

## Gene Rearrangement of Mitochondrial Genome in the Vertebrate

ZHONG Jing<sup>1</sup>, LI Guang<sup>1</sup>, LIU Zhong-Quan<sup>2</sup>, LI Qing-Wei<sup>3</sup>, WANG Yi-Quan<sup>1</sup>

(1. School of Life Sciences, Key Laboratory of SEM for Cell Biology and Tumor

Cell Engineering, Xiamen University, Xiamen 361005, China;

2. College of Life Sciences, Nanjing Normal University, Nanjing 210097, China;

3. College of Life Sciences, Liaoning Normal University, Dalian 116029, China)

**Abstract** Up to date, complete mitochondrial genomes of 321 vertebrates are deposited in GenBank. After comparing the sequence data, we summarized the gene arrangements of these species in present study. Of those species, 81 have distinct gene rearrangements all involving rRNA genes. Among them, 9 gene rearrangements display both gene order variation and gene encoding inversion. All observed gene order changes of vertebrate mDNAs fall into 3 categories characteristically: exchange of position between nearest neighbor genes or segments; the changes close to control region or the origin of light-strand replication, occasionally accompanying with duplication of control region; or the changes near the region originally characterized as t-Q-M. All birds, snakes, crocodiles and marsupials have their unique gene orders respectively. Gene inversion, another distinguished character of the rearrangement, is commonly indicated by the transfer of rRNA genes from light-strand to heavy-strand, which is more frequently observed in the fish and mammal. Three species from different classes share a complicated inversion phenomenon: all genes encoded by heavy-strand except ND6. For explanation of these rearrangements, two models "duplication-random deletion" and "recombination", were commonly adopted by scientists. This paper also gives a brief summarization on the mechanism of these models as well as the application of gene rearrangement in phylogenetic studies.

**Key words** mitochondrial DNA; gene rearrangement; gene inversion; phylogenesis

## 脊椎动物线粒体 DNA 的基因重排

钟婧<sup>1</sup>, 李光<sup>1</sup>, 刘忠权<sup>2</sup>, 李庆伟<sup>3</sup>, 王义权<sup>1</sup>

(1. 厦门大学生命科学学院, 细胞生物学与肿瘤细胞工程教育部重点实验室, 厦门 361005

2. 南京师范大学生命科学学院, 南京 210097;

3. 辽宁师范大学生命科学学院, 大连 116029)

**摘要:** 将 GenBank 上已公布的 321 种脊椎动物 mDNA 全序列, 按纲整理归类, 绘制基因排布图并进行比对。

收稿日期: 2004-04-13; 修回日期: 2004-08-26

基金项目: 国家自然科学基金 (编号: 30170505, 30470938) 教育部骨干教师资助计划 (编号: GG-180-21002403-1740) 和教育部分子生物学启动基金资助 [Supported by NSFC (No. 30170505, 30470938), Foundation for Univ. Key Teacher from State Education Ministry (No. GG-180-21002403-1740) and SRF for ROCS, SEM]

作者简介: 钟婧 (1981-), 硕士研究生, 研究方向: 动物分子遗传

通讯作者: E-mail wangyq@xmu.edu.cn; Tel: 0592-2184427

比对结果表明: 81个物种的 mDNA 中观察到基因重排现象, 涉及脊椎动物各纲, 其中 9个物种同时存在基因顺序变化和基因倒置现象, 所有的基因重排都涉及 rRNA 的变化。脊椎动物 mDNA 基因顺序变化可分为 3类: 1) 邻接的基因或片段的位置交换; 2) 接近于控制序列或轻链起始位点的基因或片段的位置变化, 有时还伴随着控制序列的倍增; 3) H-Q 区域的变化。所有鸟类、蛇类、鳄类和有袋类的 mDNA 具有各自独特的基因排列顺序。基因倒置现象常见于鱼类和哺乳类, 且多表现为 rRNA 从轻链往重链上迁移。本文就这些基因重排现象、发生重排的机制和 mDNA 基因重排在系统发生研究中的应用做一简要概述。

关键词: 线粒体 DNA; 基因重排; 基因倒置; 系统发生

中图分类号: Q953 文献标识码: A 文章编号: 0379-4172(2005)03-0322-09

动物线粒体基因组 (mDNA) 通常为 15~20 kb 的双链闭环状结构, 根据氯化铯密度梯度离心中的密度差异, DNA 双链可分为重链 (H 链) 和轻链 (L 链)。一个典型的脊椎动物 mDNA 由 13 个蛋白质编码基因、2 个 rRNA 基因、22 个 tRNA 基因、控制区 (D-loop) 和轻链复制起始区组成, 其中除 ND6 和 8 个 tRNA 基因由 L 链编码外, 其余基因都由 H 链编码<sup>[1,2]</sup>。由于 mDNA 基因组小、进化快速、世代传递过程中一般没有重组及为母系遗传等特点, 已成为比较遗传学和系统发生学研究的理想工具, 同时 mDNA 的异常与人类某些疾病的关系也一直是医学界关注的热点之一<sup>[3,4]</sup>。

长期以来人们通常认为脊椎动物线粒体基因排布相当保守, 从盲鳗、鱼类、非洲爪蟾, 一直到人类, mDNA 都具有相同的基因排布方式<sup>[5-12]</sup>。随着公共数据库中脊椎动物 mDNA 全序列数据迅速增加, 脊椎动物 mDNA 中存在的基因重排现象越来越多地被发现<sup>[13-17]</sup>。但这些报道大都比较分散, 迄今尚未见对脊椎动物线粒体基因重排现象较系统性的比较研究; 此外, 在已有关于脊椎动物 mDNA 基因重排的报道中, 除 Houe 等在研究星康吉鳗 (*Conger myriaster*) 线粒体基因组时, 注意到 rRNA<sup>Thr</sup> 基因从重链移到轻链编码的倒置现象外<sup>[16]</sup>, 多数作者均忽略了这一现象。本文就目前已完成的脊椎动物 mDNA 基因组中基因排布情况作归纳整理, 并对 mDNA 基因重排机制和应用做一概述。

## 1 脊椎动物线粒体基因组的重排

目前已在 GenBank 上公布了 mDNA 全序列的脊椎动物有 321 种, 现以纲为单位将其整理归类, 其中圆口纲 (Cyclostomata) 4 种、鱼纲 (Pisces) 170 种、两栖纲 (Amphibia) 7 种、爬行纲 (Reptilia) 11 种、鸟纲 (Aves) 26 种、哺乳纲 (Mammalia) 103

种。根据这些 mDNA 序列绘制基因排布图, 并进行比对。结果表明, 在这些脊椎动物 mDNA 基因排布中, 240 个物种的 mDNA 基因排布完全相同 (圆口纲 2 种, 鱼纲 141 种, 两栖纲 5 种, 爬行纲 6 种, 哺乳纲 86 种), 本文称之为脊椎动物线粒体基因组的“典型排布” (图 1); 其余 81 个物种的 mDNA 中观察到基因重排现象, 涉及脊椎动物各纲, 其中 28 个物种存在基因倒置现象。可见, 脊椎动物 mDNA 基因重排并非只是进化过程中的偶然现象。

### 1.1 脊椎动物各纲 mDNA 基因排列顺序的变化

#### 1.1.1 圆口纲

圆口纲中盲鳗目的蒲氏粘盲鳗 (*Eptatretus burgeri*) 和大西洋盲鳗 (*Myxine glutinosa*) 的线粒体基因顺序与脊椎动物的典型排布一致, 但七鳃鳗目中欧洲河鳗 (*Lamprologus fluviatilis*) 和海七鳃鳗 (*Petromyzon marinus*) 的 rRNA<sup>Thr</sup>, rRNA<sup>Glu</sup>, rRNA<sup>Pro</sup> 和 Cyt b 均出现位置的迁移。前两者紧密排列移至控制区序列的 3' 端, 后者则一起移入控制区序列, 将其分隔为二 (图 1, A)。

#### 1.1.2 鱼纲

在已知 170 种鱼纲动物 mDNA 全序列中, 有 19 个物种的排列顺序不同于典型的脊椎动物排布方式。这些变化归纳起来可分为 5 类: 第一类是 rRNA<sup>Glu</sup>, rRNA<sup>Thr</sup> 和 rRNA<sup>Pro</sup> 在 ND6~12S rRNA 区域间的移位。如: 长尾鲟 (*Caebrinchus kishinouyei*) 和纤钻光鱼 (*Gonostoma gracile*) 的 rRNA<sup>Glu</sup> 分别移位到控制序列 5' 端和 Cyt b 的 3' 端, 而且纤钻光鱼的 rRNA<sup>Thr</sup> 和 rRNA<sup>Pro</sup> 的位置发生交换; 阿拉斯加黑鱼 (*Dallia pectoralis*) 的 rRNA<sup>Thr</sup> 从控制序列的 5' 端移入控制区序列中; 银鲛 (*Chimaera monstrosa*) 等 3 种鱼的 rRNA<sup>Pro</sup> 均移位至控制序列中; 珍珠鱼 (*Carapus bermudensis*) 的 rRNA<sup>Pro</sup> 移

Typical arrangement in vertebrate

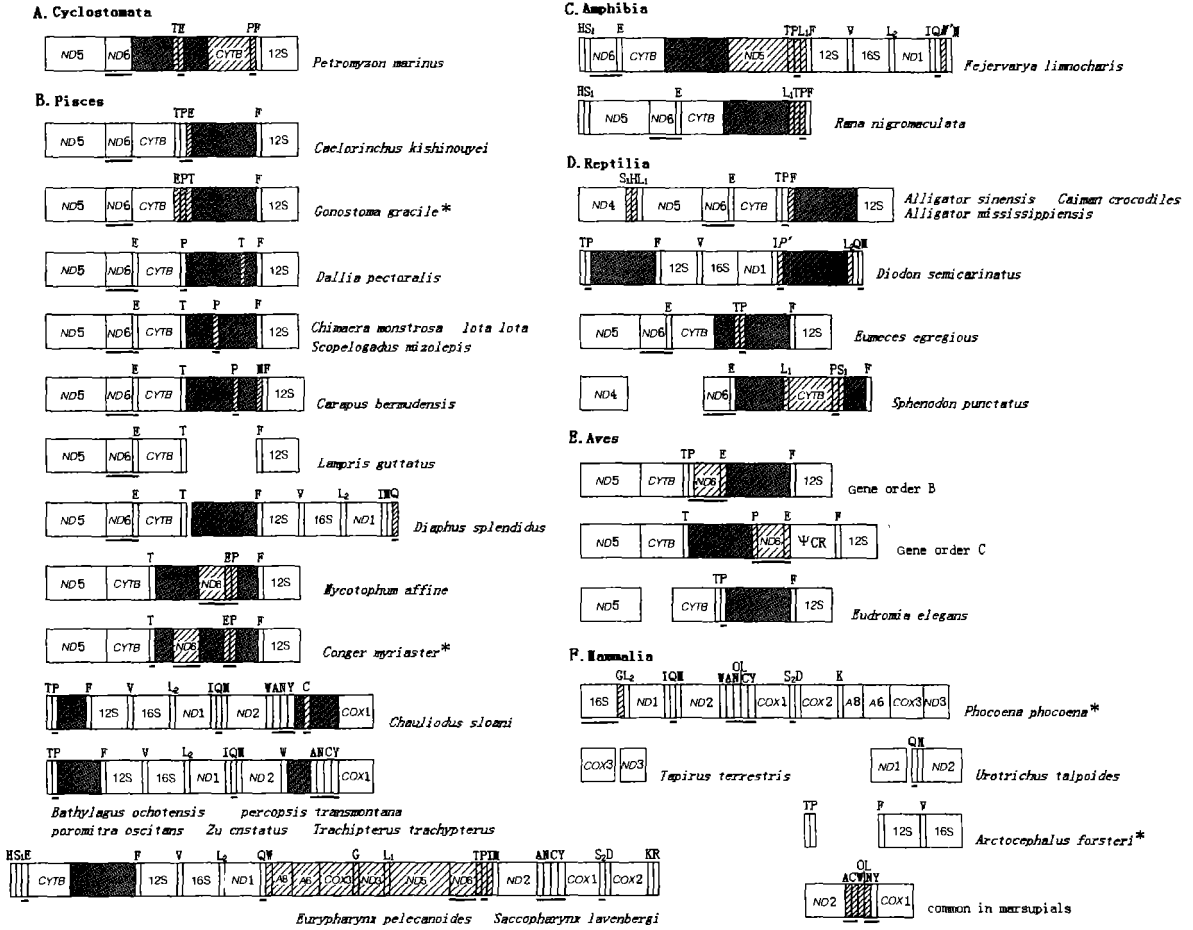
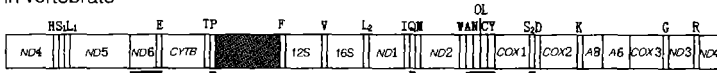


图 1 脊椎动物线粒体基因排列顺序的变化

rRNA 基因由相应氨基酸的单字母密码子表示, L<sub>1</sub>、L<sub>2</sub>、S<sub>1</sub>、S<sub>2</sub> 分别代表 rRNA<sup>Leu(CUN)</sup>、rRNA<sup>Leu(UUR)</sup>、rRNA<sup>Ser(AGY)</sup> 和 rRNA<sup>Ser(UCN)</sup>, P'、M' 分别表示 rRNA<sup>P<sup>ab</sup></sup> 和 rRNA<sup>M<sup>et</sup></sup> 的假基因; OL 表示轻链复制起始位点; ψCR 表示假控制区; 灰色区域为控制区; 斜线标注与典型排布相比, 排列位置发生变化的基因; 下划线标注轻链编码基因; 星号标注存在基因倒置现象的基因组。图中各基因长度不等于原基因组中比例。

Fig 1 Variations in gene order of the vertebrate mtDNA

rRNAs are indicated by their standard one-letter amino acid code. L<sub>1</sub>, L<sub>2</sub>, S<sub>1</sub> and S<sub>2</sub> depict rRNA<sup>Leu(CUN)</sup>, rRNA<sup>Leu(UUR)</sup>, rRNA<sup>Ser(AGY)</sup> and rRNA<sup>Ser(UCN)</sup>, respectively. The italic P' and M' represent pseudo rRNA<sup>P<sup>ab</sup></sup> and pseudo rRNA<sup>M<sup>et</sup></sup> separately. OL is abbreviated for origin of light-strand replication. ψCR designates pseudo-control region. Grey boxes are control region. Boxes with diagonal are the genes whose positions have been changed, comparing with the typical arrangement. Boxes underlined are the genes encoded by the light-strand. Asterisks mark the genome with gene inversion. The sizes of the boxes are not equal to the scale of true genes.

入控制序列的同时, 其 rRNA<sup>M<sup>et</sup></sup> 也移位到控制序列的 3 端。第二类是 rRNA<sup>P<sup>ab</sup></sup> 的缺失, 典型代表为月鱼 (*Lampris guttatus*) 和灯笼鱼 (*Diaphus splendidus*)。此外月鱼还存在控制序列的缺失, 灯笼鱼则有 rRNA<sup>G<sup>lu</sup></sup> 和 rRNA<sup>M<sup>et</sup></sup> 的位置互换。第三类是 ND 6

rRNA<sup>G<sup>lu</sup></sup> 和 rRNA<sup>P<sup>ab</sup></sup> 往控制序列内的移位。芒光灯笼鱼 (*Mycotophum affine*) 和星康吉鳗 mtDNA 中都发现此现象。所不同的是, 这 3 个基因紧密排列, 移入芒光灯笼鱼仅 300 bp 的控制序列, 将其分隔为两部分; 而全长近 3 000 bp 的星康吉鳗控制区序列则被

它们分成了 3 段, ND 6 和两个紧密排列的 rRNA 基因间隔了 607 bp。第四类是部分控制区序列往 rRNA<sup>Trp</sup> / rRNA<sup>Ala</sup> / rRNA<sup>Asp</sup> / rRNA<sup>Cys</sup> / rRNA<sup>Tyr</sup> 基因簇的移位。在蝰鱼 (*Chauliodon subani*) 中该段控制序列移至 rRNA<sup>Tyr</sup> 的 3' 端, 并为随后移入的 rRNA<sup>Cys</sup> 分隔成两部分。而在以深海鲑鱼 (*Bathylagus ochotensis*) 为代表的另 5 种鱼中, 该段控制区序列则移入 rRNA<sup>Trp</sup> 和 rRNA<sup>Ala</sup> 之间。鳗鲡目中的宽咽鱼 (*Euryparynx pekanoi*) 和囊咽鱼 (*Saccopharynx lavenbergi*) 共享第 5 类排列方式: rRNA<sup>Trp</sup>, A 8 / A 6 COX 3 / rRNA<sup>Gly</sup> / ND 3 rRNA<sup>Leu(CUN)</sup> / ND 5 / ND 6 rRNA<sup>Thr</sup> / rRNA<sup>Pro</sup> 和 rRNA<sup>Ile</sup> 这 5 个来自线粒体基因组不同位置的基因簇或基因一起转移到 rRNA<sup>Trp</sup> 和 rRNA<sup>Ala</sup> 间, 并按上述顺序紧密排列在一起 (图 1, B)。

### 1. 1. 3 两栖纲

两栖类中有尾目和蚓螈目动物的 mDNA 呈现典型的脊椎动物排布方式, 基因重排现象目前仅见于蛙科动物中。泽蛙 (*Felvaria linnocharis*) 和黑斑蛙 (*Rana nigromaculata*) 的 rRNA<sup>Thr</sup>, rRNA<sup>Pro</sup> 和 rRNA<sup>Leu(CUN)</sup> 都转移到 rRNA<sup>Phe</sup> 5' 端, 而且前者的 ND 5 基因也移位到控制序列 3' 端, 同时 rRNA<sup>Met</sup> 基因还有一个串联重复的拷贝 (图 1, C)。在测得部分序列的美洲牛蛙 (*Rana catesbeiana*) 线粒体基因组中也存在着与黑斑蛙一致的移位现象<sup>[18]</sup>, 该现象可能是无尾两栖类的 mDNA 基因排布特征。

### 1. 1. 4 爬行纲

3 种已知 mDNA 全序列的鳄类密河鳄 (*Alligator mississippiensis*)、扬子鳄 (*Alligator sinensis*) 和凯门鳄 (*Caiman crocodiles*) 具有相同的基因排列顺序, 与脊椎动物的典型排布方式相比, rRNA<sup>His</sup> 和 rRNA<sup>Ser(AGY)</sup> 基因的位置发生了互换, 同时 rRNA<sup>Phe</sup> 转移到 rRNA<sup>Pro</sup> 的 3' 端 (图 1, D)。

无毒链蛇 (*Diodon semicarinatus*) 是目前唯一测得 mDNA 全序列的蛇类, 其线粒体基因组 rRNA<sup>Ile</sup> 和 rRNA<sup>Gly</sup> 之间插入了一段与其控制区序列高度相似的片段, 随之插入的还有位于该段 5' 端的 rRNA<sup>Pro</sup> 假基因和 3' 端的 rRNA<sup>Leu(UUR)</sup> 基因 (图 1, D)。Kumazawa 等<sup>[19-21]</sup> 曾对另 4 种蛇: 颞窝毒蛇 (*Ovophis okinavensis*)、响尾蛇 (*Crotalus viridis*)、蟒蛇 (*Boa constrictor*) 和德克萨斯盲蛇 (*Leptotyphlops dubois*) mDNA 上 3 个区域的基因簇进行了测序, 结果表明前 3 种与无毒链蛇有相似的基因排序变化, 只是颞窝毒蛇在插入片段的 3' 端还有一个

rRNA<sup>Phe</sup> 拷贝, 蟒蛇 5' 端没有 rRNA<sup>Pro</sup> 假基因。而德克萨斯盲蛇在上述区域未见插入片段, 其 rRNA<sup>Ile</sup> 和 rRNA<sup>Met</sup> 之间的 rRNA<sup>Gly</sup> 则移位到 rRNA<sup>Ala</sup> 和 rRNA<sup>Asn</sup> 之间。

石龙子 (*Eumeces egregius*) 的 rRNA<sup>Thr</sup> 和 rRNA<sup>Pro</sup> 一起移位到控制序列中, 将其分隔为两部分。这与鱼纲中第一种类型的变化方式较为相似 (图 1, D)。楔齿蜥 (*Sphenodon punctatus*) 线粒体基因重排有 3 个特点: 一是基因的迁移, 即 rRNA<sup>Leu(CUN)</sup>, rRNA<sup>Pro</sup>, rRNA<sup>Ser(AGY)</sup> 和 Cyt b 一起迁至控制序列中; 二是两个 rRNA<sup>Lys</sup> 基因串联排列; 三是 ND 5 rRNA<sup>Thr</sup> 和 rRNA<sup>His</sup> 缺失 (图 1, D)。

### 1. 1. 5 鸟纲

Mindell 等根据 5 种鸟类的 mDNA 全序列和代表 13 个目的其他 132 种鸟类的 mDNA 部分序列, 将鸟类的线粒体基因排布方式分成了两类, 分别称为“gene order B”和“gene order C”<sup>[22]</sup>, 与脊椎动物典型布方式相比, 它们在 Cyt b 至 12S rRNA 的区域上具有独特的基因排列顺序: B 型顺序为 ND 5 / Cyt b / rRNA<sup>Thr</sup> / rRNA<sup>Pro</sup> / ND 6 / rRNA<sup>Gly</sup> 控制区; 而 C 型顺序为 ND 5 / Cyt b / rRNA<sup>Thr</sup> 控制区 / rRNA<sup>Pro</sup> / ND 6 / rRNA<sup>Gly</sup> 假控制区 (pseudo-control region)。随后公布的各种鸟类的 mDNA 全序列或部分序列, 也都可归为上述两种排布方式<sup>[23-25]</sup>。迄今, GenBank 上公布了 mDNA 全序列的鸟类共 26 种, 仅在凤头 (*Eudromia elegans*) 中观察到了第三种排布方式: ND 6 和 rRNA<sup>Gly</sup> 的缺失 (图 1, E)。

### 1. 1. 6 哺乳纲

哺乳动物 mDNA 多数呈现典型排布方式, 但在 5 种有袋类动物 (marsupials) 的线粒体基因中, 轻链两侧的 4 个 rRNA 基因发生了顺序上的变化, 由原先的 rRNA<sup>Trp</sup> / rRNA<sup>Ala</sup> / rRNA<sup>Asp</sup> / rRNA<sup>Cys</sup> 排列顺序变为了 rRNA<sup>Ala</sup> / rRNA<sup>Cys</sup> / rRNA<sup>Trp</sup> / rRNA<sup>Asp</sup> (图 1, F)。

日本鼯鼠 (*Urotrichus talpoides*) 缺失 rRNA<sup>Ile</sup>; 南美獭 (*Tapirus tennestris*) 缺失 rRNA<sup>Gly</sup>, 但在该基因原位置上仍有 70bp 非编码序列; 新西兰海狮 (*Arctocephalus forsteri*) 没有控制序列; 大西洋鼠海豚 (*Phocoena phocoena*) 的 rRNA<sup>Gly</sup> 从 COX 3 和 ND 3 之间移至 16S rRNA 的 3' 端, 其他基因顺序不变。

### 1. 1. 7 基因顺序变化类型及特点

目前为止, 观察到的脊椎动物 mDNA 基因顺序的变化可以分成 3 类: 1) 邻接的基因或片段的位

置交换,如: 鳄类的 S-H-L 基因簇; 2) 接近于控制序列或轻链起始位点的基因或片段的位置变化, 有时还伴随着控制序列的倍增, 如: 鸟类的 ND 6 至假控制区; 3) IQ-M 区域的变化, 这常见于蛇类。这些基因顺序的变化又具有以下两个特点: 1) 所有的基因顺序变化均涉及 rRNA 基因, 其中多数变化方式也只涉及少数 rRNA 位置的改变; 2) 已知的所有鳄类、蛇类、鸟类和有袋类的线粒体基因组都存在基因顺序上的变化。

## 1.2 基因的倒置

脊椎动物 mDNA 基因重排中还存在着另一种现象, 即基因从一条链编码转移到另一链上编码。这种基因编码链的转换, 会导致其阅读方向的改变, 在此称其为基因倒置 (inversion)。由图 2 可见, 基

因倒置在脊椎动物线粒体基因组中较普遍, 除两栖纲外, 各纲中都可观察到, 尤其多见于哺乳纲和鱼纲。除少数特例, 所有倒置现象都具有以下几个特点: 1) 绝大多数倒置只涉及 rRNA 基因的变化。大西洋鼠海豚 (*Phocoena phocoena*) 是迄今在脊椎动物中观察到的唯一特例, 其 16S rRNA 由重链编码变为轻链编码。2) 发生倒置的 rRNA 基因一般都是由轻链编码转变为重链编码。但非洲象 (*Loxodonta africana*) 的 rRNA<sup>lt</sup>, 叉尾鲷 (*Etalurus punctatus*) 的 rRNA<sup>Leu(CUN)</sup> 和星康吉鳗的 rRNA<sup>Thr</sup> 都由重链迁至轻链。3) 发生倒置的基因在基因组上的排列位置没有发生改变。虽然在 28 个观察到基因倒置现象的物种中, 有 9 种 (原口纲 1 种, 鱼纲 2 种, 鸟纲 4 种, 哺乳纲 2 种) 同时存在线粒体基因顺序上的变化, 但位置变化的基因和倒置基因各不相同。

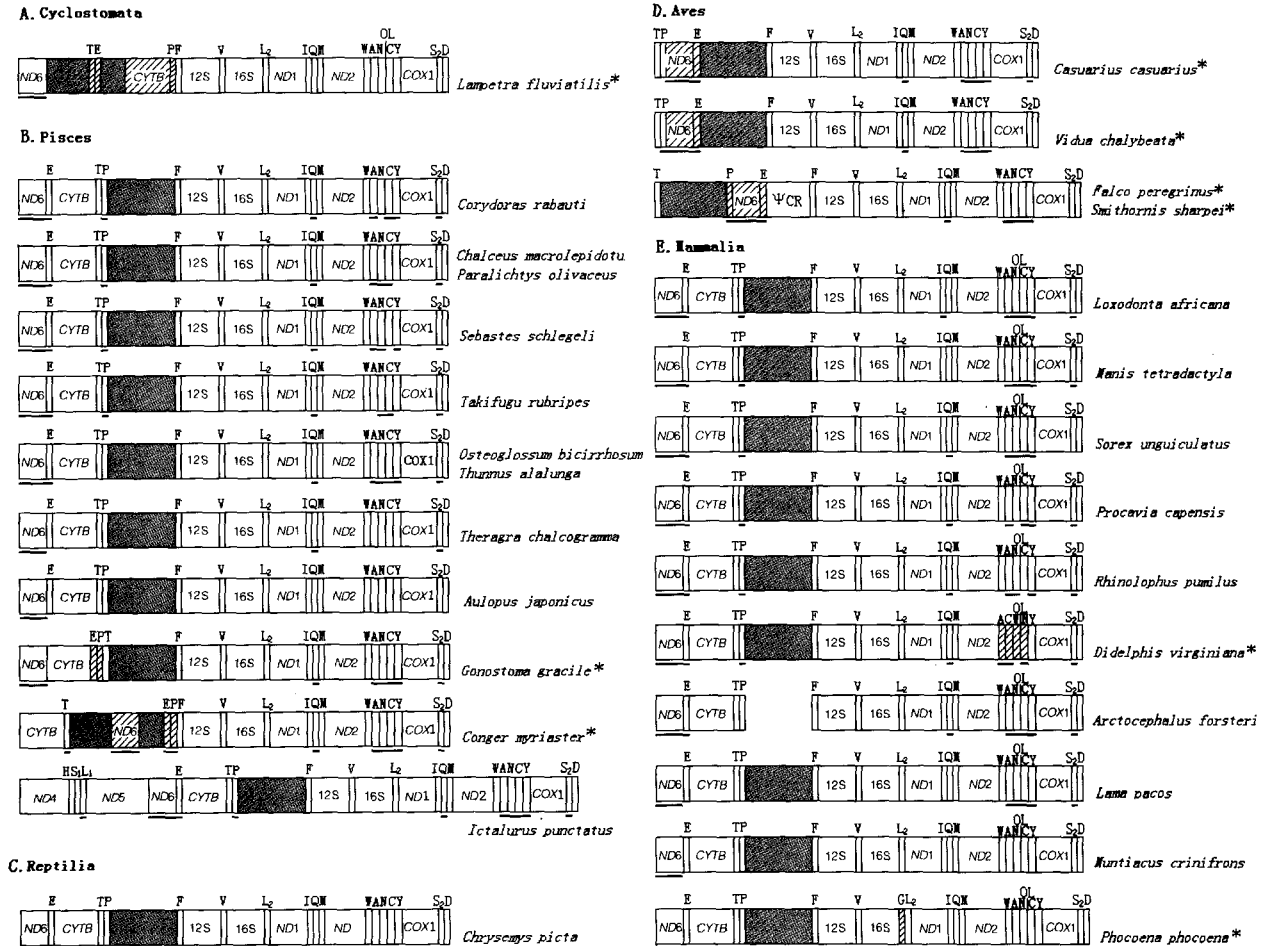


图 2 脊椎动物线粒体基因的倒置现象

星号标注的表示该基因组中同时存在基因顺序的变化, 其他符号同图 1。

Fig. 2 Inversion of the vertebrate mitochondrial genes

All symbols are depicted as in Fig. 1 with the exception of asterisks here indicating the genome with variation of gene order

另外一个值得注意的现象是, 在 3 个不同纲的物种中观察到同一种特殊的基因倒置现象, 即除 ND 6 外所有基因均由重链编码。这 3 种动物分别是圆口纲的欧洲河鳗, 鱼纲的日本比女鱼 (*Ambly japonicus*) 和哺乳纲的黑鹿 (*Muntiacus crinifrons*)。

## 2 脊椎动物线粒体基因重排的机制

### 2.1 倍增-随机删除模型

目前用于解释脊椎动物线粒体基因重排机制的最常用模型是倍增-随机删除模型 (Duplication-random deletion)<sup>[1, 13, 15, 26, 27]</sup>, 该模型认为含有两个或多个基因的 mDNA 片段发生倍增, 倍增后的基因被随机删除富余拷贝。这一过程可能产生一些假基因, 随后逐步丢失所有基因残迹, 结果可导致基因移位。但基因倍增并不是随机的, 有一种位置效应, 重排的热点多分布在两个复制起始位点附近, 仅涉及周边的基因, 并伴随控制区序列的倍增<sup>[27]</sup>。

关于线粒体基因倍增的机制有许多假说。有人猜测环状 mDNA 在复制过程中的不准确终止, 导致基因倍增<sup>[1, 28]</sup>; 另一种解释认为是轻链复制错误引起基因倍增<sup>[15, 29]</sup>, 起始轻链复制的酶可以识别 rRNA 的茎环结构, 并以此作为轻链复制的起始位点, 而终止信号仍保留在原有位置, 结果导致起始 rRNA 和终止信号间基因的倍增; 当前最普遍认可的是滑链错配 (slipped-strand mispairing) 模型<sup>[1, 15, 27]</sup>, 该模型认为复制起始后新生链的 3' 端向 5' 端滑动, 重复复制后导致相关基因的串联倍增。大多数发生线粒体基因重排的脊椎动物谱系, 如鸟类、鳄类、黑斑蛙、星康吉鳗等, 都可以用倍增-随机删除模型来解释<sup>[13, 15-17]</sup>。

Macey 等首次报道了爬行动物蚓蜥 (*Bipes biporus*) 的线粒体基因组存在 rRNA<sup>Thr</sup> 和 rRNA<sup>Pro</sup> 的串联重复拷贝, 并确认了其中的假基因<sup>[30]</sup>。随后, 楔齿蜥线粒体基因组中 rRNA<sup>Lys</sup> 和泽蛙线粒体基因组中 rRNA<sup>Met</sup> 的串联重复拷贝也相继公布<sup>[31, 32]</sup>, 这种串联重复拷贝的发现被视为倍增-随机删除模型的证据。然而, 该模型只是在 rRNA 中找到了串联重复的例证, 对于那些涉及较大的蛋白编码基因和一些发生倒置的基因重排仍不能很好的解释, 同时也无法解释少数部分片段倍增后的 mDNA 如何在

群体中最终占据主导地位。

### 2.2 重组模型

重组是在植物、真菌和原生物中导致 mDNA 变化的较常见机制, 曾认为不存在于动物 mDNA 中<sup>[28]</sup>。但近年来的研究表明, 重组可能为基因倒置提供了一个合理的解释。1997 年, Lunt 和 Hyman<sup>[28]</sup> 首次用实验验证了动物 mDNA 的重组, 他们在研究线虫线粒体基因组控制区串联重复序列时, 发现其存在明显的拷贝数目变化, 并且观察到重复序列的倒置现象, 证实这是重组的结果, 强调重组的关键在于双链断裂和重新结合: 在 mDNA 上两相邻位点可能同时断裂, 断裂产生的双链片段通过共价键作用使其两个断裂端重新结合, 从而形成一种游离于原基因组的、具有部分基因组结构的小型双链环状分子 (minicircle), 当某段序列形成 minicircle 离开原基因组时, 该基因组相应的组分也将随之丢失。楔齿蜥和凤头的线粒体基因组部分基因丢失的现象亦可作为该模型的一个佐证。2000 年, Kajander 等<sup>[33]</sup> 在人体细胞中发现了低水平分布的 sublimons, 这是一种因部分序列缺失或倍增所产生的重排 mDNA 分子 (rearranged mDNA molecules), 推测为 mDNA 双链分子断裂引发线粒体内 DNA 重组的结果。他们的依据是在这些重排的 mDNA 上找到了一系列可同时发生断裂的位点, 具有这些断裂位点的基因组往往出现部分倍增或缺失, 形成二聚体甚至多聚体的现象。随后, Mark 和 Nick<sup>[34]</sup> 整合了上述两个实验的成果, 提出一个新的模型, 用以解释基因倒置现象, 其论点仍是双链的断裂和重新结合, 只是断裂产生的 mDNA 双链片段不再自身环化脱离原基因组, 而是在相应酶的作用下反向接回断裂位点, 致使基因倒置。2001 年, Kajander 等<sup>[35]</sup> 在研究人类线粒体的病理变化时, 意外的观察到了线粒体内 DNA 重组时的中间状态, 即 Holliday 连接体。自此, mDNA 重组被越来越多地用于解释无脊椎动物谱系中的基因倒置现象<sup>[36, 37]</sup>。重组不仅可用来解释基因的缺失和倒置现象, 也同样适用于解释基因排列顺序的改变。如果一个完整的线粒体基因组有 3 个或 3 个以上的位点同时断裂, 那么由此产生的断裂片段在重新接回原基因组时, 就有可能发生连接顺序的改变, 从而导致邻接序列排列顺序的变化。与倍增-随机删除模型相比, 这可能更适合于解释长的邻接片段间的

位置互换现象,如鱼纲中第 4类基因排列顺序的变化。

### 3 mDNA 基因重排在系统发生分析上的应用

由于基因排布上的变化类型多样,使得物种间同样的基因排布一般不可能是趋同进化的结果,因此认为相同的基因排布方式,来自共同的祖先。线粒体基因的排布可为评估脊椎动物系统发生关系提供重要信息,尤其较高分类阶元间的系统发生关系上,基因排列的比较将是解决具有争议的系统发生关系的有力工具<sup>[27,30,38-40]</sup>。

Boore等<sup>[41]</sup>发现头索动物亚门的佛罗里达文昌鱼(*Banchistoma fibrilae*) mDNA缺少脊椎动物中的控制序列,基因排布方式与脊椎动物的典型排布基本相同,仅在 rRNA<sup>Gly</sup>、rRNA<sup>Phe</sup>、rRNA<sup>Met</sup>和 rRNA<sup>Asp</sup>这 4个基因的位置上有所差异。通过与果蝇、棘皮动物和半索动物 mDNA 基因排布方式的比较,推测 rRNA<sup>Gly</sup>、rRNA<sup>Met</sup>和 rRNA<sup>Phe</sup>在脊椎动物 mDNA 上排布方式反映了脊索动物最初的线粒体基因排布状态,而文昌鱼 mDNA 中 rRNA<sup>Asp</sup>的位置反映了它们共同祖先的特征。由此推断,脊椎动物在脊索动物祖先 N-W-A-C-Y 的基础上变为 W-A-N-C-Y。

Inoue等<sup>[16]</sup>测定了星康吉鳗 mDNA 全序列,发现基因重排现象,并在此基础上测定、对比了鳗鲡目另外 11 科的代表物种的 ND 5 至 Cytb 区域的基因序列,结果表明有 4 个与星康吉鳗同属一个亚目的物种 mDNA 基因排列顺序与星康吉鳗相同,提示这 5 个科动物起源于一个共同的祖先,这一结果在用 12S rRNA 和 Cytb 基因序列进行的系统发生分析中得到进一步证实。

传统的形态学分类将鹰科与隼科同归为隼形目动物,但根据近年来的一些研究结果又有人对此提出了质疑。为弄清鹰科与隼科的亲缘关系,确定它们在系统发生树上的位置,Haring等<sup>[42]</sup>测定了鹰科鸟类普通鵟的 mDNA 全序列,发现其基因排布顺序与游隼完全一致,同属本文中 C 型排布方式,并且这两种鸟类在假控制区上都有一长的串联重复片段,这一特征又使它们区别于其他共享着 C 型排布方式的鸟类。作者进一步根据公布的 mDNA 全序列,对包括上述 2 种在内的 10 种鸟类进行系统进化

分析,分析结果都支持鹰科与隼科的传统分类学地位。

Kumazawa 将<sup>[19,43]</sup> mDNA 上某种新型基因顺序的有无模式用于爬行类动物系统进化分析,根据这一模式,他将鳄类与鸟类归为同一分支:因为所有鳄类和鸟类都没有轻链复制起始所需要的 rRNA<sup>Asn</sup>和 rRNA<sup>Cys</sup>间的茎环结构,而蛇类(除德克萨斯盲蛇外)、蜥蜴类、蛙类和哺乳类都有这种茎环结构,这正好印证了鸟类和爬行类系统发生关系的学说:鸟类和鳄类为姐妹群,构成了古龙类(Archosauria)的分支,而蛇类和蜥蜴类合称鳞龙类(Lepidosauria)或有鳞类(Squamata)。另一方面,鳄类有一独特系统发生标志,即 rRNA<sup>His</sup>和 rRNA<sup>Ser(AGY)</sup>基因位置的互换;而在蛇类动物中普遍存在着一种新型排列方式,即 tQ-M 基因簇中插入了 rRNA<sup>Lys</sup>和一段近似于控制序列的非编码序列,仅德克萨斯盲蛇还保留了 tM 基因簇,此特征加上缺失的茎环结构,使得德克萨斯盲蛇与其他蛇类明显区分开来,这也符合其作为最原始蛇种的分类地位。由此也可推测,涉及 tQ-M 基因簇的多种重排现象起源于蛇系发生早期。由于乌龟没有轻链复制起始的茎环机构,也没有上述任何一种基因重排现象,因而龟类的系统位置问题一直都是生物学家争论的焦点。

### 4 展望

迄今人们只对少数后口动物 mDNA 进行了基因排布及其分子机制的研究,而且用于基因序列研究的分类群的取样具有很高的偏好性;另一方面,尽管已测得多种 mDNA 的部分序列,但大多数种系仍旧无法用线粒体基因排布信息来代表。这些都是利用 mDNA 的排布信息来重建系统进化关系的制约因素,克服这些因素的关键在于建立完整、广泛、具有代表性的系统信息资料库,以便更好地通过复杂的共享重排信息来鉴定生物的进化分支。人们已在考虑使用计算机重建基因重排<sup>[1,27]</sup>,但还未找到一个算法公式能详尽搜寻所有可能解决方案,并保证得到一个精确且简捷的重建方法。而人们对引起 mDNA 基因重排的分子过程和机制,发生重排基因的偏好性,倍增序列形成的特殊二级结构在基因删除中的作用等都所知无几,因而这方面仍有大量的工作有待深入研究。

今后线粒体基因组所要致力于解决的问题可能

存在以下几个方面: 1) 哪类基因最常重排? 即重排的热点及其在进化生物学上的意义。2) 还有哪些可能的基因重排机制? mtDNA 重排现象是由多种重排机制还是由某种单一的机制产生的? 3) 基因重排与生物体的生理状态或生活史有何关联? 4) 在同一谱系中, 序列的进化速率与基因的重排频率有关吗? 即碱基的替换速率对基因排布方式的潜在影响。

## 参考文献 (References):

[ 1 ] Jeffrey L B. Survey and summary animal mitochondrial genomes *Nucleic Acids Res*, 1999, 27(8): 1767~ 1780

[ 2 ] Taarman J W. The mitochondrial genome structure transcription, translation and replication *Biochim Biophys Acta*, 1999, 1410(2): 103~ 123

[ 3 ] Betts J, Lightowers R N, Turnbull D M. Neuropathological aspects of mitochondrial DNA disease *Neurochem Res*, 2004, 29(3): 505~ 511.

[ 4 ] Bossy-Wetzel E, Barsom M J, Godzik A, Mitochondrial R, Lipson S. Mitochondrial fission in apoptosis, neurodegeneration and aging *Schwarzenbacher Curr Opin Cell Biol*, 2003, 15(6): 706~ 716

[ 5 ] Roe B A, Din-Pow M, Wilson R K, Wong J F. The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome *J Biol Chem*, 1985, 260(17): 9759~ 9774

[ 6 ] Babb M J, Van Etten R A, Wright C T, Walberg M W, Clayton D A. Sequence and gene organization of mouse mitochondrial DNA *Cell*, 1981, 26(17): 167~ 180

[ 7 ] Anderson S, de Bruijn M H, Coulson A R, Eperon I C, Sanger F, Young I G. Complete sequence of bovine mitochondrial DNA conserved features of the mammalian mitochondrial genome *J Mol Biol*, 1982, 156(4): 683~ 717.

[ 8 ] Anderson S, Bankier A T, Barrell B G, Bruijn M H L, Coulson A R, Drouin J, Eperon I C, Nierlich D P, Roe B A, Sanger F, Schreier P H, Smith A J H, Staden R, Young I G. Sequence and organization of the human mitochondrial genome *Nature*, 1981, 290(5806): 457~ 465

[ 9 ] Johansen S, Guddal P H, Johansen T. Organization of the mitochondrial genome of Atlantic cod *Gadus morhua*. *Nucleic Acids Res*, 1990, 18(3): 411~ 419

[ 10 ] Tzeng C S, Hui C F, Shen S C, Huang P C. The complete nucleotide sequence of the *Crossostomus xanthurus* mitochondrial genome: conservation and variations among vertebrates *Nucleic Acids Res*, 1992, 20(18): 4853~ 4858

[ 11 ] Chang Y S, Huang F C, Lo T B. The complete nucleotide sequence and gene organization of carp (*Cyprinus carpio*) mitochondrial genome *J Mol Evol* 1994, 38(2): 138~ 155

[ 12 ] Zardoya R, Meyer A. The complete nucleotide sequence of

the mitochondrial genome of the lungfish (*Protopterus dolloi*) supports its phylogenetic position as a close relative of land vertebrates *Genetics*, 1996, 142(4): 1249~ 1263

[ 13 ] Sumida M, Kanamori Y, Kaneda H, Kato Y. Complete nucleotide sequence and rearrangement of the mitochondrial genome of the Japanese pond frog *Rana nigromaculata*. *Genes Genet Syst*, 2001, 76(5): 311~ 325

[ 14 ] Amdt A, Smith M J. Mitochondrial gene rearrangement in the sea cucumber genus *Cucumberia*. *Mol Biol Evol*, 1998, 15(8): 1009~ 1016

[ 15 ] Macey J R, Larson A, Ananjeva N B, Fang Z, Papenfuss T J. Two novel gene order and the role of light strand replication in rearrangement of the vertebrate mitochondrial genome *Mol Biol Evol*, 1997, 14(1): 91~ 104

[ 16 ] Inoue G, Miyama M, Tsukamoto K, Nishida M. Complete mitochondrial DNA sequence of *Conger myriaster* (Teleostei: Anguilliformes): novel gene order for vertebrate mitochondrial genomes and the phylogenetic implications for anguilliform families *J Mol Evol*, 2001, 52(4): 311~ 320

[ 17 ] Miya M, Nishida M. Organization of the Mitochondrial Genome of a Deep-Sea Fish *Gonosoma gracile* (Teleostei: Scombriformes): First Example of Transfer RNA Gene Rearrangements in Bony Fishes *Mar Biotechnol*, 1999, 1(5): 416~ 426

[ 18 ] Yoneyama Y. The nucleotide sequences of the heavy and light strand replication origins of the *Rana catesbeiana* mitochondrial genomes *J Nippon Med Sch*, 1987, 54(4): 429~ 440

[ 19 ] Kumazawa Y, Nishida M. Variations in mitochondrial tRNA gene organization of reptiles as phylogenetic markers *Mol Biol Evol* 1995, 12(5): 759~ 772

[ 20 ] Kumazawa Y, Ota H, Nishida M, Ozawa T. Gene rearrangements in snake mitochondrial genomes: highly concerted evolution of control-region-like sequences duplicated and inserted into a tRNA gene cluster *Mol Biol Evol*, 1996, 13(9): 1242~ 1254

[ 21 ] Kumazawa Y, Ota H, Nishida M, Ozawa T. The complete nucleotide sequence of a snake (*Dhond semicarinatus*) mitochondrial genome with two identical control region *Genetics*, 1998, 150(1): 313~ 329

[ 22 ] Mindel D P, Sorenson M D, Dimcheff D E. Multiple independent origins of mitochondrial gene order in birds *Proc Natl Acad Sci USA*, 1998, 95(18): 10693~ 10697

[ 23 ] Mindel D P, Sorenson M D, Dimcheff D E, Hasegawa M, Ast J C, Yuri T. Interordinal relationships of birds and other reptiles based on whole mitochondrial genomes *Syst Biol* 1999, 48(1): 138~ 152

[ 24 ] Haddad O, Baker A J. Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogenetics and the vicariance biogeography hypothesis *Proc R Soc*



- Lond B B 01Sci 2001, 268(1470): 939~ 945
- [ 25] Paton T, Haddad O, Baker A J Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds Proc R Soc Lond B B 01Sci, 2001, 269(1493): 839~ 846
- [ 26] Moritz C, Dowling T E, Brown W M. Evolution of animal mitochondrial DNA: relevance for population biology and systematics Annu Rev Eco Syst, 1987, 18: 269~ 292
- [ 27] Jeffrey L B, Wesley M B Big trees from little genomes mitochondrial gene order as a phylogenetic tool Curr Opin Genet Dev, 1998, 8(6): 668~ 674
- [ 28] David H L, Bradly C H Animal mitochondrial DNA recombination Nature, 1997, 387(6630): 247
- [ 29] Stanton D J, Daehler L L, Moritz C C, Brown W M. Sequence with the potential to form stem-and-loop structures are associated with coding-region duplications in animal mitochondrial DNA. Genetics, 1994, 137(1): 233~ 241.
- [ 30] Macey J R, Schulte J A, Larson A, Papenfuss T J Tandem duplication via light-strand synthesis may provide a precursor for mitochondrial genome rearrangement Mol Biol Evol 1998, 15(1): 71~ 75
- [ 31] Rest J S, Ast J C, Austin C C, Waddell P J, Tibbitts E A, Hay J M, Mindell D P. Molecular systematics of primary reptilian lineages and the tuatara mitochondrial genome Mol Phylogenet Evol 2003, 29(2): 289~ 297.
- [ 32] Liu Z Q, Wang Y Q, Su B. The mitochondrial genome organization of the Chinese rice frog, *Fejervarya limnocharis*: a new gene order in the vertebrate mtDNA Gene, 2005, 346: 145~ 151
- [ 33] Kajander O A, Rovio A T, Mämaa K, Poulton J, Spelbrink J N, Holt I J, Karhunen P J, Jacobs H T. Human mtDNA sublineages resemble rearranged mitochondrial genomes found in pathological states Hum Mol Genet, 2000, 9(19): 2821~ 2835.
- [ 34] Downton M, Campbell N J H. Intramitochondrial recombination: is it why some mitochondrial genes sleep around? Trends Ecol Evol 2001, 16(6): 269~ 271.
- [ 35] Kajander O A, Karhunen P J, Holt I J, Jacobs H T. Prominent mitochondrial DNA recombination intermediates in human heart muscle EMBO Rep, 2001, 2(11): 1007~ 1012
- [ 36] Shao R, Barker S C. The highly rearranged mitochondrial genome of the baguë thrips *Thrips in agines* (Insecta: Thysanoptera): Convergence of two novel gene boundaries and an extraordinary arrangement of rRNA genes Mol Biol Evol 2003, 20(3): 362~ 370
- [ 37] Downton M, Lyda R C, Campbell S L, Bargon S D, Austin A D. Frequent mitochondrial gene rearrangements at the Hymenopteran nad3-nad5 junction J Mol Evol 2003, 56(5): 517~ 526
- [ 38] Smith M J, Amtdt A, Gorski S, Faerber E. The phylogeny of echinoderm classes based on mitochondrial gene arrangement J Mol Evol, 1993, 36(6): 545~ 554.
- [ 39] Boore J L, Brown W M. Complete DNA sequence of the mitochondrial sequence of the black chiton, *Katharina tunicata*. Genetics, 1994, 138(2): 423~ 443
- [ 40] Janke A, Feldmaier-Fuchs G, Thomas W K, Haesler A, Paabo S. The marsupial mitochondrial genome and the evolution of placental mammals Genetics, 1994, 137(1): 243~ 256
- [ 41] Boore J L, Daehler L L, Brown W M. Complete Sequence, Gene Arrangement, and Genetic Code of Mitochondrial DNA of the Cephalochordate *Banchios tomia fibridae* (Amphioxus). Mol Biol Evol 1999, 16(7): 410~ 418
- [ 42] Haring E, Kruckenhauser L, Gamau A, Resing M J, Pinner W. The complete sequence of the mitochondrial genome of *Buteo buteo* (Aves: Accipitridae) indicates an early split in the phylogeny of raptors Mol Biol Evol 2001, 18(10): 1892~ 1904.
- [ 43] Kumazawa Y, Nishida M. Complete mitochondrial DNA sequences of the green turtle and blue-tailed mole skink: statistical evidence for archosaurian affinity of turtles Mol Biol Evol 1999, 16(6): 784~ 792