

动物线粒体假基因的识别及其在进化生物学中的应用*

王继文

(四川农业大学动物科技学院 雅安 625014)

摘要: 在真菌、昆虫、无脊椎动物和脊椎动物等许多分类单元中,都已发现 mtDNA 序列的核转座现象。在 PCR 扩增时,往往同时扩增出 mtDNA 和细胞核中线粒体假基因(Numts),Numts 混淆系统发育和群体遗传研究,得出错误的结果。本文综述了 Numts 的检查和避免的方法,以及在进化生物学研究中的应用进展。

关键词: 线粒体假基因;进化生物学

中图分类号: Q953 **文献标识码:** A **文章编号:** 0250-3263(2004)03-103-06

Numts Identification and Utility in Evolutionary Biology

WANG Ji-Wen

(College of Animal Science and Technology, Sichuan Agricultural University, Yaan 625014, China)

Abstract: The transposition of mtDNA sequences to the nucleus has been documented in a wide variety of taxa from fungi to insects to vertebrates. If co-amplified with mtDNA, Numts greatly complicate the accurate determination of mitochondrial sequences. The undiscovered presence of Numts could lead to incorrect results. Methods for checking and avoiding Numts have now been reviewed. However, several recent studies have demonstrated the potential utility of Numt DNA sequences serving as valuable molecular tools in evolutionary biology studies.

Key words: Numts; Evolutionary Biology

线粒体 DNA(mtDNA)具有快速进化和母系遗传等特点^[1,2],因此,线粒体 DNA 标记和以聚合酶链式反应(PCR)为基础的 DNA 测序技术的联合使用为鸟类系统学^[3]和群体遗传结构^[4]分析提供了新的机会。mtDNA 在进化研究中的应用已经在一系列综述性文献中作了详细讨论,同时也涉及到几个相关的问题,包括父系遗传、异质性、进化异速性和非中性^[5,6]。然而,从总基因组 DNA 中用 PCR 扩增 mtDNA 片段总是会遇到一个极大的问题:扩增 mtDNA 的同时可能扩增出起源于线粒体假基因 Numts (nuclear mitochondrial-like sequences, Numts,专指线粒体起源的核序列)^[7,8],而不单是 mtDNA 上的靶序列。在对 PCR 产物进行电泳分析时,往往有非特异性条带出现,增加背景噪音,对测序非常不利,不仅增加研究工作的难度,有时甚至会得出错误的结论^[9],例如,将人的 Numts 误认为恐龙^[10]和古猴^[11]的

mtDNA 序列。

1 检查和避免 Numts 的方法

造成 PCR 产物超过一条似 mtDNA 序列现象的原因有线粒体异质性^[12]、线粒体基因组的复制事件^[13,14]及样品不纯、试剂污染等,但最普遍的还是 Numts 问题。Numts 产生的背景包括 PCR 产物的影子带、杂带^[15]、序列模糊(特别是在多态位点或双链测序时遇到)、移码突变和终止密码子。检查和避免 Numts 的方法主要有以下几种:

1.1 扩增组织的选择 采用 mtDNA/核 DNA 比值高的

* 国家农业科技成果转化资金(No. 02EFN215100520)资助;
第一作者简介 王继文,39岁,男,博士,副教授;研究方向:动物遗传育种;E-mail: wjw2886166@163.com。

收稿日期:2003-06-20, 修回日期:2004-02-10

组织或器官扩增,或扩增前纯化线粒体,可以避免 Numts。Numts 在细胞核中有相当高的拷贝数^[16],相同的 PCR 引物同 mtDNA 序列相比更易与 Numts 结合^[2,17]。当组织中核 DNA 含量相对较少时,如硬的羽茎,作者使用与核 DNA 相比更富含 mtDNA 的组织肝脏和心脏,抽提 mtDNA 前先纯化线粒体(去除细胞膜和核碎片,离心沉淀即为线粒体),家鹅 mtDNA 细胞色素 *b* 和大部分控制区序列能得到良好的扩增*。

1.2 通过反转录 PCR 检查和避免 Numts Numts 能够通过反转录 PCR(RT-PCR)避免^[18],但是偶尔也有 Numts 被转录^[19]。当知道 mtDNA 和 Numts 序列,并且 mtDNA 是单系发生的,就可设计特异性引物,如果 Numts 也是单系发生的,可在 PCR 扩增前采用限制性内切酶消化^[15]。另外,如果采用 PCR 产物克隆测序,有时能够从进化模式上推断出哪些是 mtDNA,哪些是 Numts^[2,20]。

1.3 设计特异性引物 在测序比较研究中,引物设计至关重要。由于 Numts 与 mtDNA 相比进化速度更慢^[21-24],同祖先序列相比更相近,并在相关分类单元间更相似。因此,根据特定物种序列设计引物不易扩增出 Numts。而根据物种分类单元设计的引物和通用引物更容易扩增出 Numts。而且引物位于一致序列位置更接近祖先序列,因而也易于扩增出 Numts。使用非特异 PCR 引物(通用引物和简并引物)时,由于引物与核拷贝有更好的匹配,因而易于优先扩增。这是因为某些有机体核序列比它们线粒体同源部分进化更慢,以至更为保守。这已经在啮齿目动物^[17]、灵长类动物^[11,24]和雪鹑^[4]的研究中使用线粒体 Cyt *b*、12S rRNA 和 D-loop 序列作为诊断标记得到证实。

蛋白质密码子第三碱基的简并性导致编码区极少有 20~25 碱基的高度保守序列群存在,Kwok 等^[25]认为简并性引物更能从组织样中扩增出 mtDNA 序列。密码子的第三碱基使用简并引物在 mtDNA 与核 DNA 高比例的组织中比非对称引物更能得到 PCR 产物。在保守序列氨基酸密码子的第三碱基可能存在变异这一假设是合理的,已在一些已测序的物种中得到了证实。理想的测序方案是根据测序片段的末端序列设计特异性简并引物,但首先应做的工作是纯化 mtDNA^[26]。

1.4 以长链 PCR 产物为模板进行二次 PCR,排除短而模糊的 Numts 条带 扩增整个 mtDNA 基因组或大部分 mtDNA 或 Long-PCR,也是避免 Numts 一个可能有效的方法^[27]。尽管大片段 mtDNA 有可能转座到核基因组^[16],但是从已发表的大多数例子中可知 Numts 长度小于 4

kb^[2],如设计的引物扩增片段长度从 5~16 kb,则扩增出的就很可能仅仅是 mtDNA。即使整个线粒体基因组转座到核基因组,除非 Numts 的始、末段在这一区域内或是串联重复序列,否则扩增出的也是环状的 mtDNA,而不是线性的 Numts。一旦长链 PCR 产物得到后,就能以之为模板进行二次 PCR 扩增目标基因,可排除那些短的、模糊的 Numts 条带。Ruokonen 等^[28]估计 Numts 与不完整的 mtDNA 扩增产物(“跳跃 PCR”^[29])形成的异源双链核酸造成在最终的 PCR 产物上结合了一些 Numts,因而,该方法并不总是有效。

1.5 Numts 和 mtDNA 被同时扩增后的分离措施 当 Numts 和 mtDNA 被同时扩增出来时,有一些技术可以把它们分离。首先,PCR 产物可以先被连接到载体上克隆,然后对不同的克隆进行测序^[21]。其次,可以设计内部引物去区分两条不同的序列^[4,23],选择在两条序列变异最大的地方确定引物。理想的情况是,新引物 3' 端碱基最好选择 mtDNA 和 Numts 序列发生颠换位置的碱基^[30]或至少是不同的位置。这种方法对于研究种内变异比克隆测序更有效。第三,也可以在 PCR 扩增以前对基因组 DNA 用限制性内切酶消化或对 PCR 产物消化,以得到其中一种拷贝的完整 PCR 产物。根据测序结果在碱基模糊的位点确定用何种限制性内切酶。然后对消化后的 DNA 片段琼脂糖电泳,割下各片段,纯化、测序^[23]。尽管这一方法较为有效,但是须知两序列的差异,因而应用有限。一旦得到两条清楚的 DNA 序列,就可以根据上面讨论的标准区分哪个是 Numts,哪个是 mtDNA。

1.6 使用纯化 mtDNA 作探针进行 Southern 杂交 一个更直接但技术上较难的检测 Numts 方法是使用纯化 mtDNA 作探针进行 Southern 杂交。首先从血样或其他组织抽提的 DNA 用限制性内切酶消化,通过琼脂糖电泳使片段分离,然后转移到硝酸纤维素膜或尼龙膜上。在缺乏 Numts 的样品中,纯化的线粒体探针杂交后仅仅检测到 mtDNA 阳性带。如存在 Numts,则额外的带将被检测到。即使 Numts 具有同 mtDNA 一样的限制性内切酶位点,但在末端区域应生成特异长度的条带,这些限制性条带可能就是 Numts 串联重复区域^[16]。

1.7 检查和避免 Numts 的一些其它方法 如果扩增时超过一条似 mtDNA 的序列,可采用限制性酶切片段^[31]、单链构象多态(SSCP)^[32]、毛细管稳定变性电泳(CDCE)^[33]、异源双螺旋分析(HA),变性梯度凝胶电泳(DGGE)或变温梯度凝胶电泳(TGGE)^[2]、克隆 PCR 产物

* 王继文. 中国主要家鹅品种分子系统进化研究. 四川农业大学博士学位论文, 2003.

测序^[7,20]的方法解决。

2 Numts 的应用

正如上面所讨论的一样,核内线粒体序列的存在给 mtDNA 标记的系统发育和群体遗传研究带来麻烦,同时它们也为进化生物学的不同方面研究带来一个新的有力工具。

2.1 作为遗传标记、研究种间(内)变异 Numts 末端序列缺乏一致特征^[19,34],表明插入热点(hot spots)少见^[34]。因此,Numts 在多个分类单元中出现一个特别的突变位点暗示它们具有共同的祖先^[34]。在特异性位点 Numts 的出现与否可被用于确定不同种的系统发育支序^[34,35],作为群体遗传标记已在人^[36]和 *A. thaliana*^[37]中得到了应用。

在豌豆、大麦、小麦、菠菜和甜菜中发现细胞器 DNA 核拷贝数存在种内变异^[38]。拷贝数的变异通过 Southern 杂交能够被检测到,并且能被用于区分商业植物的变异^[38]。因为人^[39]和蝗虫^[7,40]的 Numts 拷贝数存在种内差异,如同其它转座元件一样^[38,41],通过 Southern 杂交或 PCR,Numts 拷贝数的变异也能被用于这些物种的群体遗传标记。

如果同源序列在分类单元间存在多样性,那么一个特异的 Numts 能够作为系统发育标记。相反,核 DNA 突变慢(如哺乳动物或鸟类),较短的 Numts 序列也许显示出极少或没有种内变异^[34,42]。

2.2 核基因组和 mtDNA 基因组的分子进化机制和进化速率 Numts 为比较核基因组和 mtDNA 基因组的分子进化机制和进化速率提供了一条有效的途径^[2,23],并且 Numts 为研究基因组进化和基因组间互作的动力提供了一个窗口^[7]。Numts 的进化速率比核基因组的编码序列快但是比 mtDNA 慢^[43]。mtDNA 和单拷贝核 DNA (single copy nuclear DNA, scnDNA) 进化的相对速率对于它们在进化研究中作为分子标记具有重要意义。哺乳动物中 mtDNA 进化速率比 scnDNA 快好几倍^[44],这一结论在其它多细胞动物中一直存在异议,尤其是在昆虫中存在相当大的争议^[6,45]。处理这一问题的一个困难是昆虫的 mtDNA 和 scnDNA 并不直接相关,它们的进化速率受很多因素影响,诸如群体历史、特定序列的功能约束等。现已知道 mtDNA 和其核插入序列具有共同的祖先,以及大多数线粒体外核序列是非功能的假基因,因此使用适宜的外群研究它们相应的进化模式和速率是可行的。

核整合频繁发生并且在分类单元中广泛分布,从而允许我们研究内共生转移和基因组内互作的分子机

理。从如此众多的生物体中得到的有效数据使我们有理由认为 DNA 从线粒体迁移到核基因组并不仅仅是一起发生过的历史事件,而是一个不断进行的动态过程^[46]。目前,功能区域内共生迁移的结果在动物方面还未见报道,但在植物和真菌中已被观察到^[47]。尽管动物 mtDNA 基因组已经是非常紧凑,但是似乎没有真正的理由认为它们已经到达内共生进化的最终状态。

动物方面已报道的例子揭示迁移到核基因组序列具有不同的基因组结构模式:低拷贝序列明显的随机分布、散在分布^[42,48](如人^[49])和串联重复^[16],这可能意味着发生转移的时间不同^[50],具有不同的分子机制。研究动物线粒体外核拷贝序列仅仅是开始,并且大多数报道还未深入到核酸序列水平。全面理解整个迁移单元、末端区域、拷贝数、分布在染色体的位置和包括其它核元件(如重复元件^[51,52]),将阐明包括核外和核内基因组互作的分子机制和本质。

2.3 物种的系统发育关系 从某种程度上讲,Numts 提供了古 mtDNA 序列的历史记录,因而也为其应用于系统发育分析提供了可能。由于 Numts 的进化速率比 mtDNA 序列更慢,其核苷酸替换模式同古 mtDNA 状态^[42]更相似,因而看起来像分子“化石”^[10,21,24,42]。Numts 为现存 mtDNA 单倍体提供了理想的外群^[4,22,42],通过以 Numts 为外群的有根树为物种间的系统发育关系提供证据^[23]。Numts 已经被用于人群^[22]、雪鹅^[4]系统发育的外群。由于人类系统发育分析合适的外群非常有限,所以 Numts 在人方面非常有用。选择 Numts 作外群是因为与黑猩猩相比同人的分歧度更小,因为核突变率低,Numts 和变异的 mtDNA 序列的共同祖先相比,变异很少。因此,Numts 是快速进化的线粒体系统发育理想的树根^[22]。

当知道大量 Numts 的同源序列时,这些序列能被用于追踪 mtDNA 特殊位点的原始状态。正如在哺乳动物^[53]和昆虫^[54,55]中所做的一样,原始状态其它方面的知识对于解决线粒体系统发育关系方面存在的疑问有所帮助。

2.4 核和线粒体 DNA 突变模式 核和线粒体 DNA 在突变方式上不同,如核苷酸替代模式^[44]、缺少或插入的片段大小。mtDNA 一旦转座到核基因组后,mtDNA 相似序列就必须服从核基因组的突变模式及其偏向性。哺乳动物^[44]和昆虫^[56]的突变显示出线粒体比核基因组具有更高的转换/颠换值;哺乳动物核基因组 GC 位点更频繁甲基化,所以 GC → GT 的突变很常见,但 mtDNA 却相反^[57]。在线粒体基因组进化方面,链不对称性的影响也有报道,但在核基因组进化却没有这种

偏向性^[54]。

与之相比,那些没有被优先进行分子研究的物种^[7,54,58],通过 PCR 产物克隆测序已经得到大量的线粒体 DNA 位于核基因组的假基因序列数据^[40]。使用限制性内切酶获得目标假基因进化的分歧时间成为可能。多细胞动物的 Numts 是失去功能的假基因^[59],所以核突变能同线粒体的变化区别开,同时通过研究 Numts 还可以进行线粒体与核序列间和核与核序列间替换模式和替换速率的多重比较^[60],结果使研究核酸的替换、插入和缺失成为可能^[55]。Numts 序列的同源性允许检测 Numts 所在位置序列对突变的影响^[55],通过已定位的 Numts 可用来检测染色体位置对突变的影响^[46]。

综上所述,尽管 Numts 给 mtDNA 分子标记的分子系统学研究带来了不少的麻烦,使以 PCR 为基础的分析方法面临挑战,但随着检查和避免 Numts 的方法、mtDNA 和 Numts 的进化模式等方面研究的不断深入,越来越多的研究表明,核基因组线粒体相似序列在以 mtDNA 作标记的进化研究方面可作为有价值的分子工具,研究者往往有目的地去发现 Numts,有助于更进一步理解进化的本质。

参 考 文 献

- [1] Moore W S. Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. *Evolution*, 1995, **49**: 718 ~ 726.
- [2] Zhang D X, Hewitt G M. Nuclear integrations: Challenges for mitochondrial DNA markers. *Trends in Ecology and Evolution*, 1996a, **11**: 247 ~ 251.
- [3] Mindell D P. Avian molecular evolution and systematics. New York: Academic Press, 1997.
- [4] Quinn T W. The genetic legacy of Mother Goose: phylogeographic patterns of Lesser Snow Goose *Chen caerulescens* maternal lineages. *Molecular Ecology*, 1992, **1**: 105 ~ 117.
- [5] Avise J C, Arnold J, Ball R M, et al. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annu Rev Ecol Syst*, 1987, **18**: 489 ~ 522.
- [6] Simon C. Molecular systematics at species boundary: Exploiting conserved and variable regions of the mitochondrial genome of animals via direct sequencing from amplified DNA. In: Hewitt G M, Johnston A W B, Young J P W, eds. *Molecular Techniques in Taxonomy*. NATO Advanced Studies Institute. Springer-Verlag, 1991.
- [7] Bensasson D, Zhang D-X, Hewitt G M. Frequent assimilation of mitochondrial DNA by grasshopper nuclear genomes. *Mol Biol Evol*, 2000 **17**(3) 406 ~ 415.
- [8] Cummins J. Mitochondrial DNA and the Y chromosome: parallels and paradoxes. *Reprod Fertil Dev*, 2001; **13**(7 ~ 8): 533 ~ 542.
- [9] Hirano M, Shuilbans A, Mayeux R, et al. Apparent mtDNA heteroplasmy in Alzheimers disease patients and normals due to PCR amplification of nucleus-embedded mtDNA pseudogenes. *Proc Natl Acad Sci USA*, 1997, **94**: 14 894 ~ 14 899.
- [10] Zischler H, Hoss M, Handt O, et al. Detecting dinosaur DNA. *Science*, 1995a, **268**: 1 192 ~ 1 193.
- [11] van der Kuly A C, Kuiken C L, Dekker J T, et al. Nuclear counterparts of the cytoplasmic mitochondrial 12S rRNA gene: a problem of ancient DNA and molecular phylogenies. *J Mol Evol*, 1995, **40**: 652 ~ 657.
- [12] Mundy N I, Winchell C S, Woodruff D S. Tandem repeats and heteroplasmy in the mitochondrial DNA control region of the Loggerhead Shrike (*Lanius ludovicianus*). *Journal of Heredity*, 1996, **87**: 21 ~ 26.
- [13] Moritz C, Brown W M. Tandem duplication of D-loop and ribosomal RNA sequences in lizard mitochondrial DNA. *Science*, 1986, **233**: 1 425 ~ 1 427.
- [14] Irwin D M. Ancient duplications of the human proglucagon gene. *Genomics*, 2002, **79**(5): 741 ~ 746.
- [15] Sorenson M D, Quinn T W. Numts: a challenge for avian systematics and population biology. *The Auk*, 1998, **115**: 214 ~ 221.
- [16] Lopez J V, Yuhki N, Masuda R, et al. Numt, a recent transfer and tandem amplification of mitochondrial DNA to nuclear genome of the domestic cat. *Journal of Molecular Evolution*, 1994, **39**: 174 ~ 190.
- [17] Smith M F, Thomas W K, Patton J L. Mitochondrial DNA-like sequences in the nuclear genome of an akodontine rodent. *Molecular Biology and Evolution*, 1992, **9**: 204 ~ 215.
- [18] Collura R V, Auerbach M R, Stewart C B. A quick direct method that can differentiate expressed mitochondrial genes from their nuclear pseudogenes. *Curr Biol*, 1996, **6**: 1 337 ~ 1 339.
- [19] Blanchard J L, Schmidt G W. Mitochondrial DNA migration events in yeast and humans: integration by a common end-joining mechanism and alternative perspectives on nucleotide substitution pattern. *J Mol Evol*, 1996, **13**: 537 ~ 548.
- [20] Lemos B, Canavez F, Moreira M A M, et al. Mitochondrial DNA-like sequences in the nuclear genome of the opossum genus *Didelphis* (Marsupialia: Didelphidae). *J Heredity*, 1999, **90**: 543 ~ 547.
- [21] Arctander P. Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. *Proceedings of the Royal*

- Society of London Series B*, 1995, **262**:13 ~ 19.
- [22] Zischler H, Geisert H, Haeseler A, *et al.* A nuclear "fossil" of the mitochondrial D-loop and the origin of modern humans. *Nature*, 1995b, **378**:489 ~ 492.
- [23] Sorenson M D, Fleischer R C. Multiple independent transpositions of mitochondrial DNA control region sequences to the nucleus. *Proceedings of the National Academy of Sciences USA*, 1996, **93**:15 239 ~ 15 243.
- [24] Collura R V, Stewart C B. Insertions and duplications of mtDNA in the nuclear genomes of Old World monkeys and hominoids. *Nature*, 1995, **378**:485 ~ 489.
- [25] Kwok S, Chang S Y, Sminsky J J, *et al.* A guide to the design and use of mismatched and degenerate primers. *PCR Methods and Applications*, 1994, **3**:S39 ~ S47.
- [26] Dowling T E, Moritz C, Palmer J D, *et al.* Analysis of fragments and restriction sites. In: Hillis D M, Moritz C, Mable B K. eds. *Molecular Systematics*, 2nd ed. Sinauer Associates, Sunderland, Massachusetts, 1996.
- [27] Cheng S, Higuchi R, Stoneking M. Complete mitochondrial genome amplification. *Nature Genetics*, 1994, **7**:350 ~ 351.
- [28] Ruokonen M K, Lumme J. Close relatedness between mitochondrial DNA from seven *Anser* goose species. *Journal of Evolutionary Biology*, 2000, **13**:532 ~ 540.
- [29] Paabo S, Irwin D M, Wilson A C. DNA damage promotes jumping between templates during enzymatic amplification. *Journal of Biological Chemistry*, 1990, **265**:4 718 ~ 4 721.
- [30] Kwok S, Kellogg D E, Mckinney N, *et al.* Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus type 1 model studies. *Nucleic Acids Research*, 1990, **18**:999 ~ 1005.
- [31] Zhang D X, Hewitt G M. Highly conserved nuclear copies of the mitochondrial control region in the desert locust *Schistocerca gregaria*: some implications for population studies. *Mol Ecol*, 1996b, **5**: 295 ~ 300.
- [32] Sunnucks P, Wilson A C C, Beheregaray L B, *et al.* SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Mol Ecol*, 2000, **9**:1 699 ~ 1 710.
- [33] Li-Sucholeiki X C, Khrapko K, Andre P C, *et al.* Applications of constant denaturant capillary electrophoresis/high fidelity polymerase chain reaction to human genetic analysis. *Electrophoresis*, 1999, **20**:1 224 ~ 1 232.
- [34] Zischler H. Nuclear integrations of mitochondrial DNA in primates: inference of associated mutational events. *Electrophoresis*, 2000, **21**:531 ~ 536.
- [35] Zischler H, Geisert H, Castresana J. A hominoid-specific nuclear insertion of the mitochondrial D-loop: implication for reconstructing ancestral mitochondrial sequences. *Mol Biol Evol*, 1998, **15**: 463 ~ 469.
- [36] Thomas R, Zischler H, Pssbo S, *et al.* Novel mitochondrial DNA insertions polymorphism and its usefulness for human population studies. *Hum Biol*, 1996, **68**:847 ~ 854.
- [37] Ullrich H, Lattig K, Brennicke A, *et al.* Mitochondrial DNA variations and nuclear RFLPs reflect different genetic similarities among 23 *Arabidopsis thaliana* ecotypes. *Plant Mol Biol*, 1997, **33**:37 ~ 45.
- [38] Ayliffe M A, Scott N S, Timmis J N. Analysis of plastid DNA-like sequences within the nuclear genomes of higher plants. *Mol Biol Evol*, 1998, **15**:738 ~ 745.
- [39] Yuan J D, Shi J X, Meng G X, *et al.* Nuclear pseudogenes of mitochondrial DNA as a variable part of the human genome. *Cell Res*, 1999, **9**:281 ~ 290.
- [40] Bensasson D, Zhang D X, Hartl D L, *et al.* Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology & Evolution*, 2001, **16** (6):314 ~ 321.
- [41] Corinne D, Hajer H, Catherine O C, *et al.* Remarkable compartmentalization of transposable elements and pseudogenes in the heterochromatin of the *Tetraodon nigroviridis* genome. *Proc Natl Acad Sci USA*, 2002, **99**: 13 636 ~ 13 641.
- [42] Fukuda M, *et al.* Mitochondrial DNA-like sequences in the human nuclear genome. *J Mol Biol*, 1985, **186**:257 ~ 266.
- [43] Vartanian J P, Simon W H. Analysis of a library of macaque nuclear mitochondrial sequences confirms macaque origin of divergent sequences from old oral polio vaccine samples. *Proc Natl Acad Sci USA*, 2002, **99**:7 566 ~ 7 569.
- [44] Brown W M, Prager E M, Wang A, *et al.* Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J Mol Evol*, 1982, **18**:225 ~ 239.
- [45] Tamura K. The rate and pattern of nucleotide substitution in *Drosophila* mitochondrial DNA. *Mol Biol Evol*, 1992, **9**: 814 ~ 825.
- [46] Tourmen Y, Baris O, Dessen P, *et al.* Structure and chromosomal distribution of human mitochondrial pseudogenes. *Genomics*, 2002, **80**(1):71 ~ 77.
- [47] Gillham N W. *Organelle Genes and Genomes*. Oxford University Press, 1994.
- [48] Kamimura N, Ishii S, Ma L D, *et al.* Three separate mitochondrial DNA sequences are contiguous in human genomic DNA. *J Mol Biol*, 1989, **210**:703 ~ 707.
- [49] Eichler E E, Johnson M E, Alkan C, *et al.* Divergent origins and concerted expansion of two segmental duplications on chromosome 16. *J Hered*, 2001, **92**(6):462 ~ 468.
- [50] Raes J, Vandepoele K, Simillion C, *et al.* Investigating ancient duplication events in the *Arabidopsis* genome. *J Struct*

- Funct Genomics*, 2003, **3**(1-4):117 ~ 129.
- [51] Gellissen G, Bradfield J Y, White B N, *et al.* Mitochondrial DNA sequences in the nuclear genome of a locust. *Nature*, 1983, **301**:631 ~ 634.
- [52] Tsuzuki T, Nomiyama H, Setoyama C. *et al.* Presence of mitochondrial-DNA-like sequences in the human nuclear DNA. *Gene*, 1983, **25**:223 ~ 229.
- [53] Hu G, Thilly W G. Evolutionary trail of the mitochondrial genome as based on human 16S rDNA pseudogenes. *Gene*, 1994, **147**:197 ~ 204.
- [54] Sunnucks P, Hales D F. Numerous transposed sequences of mitochondrial cytochrome oxidase I - II in aphids of the genus *Sitobin* (Hemiptera: Aphididae). *Mol Biol Evol*, 1996, **13**:510 ~ 524.
- [55] Bessason D, Dmitri A P, Zhang D X, *et al.* Genomic gigantism: DNA loss is slow in mountain grasshoppers. *Mol Biol Evol*, 2001, **18**:246 ~ 253.
- [56] Petrov D A, Hartl D L. Patterns of nucleotide substitution in Drosophila and mammalian genomes. *Proc Natl Acad Sci USA*, 1999, **96**:1 475 ~ 1 479.
- [57] Bulmer M. Neighbouring base effects on substitution rates in pseudogenes. *Mol Biol Evol*, 1986, **3**:322 ~ 329.
- [58] Williams S T, Knowlton N. Mitochondrial pseudogene are pervasive and often insidious in the snapping shrimp genus *Alpheus*. *Mol Biol Evol*, 2001, **18**:1 484 ~ 1 493.
- [59] Graur D, Shuali Y, Li W H. Deletions in processed pseudogenes accumulate faster in rodents than in humans. *J Mol Evol*, 1989, **28**:279 ~ 285.
- [60] Mundy N I, Pissinatti A, Woodruff D S. Multiple Nuclear Insertions of Mitochondrial Cytochrome *b* Sequences in Callitrichine Primates. *Mol Biol Evol*, 2000, **17**(7):1 075 ~ 1 080.