

# 比较基因组学在哺乳动物进化研究中的应用

李培青 刘焕民 朱必才\*

(徐州师范大学生命科学学院, 徐州 221116)

**摘要** 哺乳动物经过长期进化, 使其基因组在结构和功能上存在着明显的差异, 构成了表型进化的基础。随着人类、部分哺乳动物基因组测序的完成, 以比较基因组学为主要研究手段的哺乳动物进化研究应运而生, 从而为在基因组水平上深入认识哺乳动物进化关系、揭示生命的起源和进化提供依据。对比较基因组学的主要研究方法进行了综述, 进而探讨其在哺乳动物进化研究中的应用, 并对哺乳动物比较基因组学的发展进行了展望。

**关键词** 比较基因组学; 哺乳动物; 进化

人类基因组全序列测序工作的完成具有划时代的意义, 标志着基因组研究的新纪元已经到来。随着被成功测序的哺乳动物种类数目的增加, 相关物种间基因组的比较分析自然就成了下一步的工作重点, 随之诞生了一个新的生命科学研究方向——比较基因组学。比较基因组学(comparative genomics)是在具有生物基因组图谱和成功测序前提下, 比较物种基因组间的相似性和差异性, 进而阐述其内在分子机制, 以了解基因的功能、表达机制和物种进化的一门新兴学科<sup>[1,2]</sup>。比较基因组学的重要性在于它把不同的学科、不同的生物种类联系在一起, 架起了基础研究和应用研究之间的桥梁。跨种、跨属的基因组比较对人们认识哺乳动物间的分化、亲缘关系等表型进化过程, 以及 DNA 变化如何导致生物多样性有十分重要的意义。

在生物的进化过程中, 基因组是一个动态的结构, 不断地被修饰、演化。以基因组为研究对象的比较基因组学是常规的以基因为研究对象的分子生物学的一次飞跃。对伴随人类基因组而完成的大量生物完整基因组的信息分析, 不仅将直接帮助破译人类遗传密码, 其本身也可能解决关键问题<sup>[3]</sup>。近年来, 比较基因组学的兴起对人们全面理解哺乳动物的进化起着积极作用。在不同物种、不同进化水平的生物的相关基因之间进行比较分析, 是基因研究的重要手段。目前, 人们有了越来越多的哺乳动物全基因组序列数据, 因此, 对哺乳动物进化研究, 也必须从基因水平上的比较, 上升到对不同进化水平的生物在全基因组水平上的比较研究。这样的研究将更有效地阐述哺乳动物进化的历程, 揭示

基因在生命系统中的地位和作用, 解释整个生命系统的组成和作用方式<sup>[4]</sup>。由完整基因组研究所导致的比较基因组学必将为哺乳动物进化研究开辟新的领域。

## 1 比较基因组学的研究方法

### 1.1 荧光原位杂交技术

荧光原位杂交技术(fluorescence *in situ* hybridization, FISH)技术是20世纪80年代在已有的放射性原位杂交技术的基础上发展起来的一种分子细胞遗传学技术, 正在引起分子细胞遗传学的深刻变化, 它已经深入的改变了基因组研究的各个方面<sup>[5]</sup>。通过特异的探针与裸露 DNA 纤维的杂交, 以显示其杂交信号, 可以检测仅有数千碱基的缺失。FISH 技术的应用会加深加快在哺乳动物的染色体进化研究及亲缘关系等领域的认识<sup>[6]</sup>。

比较染色体涂染技术(comparative chromosome painting)是对荧光原位杂交技术的发展, 已被广泛应用<sup>[7]</sup>。作为荧光原位杂交的一种特殊应用, 染色体涂染技术是用某个染色体的 DNA 片段混合物, 通过荧光标记检测中期染色体的最有效的方法<sup>[8]</sup>。亦可用于比较哺乳动物相近物种中染色体的同源性<sup>[9]</sup>。该技术所揭示的同源片段是建立在 DNA 序列同源的基础上, 可以从整体水平比较基因组中所有的染色体片段, 因而比传统显带和基因定位技术具有更准确、可靠、快速等优点, 正日益成为目前比较基因

收稿日期: 2005-05-09 接受日期: 2005-07-20

国家自然科学基金资助项目(No.30470966, No.30270738)

\* 通讯作者。Tel: 0516-83500082, E-mail: bicaizhu0054@xznz.edu.cn

组和染色体进化研究的主要手段<sup>[10]</sup>。

## 1.2 比较基因组杂交技术

比较基因组杂交技术(comparative genomic hybridization, CGH)是在1992年创建的<sup>[11]</sup>,原理是在荧光原位杂交基础上发展起来的一种新的分子细胞遗传学技术,用不同的荧光染料标记基因组DNA和待测细胞的DNA,再与中期染色体杂交,通过检测染色体上的荧光信号,确定待测组织DNA拷贝数的改变,同时在染色体上定位。它使间期基因组快捷可靠的检测成为可能。和传统细胞遗传分析技术不同,它无需体外培养细胞。与FISH技术相比,它克服了需制备染色体特异区域的探针且仅能检测个别染色体或部分位点的局限性,它是一种全方位检测法,一次实验即可检测出待测标本整个基因组DNA拷贝数增减,亦可在物种间进行染色体同源性比较。染色体物质的丢失或增多也可以通过比较基因组杂交方法而核实<sup>[12]</sup>。最近,比较基因组杂交的广泛用途在研究灵长类的进化上已经得到有力的证明<sup>[13]</sup>。

CGH技术也有局限性,如中期染色体的应用限制了基因组小片段(20 mb)和紧密连锁检测, Pollack等<sup>[14]</sup>用cDNA微阵列代替中期染色体,具有了较高的分辨率,能比较DNA拷贝数与基因表达的差异,从而克服了此局限性。

## 1.3 微阵列-比较基因组杂交技术

微阵列-比较基因组杂交(microarray-CGH, 又称array CGH)技术是将基因芯片技术和传统比较基因组杂交技术集为一体的新技术<sup>[14,15]</sup>。它通过在一微小的基片表面固定大量的基因探针,用标记后的待检测样本与已固定的核苷酸序列杂交,然后检测信号,确定样本中该基因或核苷酸序列的含量。由于该技术具有较高的灵敏度和精确性,而且又借助了计算机自动操作系统,实现了程序化,快速又直观,所以在哺乳动物进化研究中具有广阔应用前景。

随着基因组研究的深入和对生物学问题的探索,array CGH技术很快被应用于很多研究领域。如在基因诊断和基因治疗方面,array CGH技术更易于发现新的癌变基因。由于对基因的定位更加精确,一旦发现某一序列上存在癌变基因,通过基因组数据库就很容易找到此基因<sup>[16]</sup>。此外,array CGH技术还能高分辨率、大规模地检测染色体的异常,进行同源染色体的比较<sup>[17]</sup>,以鉴定哺乳动物间亲缘和进化关系。

## 2 比较基因组学的应用

### 2.1 鉴定哺乳动物间的亲缘和进化关系

由于近年来基因组学的迅猛发展和基因组数据的大量积累,衍生出一个新的研究方向——进化基因组学<sup>[18]</sup>。它在基因组水平阐释生物学功能和生命现象,或依据基因组数据研究基因组自身的进化过程和规律。它让人们能在更大程度上探究功能基础上的进化限制因素<sup>[19]</sup>。哺乳动物的复杂进化过程同时也是自然选择的过程和大量基因差异表达的结果。基因的差异表达不仅表现在转录水平上,也表现在翻译水平上,最终则表现为特异的生物功能的差异。因而在不同的水平上,对基因在不同组织器官、不同发育阶段和不同环境条件下表达模式的研究可为基因的生物学功能提供证据。在鉴定哺乳动物间的亲缘和进化关系时,以往的鉴定参数如:哺乳动物间基因组的大小、(G+C)的含量等,但是可靠性不强,灵敏度较低<sup>[11]</sup>。比较基因组学的出现使得快速的高效定性分析成为可能。目前,哺乳动物的进化研究,已经在3个水平即DNA水平、基因水平和基因组水平上展开,主要是在基因组水平的研究。

**2.1.1 DNA水平上的比较** 常见的哺乳动物基因进化模式为染色体的断裂破坏和染色体片段的交换。这些现象可以通过DNA水平上基因组结构的比较而验证<sup>[11,20]</sup>。Chureau等<sup>[21]</sup>分析比较了人和小白鼠X染色体上L1元件(一种重复的DNA序列)的分布,发现了DNA的两条链上的L1元件分布不等,并猜测L1元件可能具有某种潜在功能。Mural等<sup>[22]</sup>证实了DNA断裂位点附近具有显著的不连续性,他们推测了DNA断裂位点的功能有利于理解哺乳动物基因组的进化。这样使经典的遗传学和现代基因组学、表型与序列相结合,从而可对哺乳动物进化有更深刻的理解。

**2.1.2 基因水平上的比较** 染色体的破坏和染色体片段的交换导致了基因序列的中断。基因序列的研究可通过相近哺乳动物基因间的比较而实现,包括分析基因序列的保守性、定向进化同源基因的产生位点等。人类和大量哺乳动物基因组序列区域已经被应用于比较研究,研究的结果表明编码序列是很保守的<sup>[23,24]</sup>。对保守区域的详细分析和比较可以帮助人们理解基因组中基因和调控元件的组织结构。最近研究表明,哺乳动物基因组里普遍存在的非编码DNA在进化中有重要作用,很可能成为研究

基因进化的转折点<sup>[25]</sup>。此外, 在哺乳动物基因组中存在着大量的重叠基因。人类和小白鼠的基因组中鉴定了许多的重叠基因, 然而, 这些重叠基因是如何起源和进化的仍是一个未知数<sup>[26]</sup>, 还需要实验证据的进一步论证。

**2.1.3 基因组水平上的比较** 哺乳动物的基因组中, 需要大量的序列分析才能鉴定极少量的基因。为了收集这些序列中的相关信息, 比较分析相近哺乳动物间完整基因组序列, 不仅能帮助快速鉴定相关基因和它们的调控序列, 而且还能洞察它们基因组的结构和进化上的关系<sup>[27,28]</sup>。作为模式生物的小白鼠详尽的基因组物理图已被构建<sup>[29]</sup>, 通过对整个基因组序列的比较分析, 既有助于快速鉴定基因, 也为研究基因组间的结构和进化关系开拓新的视野。黑猩猩的基因组草图已经被成功构建, 关于黑猩猩的大量 cDNA 序列被应用于比较研究<sup>[30]</sup>, 这可以用来鉴定自然选择下的人的特殊基因的变化, 还有利于理解两个最相近的哺乳动物基因组的分子进化的动力学。已经报道了关于人类和黑猩猩的染色体重排在它们分子进化上有重要影响<sup>[31]</sup>。最近研究表明, 在两者的进化过程中重排染色体上的蛋白质进化的速度显著快于共线性染色体(colinear chromosomes)上的蛋白质进化的速度<sup>[32]</sup>。通过比较人类和黑猩猩二者的染色体发现: 他们的许多染色体发生了重排, 其余的则是呈共线性, 可以认为染色体重排在人类和黑猩猩的物种形成中起重要作用。

## 2.2 比较基因组学与比较遗传学的结合

目前, 比较基因组学的一个明显发展趋势是和比较遗传学相互结合, 相互渗透。随着遗传作图的不断深入, 大量已知功能基因被定位在染色体上, 比较遗传学在比较基因排列顺序的基础上, 开始分析已知基因的结构和功能。这样的比较能更深入的揭示染色体上基因数目和排列顺序变化所产生的效应, 从而真正解决基因组起源、进化以及个体发育中的一系列重大问题。O'Brien 等<sup>[4]</sup>比较了人和一些模式生物的基因组结构和功能, 同时比较了包括人类基因组在内的 11 种哺乳动物的整个基因组水平上的基因排列顺序, 他们的工作将比较遗传学和比较基因组学研究结合起来。随着对基因功能的深入研究, 人们意识到基因的功能往往受到其他基因的影响, 而基因在染色体上的位置及其排列顺序和基因的功能并非全无关系<sup>[9]</sup>。例如, 已经将遗传作图和基因组的比较分析结合起来的方法, 在阐明人类和

小白鼠的染色体上组蛋白的修饰方面有重大进展<sup>[33]</sup>。最近, 应用 FISH 技术和遗传作图的结合在研究灵长类动物进化中染色体的可塑性也取得成功<sup>[13]</sup>。

## 2.3 比较基因组学与 DNA 芯片技术的结合

应用高密度芯片对模式生物基因组进行大规模的测序, 进而对进化过程中各个阶段不同生物基因序列的保守性及差异性进行比较研究, 将有助于揭示基因功能和疾病分子机制, 阐明物种的进化关系<sup>[34]</sup>。Hacia 等<sup>[35]</sup>利用芯片技术对与人进化相近的黑猩猩和大猩猩等物种的同源基因 BRCA1 的外显子 11 的序列进行比较分析, 证明在 314 kb 中有 83.5% 至 98.2% 的核苷酸相同, 与杂交法和双脱氧测序分析所得的结果一致, 这说明芯片技术可大大加快比较基因组学研究的进程。

## 3 展望

人们已经开始认识到比较并分析基因组的重要性, 并且正在制定有关大量模式生物基因组测序的计划, 已从模式生物之间以及与人类之间发现了一些共性特征以及各自的独特性<sup>[36]</sup>。目前正在兴起的分子进化生物学显示了研究哺乳动物进化重要性的趋势, 而比较基因组学则提供了人们研究哺乳动物进化的新的视角, 比较基因组学研究方法的改进大大加速了进化研究的进程。随着比较基因组学的发展, 人们定会获得大量与哺乳动物进化相关的序列信息, 从而对其进化复杂性会有更全面的理解。

## 参考文献 (References)

- [1] Wei L et al. *J Biomed Inform*, 2002, **35**: 142
- [2] Koonin EV et al. *Cell*, 2000, **101**: 573
- [3] Eichler EV et al. *Curr Opin Genet Dev*, 2003, **13**: 559
- [4] O'Brien SJ et al. *Science*, 1999, **286**: 458
- [5] van Ommen GJ et al. *Curr Opin Genet Dev*, 1995, **5**: 304
- [6] 朱必才等. *遗传*, 2003, **25**: 517
- [7] 朱必才等. *动物学报*, 2004, **50**: 245
- [8] Pinkel D et al. *Proc Natl Acad Sci USA*, 1988, **85**: 9138
- [9] Wienberg J et al. *Curr Opin Genet Dev*, 1995, **5**: 792
- [10] O'Brien SJ et al. *Trends Genet*, 1997, **13**: 393
- [11] Kallioniemi A et al. *Science*, 1992, **258**: 818
- [12] Struski S et al. *Cancer Genet Cytogenet*, 2002, **135**: 63
- [13] Tsend-Ayush E et al. *Genomics*, 2004, **83**: 193
- [14] Pollack JR et al. *Nat Genet*, 1999, **23**: 41
- [15] Solinas-Toldo S et al. *Genes Chromosomes Cancer*, 1997, **20**: 399
- [16] Giglio S et al. *Am J Hum Genet*, 2001, **68**: 874
- [17] Veltman JA et al. *Am J Hum Genet*, 2003, **72**: 1578
- [18] Eyre-Walker A. *Trends Ecol Evol*, 1999, **14**: 176
- [19] Jones SJ. *Curr Opin Genet Dev*, 1995, **5**: 349
- [20] Naruya S. *Trends Genet*, 2002, **18**: 239

- [21] Chureau C *et al. Genome Res*, 2002, **12**: 894  
[22] Mural RJ *et al. Science*, 2002, **296**: 1661  
[23] Hardison RC *et al. Genome Res*, 1997, **7**: 959  
[24] Jareborg N *et al. Genome Res*, 1999, **9**: 815  
[25] Gaffney DJ *et al. Trends Genet*, 2004, **20**: 332  
[26] Veeramachaneni V *et al. Genome Res*, 2004, **14**: 280  
[27] Muthusamy B *et al. Trends Genet*, 2002, **18**: 553  
[28] Sorek R *et al. Trends Genet*, 2004, **20**: 68  
[29] Gregory SG *et al. Nature*, 2002, **418**: 743  
[30] Fujiyama A *et al. Science*, 2002, **295**: 131  
[31] Navarro A *et al. Science*, 2003, **300**: 321  
[32] Vallender EJ *et al. Genomics*, 2004, **84**: 757  
[33] Bernstein BE *et al. Cell*, 2005, **120**: 169  
[34] Müller UR *et al. Clin Appl Immunol Rev*, 2001, **1**: 255  
[35] Hacia JG *et al. Nat Genet*, 1998, **18**: 155  
[36] Clark AG *et al. Curr Opin Genet Dev*, 2002, **12**: 631

## Application of Comparative Genomics in Mammalian Evolution Research

Pei-Qing Li, Huan-Min Liu, Bi-Cai Zhu\*

(School of Life Science, Xuzhou Normal University, Xuzhou 221116, China)

**Abstract** Mammalian genome has obvious diversities in structure and function after long time of evolution, which also forms the foundation of evolution. As the number of completed mammalian genome sequences increases, there is an increasing emphasis on comparative genomic analysis of closely related organisms in mammalian evolution. Mammalian evolution will be intensively cognized in genome level by comparative genomics, in order to discover the origin of life and also provide the proof for evolution. This article summarizes the methods of comparative genomics, then discusses its application in mammalian evolution, and also expresses the prospect on mammalian comparative genomics.

**Key words** comparative genomics; mammalian; evolution

Received: May 9, 2005 Accepted: July 20, 2005

This work was supported by the National Natural Science Foundation of China (No.30470966, No.30270738)

\*Corresponding author. Tel: 86-516-83500082, E-mail: bicaizhu0054@xznu.edu.cn