop) was reviewed. In terms of nucleotide substitutions, short insertions/deletions (indels), and dynamics of variable-number tandem repeats (VNTRs), D-loop are usually considered to be the most variable parts of mtDNAs. However, site mutability and structural rearrangements are not distributed randomly across the entire region, but affect particular hypervariable sites and domains. The research of mitochondrial DNA genome of D-loop focus on conserved domain and heteroplasmy contributed to nucleotide variance and number variation that altered in mitochondrial DNA hypervariable region. Most of the analysis of D-loop sequences can aid in researching the origin of animal, the animal genetic relationship identification, analysis of population variance, evolution, individual or cell recognition, and species way.

**Key words:** animal; mitochondrial DNA; control region; D-loop; conserved; heteroplasmy

动物线粒体 DNA (mitochondrial DNA, mtDNA) 是一个闭合双链环状 DNA 分子。mtDNA 控制区 (control region) 是 mtDNA 复制与转录的关键部位, 对 mtDNA 的转录和复制起调控作用<sup>[1~3]</sup>, 是分子中进化最快的一个区域, 使之成为 mtDNA 基因组中序列和长度变异最大、进化速率最高、最具多态的区域<sup>[4]</sup>。mtDNA 控制区可忠实的揭示不同群体的遗传分化关系, 而被广泛地应用于群体的起源研究, 作为进化生物学研究一个强有力的工具, 在高等动物微进化研究中开拓了一个崭新的领域<sup>[4,5]</sup>。

### 1 动物线粒体控制区 (control region)

动物 mtDNA 可分为编码区和非编码区。编码区 mtDNA 编码着大约 22 个 tRNAs, 2 个 rRNAs (12S rRNA 和 16S rRNA) 和 13 个蛋白质多肽。

非编码区即线粒体基因组的控制区 (control region) 或称 D-loop (displacement loop) 区, 是 mtDNA 中唯一不编码多肽链的核苷酸片段, 序列长度  $1155 \pm 2$  核苷酸 (nt)<sup>[3]</sup> 占全部 mtDNA 分子的 6% 左右。

#### 1.1 mtDNA 控制区 (control region) 基因组顺序

大部分动物 mtDNA 基因排列的顺序基本一致, 哺乳动物中 Cyt b 和 ND5 基因间由 tRNA<sup>Glu</sup> 和 ND6 间隔开 (5' cytochrome b-tRNA<sup>Thr</sup>-tRNA<sup>Pro</sup>-D-loop3')。鸟类控制区序列研究表明, Cyt b 基因和 ND5 是连续的, 它们之间仅被少数几个核苷酸所分开, 连接 tRNA<sup>Glu</sup> 和 ND6 的序列的位置紧接 D-loop 区 (5' ND6-tRNA<sup>Glu</sup>-D-loop3'), 家鸡、鸭测序结果相同<sup>[6]</sup>, 反应了鸟类的线粒体基因组出现的 tRNA<sup>Glu</sup> 和 ND6 基因转座。

收稿日期: 2004-11-30 修回日期: 2005-04-06

作者简介: 苏瑛 (1962~), 女, 在读博士, 副教授, 主要从事动物科学及生态学研究

在多数脊椎动物 D-loop 的 5' 端与 tRNA<sup>pro</sup> 相接, 位于 tRNA<sup>pro</sup> 和 tRNA<sup>phe</sup> 之间, 但鸟类的 D-loop 的 5' 端和 3' 端旁分别与 tRNA<sup>glu</sup> 和 tRNA<sup>phe</sup> 连接。多数低等动物的控制区特征不明显, 黑腹果蝇的 D-loop 位于 tRNA<sup>ile</sup> 与 12sRNA 之间, 海虾的 D-loop 则位于 tRNA<sup>met</sup> 和 12sRNA 之间<sup>[7]</sup>。

1.2 动物 mtDNA 控制区的大小

Larizza (2002)<sup>[9]</sup> 研究动物 D-loop 区大约 1100 bp 碱基, Shi Yanfeng 研究鹿科长度在 909~1049 bp 之间, 鸟类为 1072 bp-1240 bp (黑腹滨鹬、绿雀), 平均值为 1168 bp, 人的为 122 bp、小鼠 879 bp、牛 910 bp、爪蟾 2134 bp, 品种内序列大小差异认为是由于 5' 和 3' 区域有较小片段 (1~20 bp) 插入和缺失及 3' 区域串联重复序列的数量变化造成的, 这为研究 mtDNA 群体间和群体内长度变异创造了条件, 而控制区序列的大小差异是引起线粒体基因组长度变化的主要原因。

1.3 动物 mtDNA 控制区的碱基组成

D-loop 序列中 A、T 含量明显高于 G、C 含量, 为 AT 富集区, 约 62.06%。左功能区 (left

domain, 包含 D-loop 的 5' 端) 和右功能区 (right domain, 包含 D-loop 的 3' 端) 为 A 碱基富集区, 即遗传上的高变区。由保守序列区 (conserved sequence block, CSB) 构成的中间保守区 (central domain) 为 G 碱基富集区, 遗传上为保守区<sup>[1]</sup>。Pierre-André Crochet 等 (2000)<sup>[9]</sup> 研究 domain I 区富含 C, domain II 区富含 G, 而 domain III 区富含 A、T。

2 动物 mtDNA 控制区 (D-loop) 保守区

2.1 保守区结构

动物 mtDNA D-loop 保守区的结构分为 3 个部分: 中央保守区 (central conserved domain, CD)、终止序列区 (extended termination associated sequences, ETAS) 和保守序列区 (conserved sequence blocks, CSBs), 在哺乳动物中, 识别了几个短的保守序列 B、C、D、E、F。

鸟类将 D-loop 中央保守区命名为 domain II, 两边相接的两个 HVS I 和 HVS II 区域分别为 domain I 和 domain III。domain II 包含 3 个保守区 (F、D、C、boxes)。如图所示。

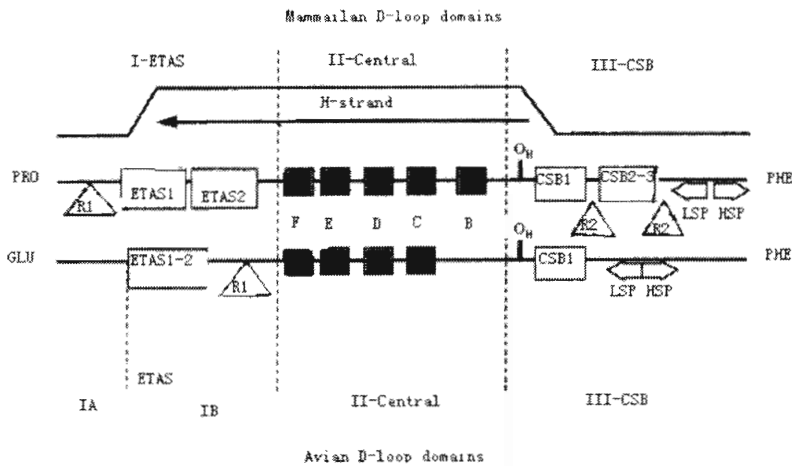


图 哺乳动物和鸡形目 mtDNA D-loop 结构的比较图

ETAS 终止序列区和哺乳动物的一致序列 (ETAS1 and ETAS2); R1 和 R2 - 在 ETAS 和 CSB 边缘区域的串联重复; B、C、D、E、F 控制区的保守序列; O<sub>H</sub>-H 链的起始复制区; CSB1、-2、and -3-CSB 边缘区域的保守序列区; LSP-L 链启动子; HSP-H 链启动子

2.2 保守区特点

CD 在整个控制区 DNA 中是最为保守的区域, ETAS 和 CSB 进化比 CD 迅速得多。ETAS 区跨度在线粒体 D-loop 3' 端的 tRNA<sup>Pro</sup> 到 CD 之间, 包含两个保守区 ETAS1 和 ETAS2 及 TAS 序列。CSB

区是重要的功能区域, 位于 5' 端 CD 到 tRNA<sup>Phe</sup> 范围, 包含 H 链的起始复制区 (O<sub>H</sub>) 和重链和轻链的启动子 (HSP and LSP) 以及 3 个短的保守区 (CSB1、CSB2 和 CSB3)。CSB 中 CSB1 是最保守的成分, 几乎所有的脊椎动物都含有 CSB1。但研究发现, 肺鱼只有 CSB2 和 CSB3, 鳕鱼只存在

CSB2, 鲤形目一些鱼类有 CSB1、CSB2 和 CSB3, 其中 CSB2 的序列终止序列区是控制区中变异最大的部分<sup>[10,11]</sup>, 它包含 DNA 复制终止相关的序列。研究中发现鸡无 CSB2, 鸟类中 CSB2 与 CSB3 合并, 而牛、海豚、羊、鲸中 CSB3 完全缺失<sup>[3,7]</sup>, 啮齿类存在 CSB1, CSB2 和 CSB3 的存在无规律性。此外, 鸟类 CSB1 和重复序列中发现了交感序列 (consensus sequence, CS), CS 在 mtDNA 的复制和转录中起着至关重要的作用。

### 3 动物 mtDNA 控制区基因重复序列 (repeated sequence, RS) 及异质 (heteroplasmy)

Kumazawa 等<sup>[12]</sup>曾发现动物 mtDNA 双链 3' 端是一段区域多是串联重复的序列 (repeated sequence, RS), 它与转座、错配表达和线粒体基因异质 (heteroplasmy) 相关。在脊椎动物线粒体 DNA 的 D-loop 5' 端发现含有串联重复序列, 拷贝数在 1~8 个之间。不同的物种重复序列不同, 重复序列所在的位置及重复单位序列 (motif) 不同, 同种动物不同个体的序列重复数也不尽相同。不仅种间有差异, 种内个体间也存在差异, 只是小于种间。这种差异在个体间表现为 mtDNA 分子的长度变异 (molecule size variation), 个体内则存在多种重复序列数目不同从而表现为异质。

#### 3.1 异质

异质 (heteroplasmy) 是产生在单个个体内的多样性<sup>[13]</sup>。异质有 D-loop 内重复序列数目的不同形成的也有碱基变异形成的, 串联重复引起的长度变异造成的异质较点突变引起的常见。重复序列能够形成发卡结构, 发卡结构引起的高频率回复突变可能也是异质形成的原因。绵羊<sup>[14]</sup>、白鲟、大西洋鳕鱼 (*Gadus morhua*) 均因 D-loop 区串联重复序列而引起异质。Arne Ludwig 等 (2000)<sup>[15]</sup>研究发现鲟鱼重复序列长度 74~83 bp, 存在异质。张四明等 (2000)<sup>[16]</sup>发现了我国一级珍稀保护动物中华鲟 (*Acipenser sinensis*) 控制区 (D-loop) 存在数目不等的串联重复序列, 该重复序列造成了中华鲟广泛的异质现象。

#### 3.2 异质与进化

D-loop 区包含 O<sub>H</sub>、HSP、LSP 等重要元件, 所以 D-loop 长度的变异势必会影响线粒体基因组的复制和转录, 从而影响到整个生物体的进化速

率。根据异质的区分及重复序列长度和结构分析表明, 异质的存在是进化活动产生的唯一原因。哺乳动物 mtDNA D-loop 区位点中的异质、串联重复、高频率的插入和缺失及特质连锁等是进化的特征。国内研究了中国大陆梅花鹿 (*Cervus nippon*) 不同种群、红原鸡亚种、关中驴、羚牛等的基因序列差异, 结果说明了这些物种的起源、进化和分布格局<sup>[17~21]</sup>。

综上所述, mtDNA D-loop 核苷酸序列分析已成为当前动物 mtDNA 研究的重点并已取得了一些进展, 而 mtDNA D-loop 的研究热点集中在 mtDNA D-loop 区内的保守区和高变异区碱基突变、重复序列数目的变异导致的异质方面, 不仅群体之间有明显差异, 而且群体中个体也存在差异, 通过对 D-loop 序列分析, 能较好地阐明动物的起源, 在动物亲缘关系鉴定、遗传纯度分析、进化研究等领域具有广阔的研究和应用前景。

### 4 参考文献

- [1] 郑冰蓉, 张亚平, 等. 南倒刺鱼巴 mtDNA D-loop 区序列的遗传多样性研究 [J]. 水利渔业, 2002, 22 (3): 15~16.
- [2] Shi Yanfeng, et al. Sequence analysis and phylogeny of deer (*Cervidae*) mtDNA control regions [J]. Acta Genetica Sinica, 2004, 31 (4): 395~402.
- [3] Ettore Randi, Lucchini. Organization and Evolution of the Mitochondrial DNA Control Region in the Avian Genus *Lectotortrix* [J]. J Mol Evol, 1998, 47: 449~462.
- [4] Giuffra EK, et al. The origin of the domestic pig; independent domestication and subsequent introgression [J]. Genetics, 2000, 154: 1785~1791.
- [5] Gares R. Drosophila melanogaster mitochondrial DNA gene organization and evolutionary consideration [J]. Genetic, 1998, 118 (4): 649~663.
- [6] Liu HZ, et al. Sequence variations in the mitochondrial DNA control region and the implications for the phylogeny of the Cypriniformes [J]. Can J Zool, 2002, 80: 569~581.
- [7] 赵兴波, 李宁, 吴常信. 线粒体 DNA 指纹技术的研究及其应用 [J]. 高技术通讯, 2004, 5: 94~98.
- [8] Larizza A, et al. Lineage specificity of the evolutionary dynamics of the mtDNA D-loop region in rodents [J]. J Mol Evol, 2002, 54: 145~155.
- [9] Pierre-André Crochet and Eric Desmarais. Slow rate of evolution in the mitochondrial control region of gulls (Aves: Laridae) [J]. J Mol Evol, 2000, 17: 1797~1806.

# 脊椎动物基因组与基因复制研究进展

刘智皓, 张修月, 王德寿\*

(西南师范大学生命科学学院, 重庆市水产科学重点实验室, 重庆北碚 400715)

**摘要:** 脊椎动物的出现是动物进化历史上一次质的飞跃。由于所有的脊椎动物在其胚胎发育中都呈现连续的解剖学特征, 因此过去很多学者都根据现存脊椎动物的形态特征和在其发育过程中的解剖学特征假想原始脊椎动物, 并推导其进化过程和起源。近年来的研究表明, 通过对脊椎动物和与之亲缘关系接近的物种之间进行基因家族、染色体结构分析, 可以对脊椎动物进化提供很多线索和证据。更多的研究表明, 脊椎动物在进化过程中很可能发生过整体基因组的复制, 基因和/或基因组的复制可能是引起脊椎动物形体结构复杂性增加的根本原因。因此, 基因和基因组的复制正在成为生物进化研究的热点问题。但这两种复制方式中哪一种产生动物形体结构和功能复杂性增加最重要的原因尚有争论。

**关键词:** 脊椎动物; 基因组复制; 基因复制; *Hox* 基因簇; 起源; 系统分析

**中图分类号:** Q78 **文献标识码:** A **文章编号:** 1000-7083(2005)04-0672-06

## Advance on Genome and Gene Duplication of Vertebrates

LIU Zhi-hao, ZHANG Xiu-yue, WANG De-shou\*

(Faculty of Life Science, South China Normal University, Key Lab of Aquaculture Technology, Chongqing 400715)

**Abstract:** The origin of the vertebrates is one of the most important transitions in animal evolution. Because the vertebrates express anatomical characteristics consistently during the process of fetation, many researchers conjectured ancient vertebrates by the characteristics of the animals living nowadays and speculate the origin and the process of evolution. Recent researches show that the analyses of gene families and chromosomal structures between vertebrates and those close to

收稿日期: 2005-03-24

作者简介: 刘智皓 (1979~), 男, 硕士研究生; 研究方向: 动物分子生物学; E-mail: mininet@swnu.edu.cn \* 通讯作者

- [10] Sbisà E, Janzariello F, Reyes A, *et al.* Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications [J]. *Gene*, 1997, 205: 125~140.
- [11] Liu HZ, *et al.* Sequence variations in the mitochondrial DNA control region and their implications for the phylogeny of the Cypiniformes [J]. *Can J Zool*, 2002, 80: 569~581.
- [12] Liu HZ. The structure and evolution of mitochondrial DNA control region of fish: a case study to bitterlings [J]. *Prog Nat Sci*, 2002, 12 (3): 266~271.
- [13] Kumazawa Y, *et al.* The complete nucleotide sequence of a snake (*Dinodon semicarinatus*) mitochondrial genome with two identical control regions [J]. *Genetics*, 1998, 150 (1): 313~329.
- [14] 郭晓华, 等. 人类线粒体 mtDNA 的分子遗传特性 [J]. *生物学通报*, 2003, 38 (12): 19~20.
- [15] Stefan Hiendleder. The complete mitochondrial DNA sequence of the domestic sheep (*Ovis aries*) and comparison with the other major ovine haplotype [J]. *J Mol Evol*, 1998, 47: 441~448.
- [16] Arne Ludwig, Bernie May, Lutz Debus, and Ingo Jenneckens. Heteroplasmy in the mtDNA control region of sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*) [J]. *Genetics*, 2000, 156: 1933~1947.
- [17] 张四明, 吴清江, 张亚平. 中华鲟 (*Acipenser sinensis*) 及相关种类的 mtDNA 控制区串联重复序列及其进化意义 [J]. *中国生物化学与分子生物学报*, 2000, 16 (4): 458~461.
- [18] 刘海, 等. 中国大陆梅花鹿 mtDNA 控制区序列变异及种群遗传结构分析 [J]. *动物学报*, 2003, 49 (1): 53~60.
- [19] 傅衍, 等. 中国家鸡的起源探讨 [J]. *遗传学报*, 2001, 28 (5): 411~417.
- [20] 刘益平, 等. 原鸡线粒体 DNA 部分序列多态性分析 [J]. *畜牧兽医学报*, 2004, 35 (2): 134~140.
- [21] 雷初朝, 等. 关中驴线粒体 DNA D-loop 多态性分析 [J]. *中国畜牧杂志*, 2004, 40 (4): 10~12.
- [21] 李明, 等. 羚牛的遗传多样性及其种群遗传结构分析 [J]. *兽类学报*, 2003, 23 (1): 10~16.