

STUDY ON MOLECULAR  
**B**IOLOGY OF GENETICS,  
BREEDING AND CLONING  
IN ANIMALS

李 宁 朱作言 等 著

动物遗传育种与  
克隆的分子生物学  
基础研究



科学出版社  
[www.sciencep.com](http://www.sciencep.com)

# 动物遗传育种与克隆的 分子生物学基础研究

李 宁 朱作言 等著

科学出版社

北 京

## 内 容 简 介

本书介绍了“973”计划项目“农业动物遗传育种与克隆的分子生物学基础研究”的研究成果,内容包括:肉、奶产量及其品质性状主基因的分离、克隆、测序、定位分析和性状基因的表达调控,农业动物杂种优势的分子机理和分子遗传标记辅助选择改良机理的研究进展,以及体细胞克隆胚胎核质互作中细胞核活动规律、农业动物克隆繁殖的分子生物学机理的研究成果。这些重大科学问题的阐明,将为我国在21世纪的动物农业提供高产、优质、高效发展的遗传理论基础和高新技术。

本书可作为农业院校动物遗传育种专业和综合性大学动物科学相关专业研究生的参考用书,也可供以动物为研究对象的科研院所的研究人员参考阅读。

### 图书在版编目(CIP)数据

动物遗传育种与克隆的分子生物学基础研究/李宁,朱作言等著. —北京:科学出版社,2009

ISBN 978-7-03-023246-5

I. 动… II. ①李…②朱… III. ①畜禽育种-遗传育种-分子生物学-研究②畜禽-无性系-遗传工程-分子生物学-研究 IV. S813.2

中国版本图书馆CIP数据核字(2008)第164828号

责任编辑:李秀伟 刘 晶 王 静/责任校对:包志虹

责任印制:钱玉芬/封面设计:美光制版

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

北京佳信达欣艺术印刷有限公司印刷

科学出版社发行 各地新华书店经销

\*

2009年6月第 一 版 开本:787×1092 1/16

2009年6月第一次印刷 印张:23 3/4 插页12

印数:1—1 500 字数:548 000

定价:86.00元

如有印装质量问题,我社负责调换

《动物遗传育种与克隆的分子生物学基础研究》  
编写委员会

- |     |                  |
|-----|------------------|
| 张亚平 | 中国科学院昆明动物研究所     |
| 李 宁 | 中国农业大学           |
| 赵兴波 | 中国农业大学           |
| 张细权 | 华南农业大学           |
| 黄路生 | 江西农业大学           |
| 张 勤 | 中国农业大学           |
| 李 奎 | 华中农业大学           |
| 吴常信 | 中国农业大学           |
| 陈瑶生 | 华南农业大学           |
| 徐宁迎 | 浙江大学             |
| 张 沅 | 中国农业大学           |
| 熊远著 | 华中农业大学           |
| 郑友民 | 中国农业科学院北京畜牧兽医研究所 |
| 魏 泓 | 第三军医大学实验动物中心     |
| 陈清轩 | 中国科学院遗传与发育生物学研究所 |
| 杜 森 | 中国科学院遗传与发育生物学研究所 |
| 季维智 | 中国科学院昆明动物研究所     |
| 谭景和 | 山东农业大学           |
| 周 琪 | 中国科学院动物研究所       |
| 陈大元 | 中国科学院动物研究所       |
| 汪亚平 | 中国科学院水生生物研究所     |
| 朱作言 | 中国科学院水生生物研究所     |

## 前 言

国家重点基础研究发展计划（“973”计划）项目“农业动物遗传育种与克隆的分子生物学基础研究”（项目编号：G2000016100）经过五年艰苦的努力工作，不仅发表了一批高水平的学术论文，多次获得国家级和省部级奖励，而且培养了一大批青年才干，积累了丰富的研究资源和阶段性成果，为我国动物遗传育种与克隆科学的持续原始创新奠定了基础。本项目研究人员团结协作，努力工作，执行情况优秀，在论文产出和人才培养等方面超额完成了预定的研究计划，同时也建立了优秀的研究队伍和研究平台，积累了重要的实验材料和研究思路，为进一步的深入研究创造了很好的基础和条件。项目组织了在国内从事动物遗传育种与克隆研究领域优秀的研究队伍进行综合研究，在项目实施的五年中，总体研究工作从原来的独立零散的研究小组发展为目前强大而且有创造力的群体，研究水平和规模有了很大的提高，走在国际先进的队伍行列，取得了许多重大或重要的研究成果，主要的创新性成果包括：①完成了我国主要农业动物的品种资源调查和 DNA 样品收集工作，初步探明了家养动物猪、鸡、狗、山羊的起源、驯化和进化路线；②利用我国特有的农业动物资源群，定位、分离和克隆了一批重要经济性状的功能基因，具有自主知识产权，也为我国分子育种奠定了基础；③动物克隆机理研究取得重大突破，不仅克隆了一批经济动物，也为动物克隆效率的提高奠定了理论基础。

经过 5 年的努力，项目取得了一批重要的研究成果，累计发表论文 634 篇，其中 SCI 收录的论文 245 篇，国际学术刊物上发表的论文 209 篇，EI 收录论文 1 篇；出版专著 2 部；获得发明专利 9 个，申请专利 26 项；获国家级科技奖 5 项和省部级奖励 8 项。这支研究队伍充满活力，不仅在理论和动物品种改良上取得了创新性的重要进展，超额完成了预定任务，而且更加明确了今后的研究目标以及突破的重点，为今后进一步的研究和持续原始创新打下了坚实的基础。这些研究成果在国内外同行中引起了广泛的关注，其中不乏 *Nature*、*Science*、*PNAS* 等国际顶尖学术期刊的文章和国家级的科技奖励和发明专利，相信本项目的研究成果汇编成书必将推动我国动物遗传育种与克隆的研究发展和深入。

作 者

2008 年 10 月 1 日

# 目 录

## 前言

第一章 家养动物的起源、进化与遗传多样性	1
第一节 狗的东亚起源：来自遗传学的证据	2
第二节 中国家养山羊的遗传多样性和系统发育地理结构研究	7
主要参考文献	18
第二章 农业动物重要经济性状的主基因定位	21
第一节 基因组扫描精细定位重要经济性状的主基因	21
第二节 利用候选基因法鉴定重要经济性状的主基因	37
第三节 重要经济性状基因的克隆、测序和表达分析	59
第四节 重要经济性状基因的 DNA 芯片研究	65
主要参考文献	78
第三章 农业动物遗传资源评估与优良基因的分离和利用	82
第一节 我国地方品种遗传资源的分子评估	82
第二节 我国地方品种优良基因的挖掘	88
第三节 优良基因在畜禽育种中的利用	101
第四节 QTL 定位方法研究	106
主要参考文献	118
第四章 农业动物“超高产”育种的分子生物学基础研究	120
第一节 重要经济性状超高产的分子生物学机理	120
第二节 重要经济性状主基因与多基因互作机理	122
第三节 分子遗传标记辅助育种机理	126
第四节 超高产优良基因组集成的分子育种	132
第五节 金华猪重要经济性状主效基因的研究	133
主要参考文献	158
第五章 农业动物杂种优势的分子遗传机理研究	161
第一节 农业生物杂种优势遗传机理研究进展	161
第二节 试验鸡群的构建	169
第三节 mRNA 差异显示技术平台的建立	172
第四节 鸡肝脏组织基因差异表达及与主要屠体性状杂种优势的关系	173
第五节 鸡心肌和肝脏组织基因差异表达及与屠体性状杂种优势关系	178
第六节 鸡卵巢组织基因差异表达及与产蛋性状杂种优势的关系	182
第七节 鸡 DNA 甲基化与杂种优势的关系	188
第八节 讨论与展望	190
主要参考文献	192

<b>第六章 近交动物资源评估、利用及其机理研究</b> .....	194
第一节 近交动物遗传资源评价.....	194
第二节 农业动物近交系培育.....	199
第三节 近交动物分子遗传学研究.....	200
主要参考文献.....	205
<b>第七章 细胞核在卵质中去分化和再程序化及发育全能性的研究</b> .....	207
第一节 体细胞核在卵质中去分化和再程序化及发育的分子机理研究.....	207
第二节 克隆胚胎中体细胞核活动规律与核全能性研究.....	235
主要参考文献.....	259
<b>第八章 动物体细胞克隆的细胞质与核的相互作用关系</b> .....	261
第一节 影响猕猴卵母细胞成熟的因素研究.....	261
第二节 动物克隆受核卵母细胞质成熟的调控.....	291
第三节 小鼠体细胞克隆的研究.....	306
第四节 猕猴胚胎干细胞的培养与诱导分化.....	325
主要参考文献.....	346
<b>第九章 种间细胞核质相互作用和核质杂交研究</b> .....	350
第一节 鱼类细胞核移植技术.....	350
第二节 鱼类细胞核移植技术的发展.....	351
第三节 鱼类细胞核移植研究.....	355
第四节 哺乳动物异种克隆中线粒体的命运.....	360
主要参考文献.....	368

图版

图版 1	.....	194
图版 2	.....	194
图版 3	.....	199
图版 4	.....	200
图版 5	.....	205
图版 6	.....	207
图版 7	.....	207
图版 8	.....	235
图版 9	.....	259
图版 10	.....	261
图版 11	.....	261
图版 12	.....	291
图版 13	.....	306
图版 14	.....	325
图版 15	.....	346
图版 16	.....	350
图版 17	.....	351
图版 18	.....	355
图版 19	.....	360
图版 20	.....	368

## 第一章 家养动物的起源、进化与遗传多样性

“973”计划项目“农业动物遗传育种与克隆的分子生物学基础研究”立项以来，课题组紧密围绕《国家重点基础研究发展规划课题任务书》规定的研究内容和指标，运用分子生物学技术解决农业动物系统与演化中重要的、长期悬而未决的疑难问题，系统深入地开展了猪、鸡、狗、山羊等农业动物的起源分化及基因多样性的研究工作，获得了一些新的理论认识。这些研究工作从遗传学角度探讨了一些农业动物起源中驯化事件发生的次数、地理位置和时间等十分富有挑战性的重要问题，研究了家养动物系统发育和群体历史中的精细事件与人类活动及迁移之间的密切关系。

家养动物向人类提供了稳定的动物蛋白来源，是人类社会由散居、渔猎型向群居、农业型转变的重要基础和必然产物。野生动物被驯化为家养动物后，其扩散是依赖于人类的迁移而实现的。因此，对家养动物遗传多样性的研究不仅是遗传育种的物质基础，而且对于了解人类社会的发展也有重要的价值。

与一般的野生动物相比，虽然家养动物的历史非常短，但在强烈的人工选择作用下，却呈现出极其丰富的形态变异，远远超过了其野生祖先。这为研究选择的作用及形态变异的发育遗传基础提供了绝佳的材料。此外，家养动物的遗传多样性是遗传育种的物质基础，也是人类社会的宝贵资源。

虽然达尔文在《物种起源》一书中专门论述了家养动物的起源问题，但在很长的时期内，人们对于家养动物的起源、进化与遗传多样性却知之甚少。近年来，基因组计划的发展及相关技术与信息的运用，推动了家养动物的研究。

中石器时代的考古学遗址和现代的遗传学证据都提示，狗可能是第一种被驯化的家养动物（Clutton-Brock, 1995），而其历史也不过大约 1.5 万年。张亚平研究组与瑞典皮特·赛沃雷尼博士研究组合作，通过研究狗的线粒体 DNA 多样性，揭开了狗的起源之谜，发现世界范围内的所有家犬具有共同的遗传基础，并有多个母系来源。他们的研究成果被选为封面论文，发表在 2002 年 11 月 22 日出版的 *Science* 上，引起了广泛注意（Savolainen et al., 2002）。该杂志同期还在“新闻聚焦（News Focus）”栏目发表了题为“A shaggy dog history”的评论（Pennisi, 2002）。张亚平研究组还系统地研究了我国山羊土著品种的遗传多样性及它们的起源、分化，提出了家养山羊有多个母系起源，且至少有一个亚支是在东亚起源的观点。对中国山羊的群体遗传结构的分析显示，在中国山羊群体中没有显著的系统发育地



——选自 Savolainen P, Zhang Y P, Luo J, Lundeberg J, Leitner T. 2002. Genetic evidence for an East Asian origin of domestic dogs. *Science*, 298: 1610~1613



理结构,提示在山羊的群体历史过程中发生了伴随人类迁移的广泛的运输。该研究结果已发表在国际知名 SCI 刊物 *Molecular Phylogenetics and Evolution* (《分子系统发育与进化》) 上 (Chen et al., 2005)。

## 第一节 狗的东亚起源:来自遗传学的证据

狗是人类最忠实的朋友,它在东西方文化中都扮演着重要的角色。但是,狗来自何方?它们有着怎样的遗传背景?瑞典皮特·赛沃雷尼博士领导的研究组与中国科学院昆明动物研究所张亚平研究组合作,深入研究了狗的线粒体 DNA 多样性,揭开了狗的起源之谜。

中石器时代的考古学遗址的证据提示,狗可能是第一种被驯化的家养动物 (Clutton-Brock, 1995)。遗传学、行为学和形态学的研究已经证明,狼是狗的祖先 (Clutton-Brock, 1995; Vilà et al., 1997; Wayne, 1993)。但是,狗是如何由狼驯化而来的过程却知之甚少,人们对于狗的驯化是在何时、何地以及如何发生的仍存在诸多疑问,长久以来有关这些问题的争论一直没有停息。

长期以来,科学家们只能借助考古学的方法来证明狗和狼的关系。最早的狗化石证据来自于德国 14 000 年前的一个下颌骨化石 (Nobis, 1979; Davis and Valla, 1978),另外一个来源于以色列大约 12 000 年前的小型犬科动物骨架化石 (Tchernov and Valla, 1997; Clutton-Brock, 1999)。这些考古学证据支持狗是起源于西南亚 (Olsen and Olsen, 1977) 或欧洲的。此外,狗的骨骼鉴定特征提示狗可能起源于中国的狼,由此提出了狗的东亚起源说 (Benecke, 1987; Lawrence, 1967)。然而,由于在形态上很难将小狼和狗区分,使得仅用考古学的方法研究狗的起源存在很大问题 (Benecke, 1987; Nobis, 1979)。

此外,不同品种狗在形态上极其丰富的多样性,似乎又倾向于狗起源于不同地理群体的狼的假说。由于狼的迁移性特别强,通过不同地理群体狼的研究难以提供狗起源的可靠线索。追溯狗起源过程中驯化事件发生的次数、地理位置和时间是十分富有挑战性的重要问题。最近,中国和瑞典科学家在这方面的研究上取得了突破性进展。

为确定家犬是在一个地区还是几个地区被驯化的,以及可能的驯化地和时间,我们和皮特·赛沃雷尼开始用另一种方法去探寻狗的起源之谜。我们选择母系遗传、进化速度快、缺乏重组的线粒体 DNA (mtDNA) 作为遗传标记,第一次系统研究了来自欧洲、亚洲、非洲和北美地区的几乎代表了世界范围内所有主要品种的 654 只狗和 38 只欧亚大陆狼的遗传变异。研究结果显示,虽然狗有多个母系起源,但是大于 95% 狗的单倍型属于三个主要支系——A、B 和 C,而且支系 A、B、C 在所有地理区域内出现的频率都十分相近 (图 1-1)。这样的结果提示,世界范围内不同地理区域的狗具有相同的遗传基础,即有一个共同的祖先。而东亚地区的核苷酸多样性显著高于其他区域,且系统地理结构的显著差异提示,来自全世界狗的线粒体 DNA 都可追溯到东亚,也就是说,狗起源于东亚,其驯化时间大约在 15 000 年前。

我们采集了来自欧洲、亚洲、非洲和北美地区的 654 个家犬和 38 个欧亚大陆狼的样品,这也是在世界上第一次收集到如此多的、能够覆盖世界范围的家犬样品 (Oku-

mura et al., 1996; Vilà et al., 1999)。我们对这些样品的 582bp 线粒体 DNA 遗传变异进行了分析, 所得到的数据可以较为细致地研究世界范围家犬的遗传变异。家犬在欧洲、西南亚和东亚的单倍型变异表明, 这三个地理区域很可能是家犬的起源地。先前的工作表明, 家犬至少来源于 4 个狼的母系群体 (Wayne, 1993; Kurtén and Anderson, 1980)。由我们的数据分析得到的系统发育关系显示, 共存在 6 个狗的单倍型类群 (支系 A~F)。除存在先前报道的 4 个系统发育类群 (支系 A~D) 外, 还存在第 5 个独立的单倍型类群 (支系 E), 它由一个单独的单倍型组成 (Vilà et al., 1999)。此外, 我们认为存在第 6 个支系 F, 理由是: 由系统发育树上看, 第 6 个类群虽然没有由支系 A 中单独地分化出来, 但由于与支系 A 之间存在较大的遗传距离, 因此足以确立其独立的支系地位。

我们对这 6 个类群进行了单独分析: 支系 A 中包含了 3 个来自于中国和蒙古的狼的单倍型; 支系 B 包含了 3 个狼的单倍型, 其中 2 个来自于东欧, 1 个来自于阿富汗。这似乎支持支系 A 起源于东亚而支系 B 起源于欧洲或者西南亚。然而, 狼是一种具有很强迁移能力的动物, 造成地区之间狼的线粒体 DNA 的区别相当模糊。Vilà 等发现, 1 只沙特阿拉伯狼的序列与 1 只保加利亚狼的序列完全相同, 而另 1 只沙特阿拉伯狼的序列又与蒙古狼的序列完全相同 (Fu, 1997)。因而, 部分狼的单倍型与家犬的单倍型有较近的亲缘关系就不足以用来证明与狼具有相同的单倍型的家犬的地理区域就是家犬的驯化地。在这种情况下, 世界范围内狗的遗传变异格局将是确定狗起源的一个关键指标。

总体上, 71.3% 的狗的单倍型属于支系 A, 95.9% 的狗的单倍型属于支系 A、B 和 C (表 1-1)。支系 A 遍布世界各地, 支系 B 和 C 分布于除美洲以外的其他所有地理区域, 显然, 支系 A、B、C 是世界不同地理区域内家犬的主流类群。有趣的是, 主流支系 A、B、C 在所有地理区域内的出现频率都十分相近 (表 1-1)。这表明, 除非在欧亚大陆存在非常强大的基因流, 否则当今的家犬群体不可能在任何一个地理位置上都能拥

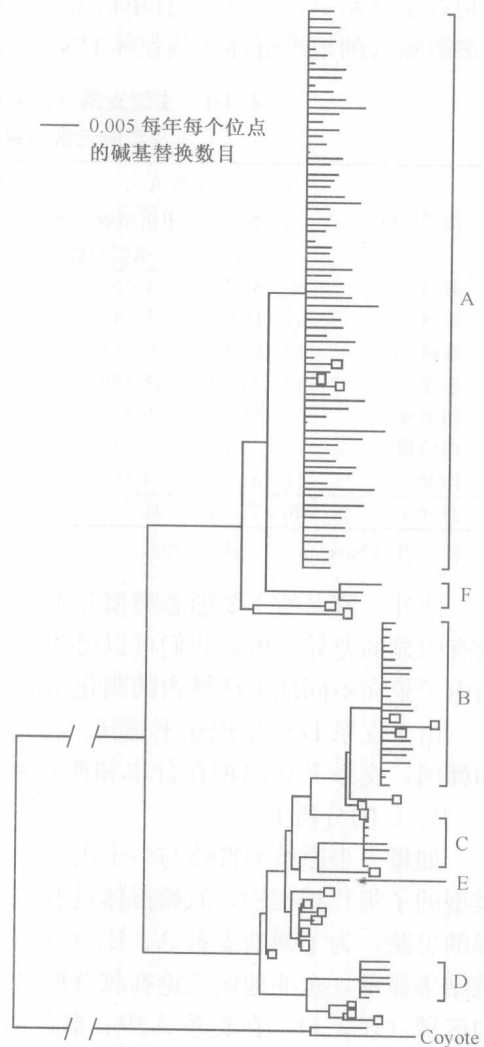


图 1-1 基于线粒体 DNA 单倍型类型构建世界家犬 (无标记) 和狼 (方块) 的系统发育树, 显示了家犬的 6 个单倍型类群 (A~F)。图中所示为以美洲山狗为外群, 在 HKY+I+ $\Gamma$  ( $\alpha=0.5960$ ,  $I=0.7367$ ) 核苷酸替代模型下构建的最大似然树。分枝长度如图所示 (Savolainen et al., 2002)

有包含了支系 A、B、C 所有单倍型类型在内的共同祖先的单一基因库。由此可见，绝大多数家犬的地理群体在线粒体 DNA 遗传变异上都具有共同的遗传基础。

表 1-1 主流支系 A、B 和 C 在不同地理区域的样本信息  
及其所包含的单倍型数和独有单倍型数

地理区域	支系 A		支系 B		支系 C	
	样本数 (%)	单倍型数 (独有 单倍型数)	样本数 (%)	单倍型数 (独有 单倍型数)	样本数 (%)	单倍型数 (独有 单倍型数)
非洲	30 (85.7)	14 (5)	4 (11.0)	1 (0)	1 (2.9)	1 (0)
美洲	25 (100)	5 (1)	0	0	0	0
欧洲	140 (67.6)	20 (9)	36 (18.0)	4 (1)	12 (5.9)	3 (0)
东亚	192 (73.8)	44 (30)	39 (15.0)	10 (7)	24 (9.2)	4 (1)
东南亚	51 (56.7)	16 (4)	32 (35.0)	4 (2)	5 (5.4)	2 (0)
西伯利亚	17 (70.8)	9 (1)	2 (8.3)	1 (0)	4 (17.0)	2 (1)
印度	11 (84.6)	4 (0)	1 (7.7)	1 (0)	1 (7.7)	1 (0)
总计	466 (71.3)	71	114 (17.4)	13	47 (7.2)	5

注：引自 Savolainen et al., 2002。

此外，家犬的主要形态类型介于大型犬和小型犬之间，在三个主要支系的构成上也并没有明显的差异。由此我们可以得到这样的结论：不同品种间形态上的巨大差异，并不是由于狼在不同地理区域内的驯化造成的。

由于支系 D 仅出现在土耳其、西班牙和斯堪的纳维亚地区，支系 E 仅出现在日本和韩国，支系 F 仅出现在日本和西伯利亚。因而，我们的分析主要集中于对主流支系 A、B、C 的分析上。

如果一个原始的群体与一个由其分化而来的群体（即来源于同一个原始类群的遗传类型的子集）相比较，原始群体具有较多的单倍型数量和较高的核苷酸多样性。依据这样的思路，为了判断支系 A、B、C 的起源地，我们比较了群体间的一些遗传参数，比较结果显示，东亚地区无论在核苷酸多样性上还是在独有单倍型的数量上都显著高于其他区域（表 1-1）。在支系 A 中，群体内两两序列间的平均遗传距离粗略的估计结果是：东亚的序列间的碱基替代数为  $3.39 \pm 0.13$ ，欧洲为  $2.97 \pm 0.08$ ，东南亚为  $2.28 \pm 0.23$ 。相应地，东亚比其他地区拥有更多的单倍型。在支系 A 中，总共含有 71 个单倍型。其中，东亚有 44 个单倍型，欧洲有 22 个单倍型，东南亚有 16 个单倍型。而且，东亚的 44 个单倍型中的 30 个（占 68%）为其独有。其他两个地区所独有的单倍型总和还不到 50%（表 1-1）。此外，东亚具有一些变异较大的独有单倍型，而其他区域独有单倍型的变异较小。

图 1-2A，利用最小跨度网络图法分别显示了家犬线粒体 DNA 控制区序列三个主要单倍型类群支系 A、B、C 不同地理区域单倍型的分布和频率：方块代表狼的单倍型，圆圈代表家犬的单倍型；圆圈的颜色表示地理区域（橘黄色代表东亚，蓝色代表欧洲，绿色代表东南亚，空心圆圈代表未在该区域发现）；圆圈大小表示单倍型的频率。从图上可以看出：东亚分布有比其他区域丰富的单倍型类型，其中含有大量的独有单倍型。这表明东亚具有高于其他区域的遗传多样性，结合其他信息提示了家犬的东亚起源。图 1-2B 是以喜马拉雅山脉和乌拉尔山脉为界，东部和西部地区独有单倍型和共享单倍型之间的最小跨度网络关系图。黑点表示假定的中间节点；白色圆圈代表东西部共享的单

倍型；蓝色圆圈代表西部独有的单倍型；深蓝色代表西部独有的单倍型，与东部的单倍型间存在两步的突变；橙色圆圈代表东部独有的单倍型；红色圆圈代表东部独有的单倍型，与西部的单倍型间存在两步的突变；红色粗圆圈代表东部独有的单倍型，与西部的单倍型间存在三步或三步以上的突变。

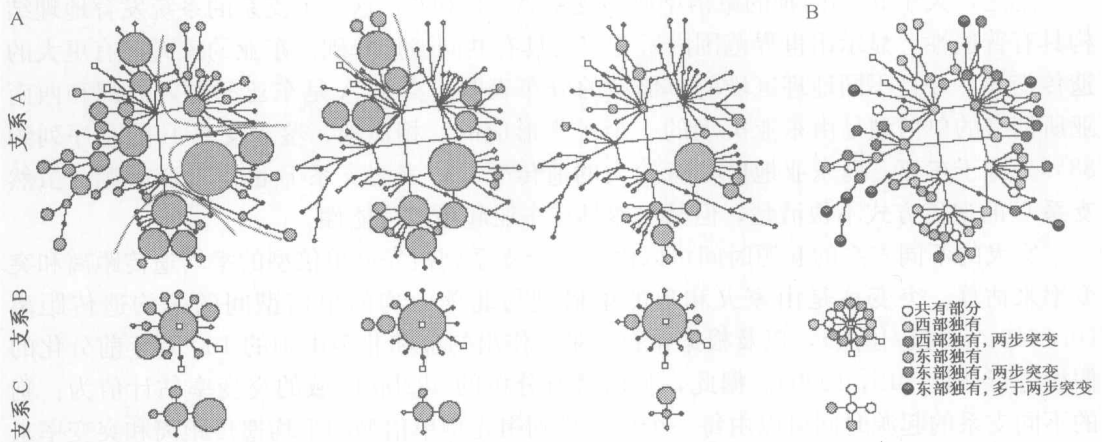


图 1-2 狗不同 mtDNA 单倍型之间的最小跨度网络关系图 (Savolainen et al., 2002)  
(见彩色图版)

不同群体间的遗传变异方式可以采用最小跨度网络法进行较为细致的研究 (图 1-2A)。在支系 A 中，东亚所具有的单倍型分布于网络图的各处，而欧洲和西南亚所具有的单倍型仅分布于网络图的部分区域，大部分区域处于“空白”状态。为进一步研究东亚所具有的较大的遗传变异和家犬东亚起源（而非欧洲或西南亚起源）的可能性，我们以喜马拉雅山脉和乌拉尔山脉为界，将旧大陆划分为东部和西部两个地区，并就这两个地区进行了比较。在支系 A 中，13 个单倍型为东西部共享，35 个单倍型为东部独有，23 个单倍型为西部独有 (图 1-2B)。那些欧洲和西南亚单倍型的出现频率，在东西部地区也几乎相同。西部所共享的比例 (71.9%) 要远远大于东部 (48.5%)，而且不同的地区也表现出相同的趋势。例如，在西部单倍型共享的比例都很高。图 1-2B 也表现了分别由东西部独享的单倍型的情形，以及一个单倍型与其他地方的单倍型存在起码两步突变的情况。在东部，有 19 个单倍型与西部的某一个单倍型存在起码两步的突变；而在西部，只有三个单倍型与东部的单倍型存在两步的突变。而且，在东部有 8 个单倍型与西部存在起码三步的突变，并且有两个存在 4 或 5 步的突变；但是在西部，没有一个单倍型是与东部单倍型存在三步或三步以上的突变。并且，来自于西部的单倍型要么与东部的单倍型相同，要么与东部的单倍型保持最多两步的突变，这表明欧洲或西南亚的单倍型是由东亚单倍型的一个子集形成的。与西部单倍型保持起码三步突变的东部单倍型来自于泰国 ( $n=2$ )、柬埔寨 ( $n=1$ )、中国 ( $n=2$ )、日本和韩国 ( $n=1$ )。由此可以看到，东亚和西南亚有非常大的遗传分化。

相似地，在支系 B 中东部所拥有的单倍型比西部的多，且独有单倍型也较多。一个主体单倍型连接两个共享单倍型这种结构是东西部的一种普遍形式，同时东部拥有 7 个独有的单倍型，而西部只拥有 3 个 (图 1-2B)。而且，西部所拥有的共享单倍型的比

例 (93.2%) 要比东部 (58.8%) 的高, 并且其亚区也表现出相同的方式。在这个支系中, 两两序列平均遗传距离也是东亚 (0.93) 大于欧洲 (0.45) 和西亚 (0.36)。支系 C 的序列变异程度普遍较低, 但同样呈现这种趋势, 构成支系 C 的单倍型序列在西部都是与东部共享的, 而东部的单倍型中有两个是独享的 (图 1-2)。

总之, 大于 95% 的狗的单倍型属于支系 A、B 和 C, 这三个支系的系统发育地理结构具有普遍性, 显示出世界范围内所有家犬具有共同遗传基础。东亚的家犬拥有更大的遗传变异, 并且不同地理区域的单倍型的分布表明, 支系 A 是东亚起源, 欧洲和西南亚所拥有的单倍型是由东亚类型的一个子集形成的。相似地, 鉴于支系 B 的总序列约 88% 来源于东亚, 而东亚地区具有较大的遗传变异, 表明支系 B 也是东亚起源。虽然支系 C 的遗传方式不很清楚, 但是仍然具有东亚起源的可能性。

家犬的不同支系的起源时间可以由每一个支系到祖先型单倍型的平均遗传距离和突变率来估算。突变率是由家犬和狼的单倍型与北美山狗的单倍型间的平均遗传距离 (0.1419 个替换/位点), 以及根据化石记录所作出的狼和北美山狗的 100 万年前分化的假设进行估计 (Li, 1990)。据此, 我们对所分析的 582bp 区域的突变率估计值为: 狗的不同支系的起源时间可以由每一个单倍型到祖先型单倍型的平均遗传距离和突变率来计算, 大致为每百万年 ( $7.1 \pm 0.4$ )%。

家犬在被驯化后会出现一个群体扩张, 在系统发育上表现为: 一个建群单倍型处于中央, 新形成的单倍型呈放射状的“星状”系统发育关系。支系 B、C 就表现为这种结构, 表明其是从一个狼的单倍型起源的。支系 A 相对复杂, 看不出一个共同的起源中心。在支系 A 存在一个 11 步的替换, 表明支系 A 比支系 B 和 C 更加古老, 可能是最初由狼驯化形成的一个单倍型类群。在支系 A 中, 有几个亚类群呈现星状结构, 表明支系 A 可能是来源于几个狼的单倍型。目前的数据尚不能说清支系 A 是由多少个狼的单倍型建群的。如果把支系 A 假定为单一起源的话, 在东亚群体中, 582bp 的 DNA 序列的碱基平均替换数为 3.39。根据每百万年 7.1% 的进化速率, 得到支系 A 的形成年代在 40 000 年前。如果假设支系 A 为多起源, 那么有 3 个亚类群可以用以推测支系 A 的年代, 它们的碱基替换数分别为 0.45、0.65 和 1.07, 相应的推算年代为 11 000、15 500 和 25 500 年。支系 B 和 C 也是星状结构, 并且假定是单一单倍型的狼建群的, 由东亚序列到节点的平均距离 (0.54 和 0.71 个替换) 所估计的支系 B 和 C 的产生年代分别为 13 000 年和 17 000 年。

根据现有数据推测, 如果将支系 A 作为最先出现的家犬的支系, 其形成时间应在 40 000 年前。如果同时考虑支系 A、B 和 C, 那么家犬的形成时间应在 15 000 年前。然而, 欧洲的属于支系 A 的最古老的亚类群, 按 0.39 个替换, 所估计的驯化年代为 9000 年前 (图 1-2A)。以支系 A 最早可能的形成年代 40 000 年为依据, 东亚家犬的独立形成时间最长, 之后才逐渐扩散到世界各地。考古学的证据支持支系 A 和 B (支系 C 也可能) 同时在东亚起源, 时间大约发生在 15 000 年前。

考古学和形态学的证据似乎支持上面这样的假说。因为绝大多数的考古学证据都是残片, 并且小体型的狼与家犬又很近似, 因此有关犬科动物与人的关系的一些解释都相当模糊。中国的有关家犬的考古学证据都是新石器时代的, 时间可以追溯到 7500 年前 (Nobis, 1979), 旧石器时代的确凿证据相对缺乏。考虑到东亚的考古学证据的有限性,

虽然在这一地区缺乏旧石器时代的考古学证据,但是也不能排除家犬的东亚起源的可能性。以色列的考古学证据显示家犬出现在12 000年前(Tcherncov and Valla, 1997; Clutton-Brock, 1999),明显为家犬的证据特征出现在9000年前(Benecke, 1987),在北美出现在8500年前(Benecke, 1987; Swofford, 2002)。总之,家犬的考古学证据说明家犬的驯化事件发生在9000~14 000年前。我们的线粒体DNA的数据显示,家犬的驯化时间应在15 000~40 000年前。将两者的起源时间综合考虑,家犬的起源时间应为15 000年前。本次研究表明,支系A很可能是由几个狼的单倍型形成的。狼的驯化时间不大可能是一个孤立的事件,更为可能的是人类在探索过程中的一个共同实践过程。

通过本次研究,我们提供了家犬的东亚起源的证据。更为准确的家犬起源时间的确定有赖于线粒体基因组的大规模测序。而且,基于母系遗传物质线粒体DNA的研究已经表明家犬具有一个共同的起源地。但是,随后发生的母家犬与公狼的杂交对地方家犬品种形态的影响目前尚不清楚。通过Y染色体DNA的分析将能够揭示北极品种似狼的形态特征是由家犬和狼杂交形成的,还是单纯是一个品种特征。中国和瑞典科学家的研究表明,狗起源于东亚,之后逐渐扩散到世界各地。这为揭示狗的起源之谜,为狗的进一步遗传学研究奠定了基础。

## 第二节 中国家养山羊的遗传多样性和系统发育地理结构研究

家养山羊(*Capra hircus*)为人类社会提供了一系列有用的动物性产品如肉、奶及毛等,这使得其成为人类曾经驯化过的最为有用的动物之一(Porter, 1996)。在人类文明早期,山羊在农业、经济、文化乃至宗教方面也起过重要作用(Joshi et al., 2004)。考古学方面的证据显示,山羊最早可能在大约1万年前于近东新月地带被驯化(Porter, 1996; Pringle, 1998; Zeder and Hesse, 2000);而其他一些研究提示,在巴基斯坦发生了一次独立的驯化事件,产生了类似绒山羊的品种(Meadow, 1993; Porter, 1996)。虽然从新石器时代起山羊就在全世界范围内起过重要作用,但有关家养山羊的起源问题仍然不太清楚。

线粒体DNA(mtDNA)为解开家养动物起源之谜提供了最为丰富的信息。目前,线粒体DNA序列被广泛应用于研究牛、猪、绵羊、马、狗、驴及山羊等家养动物的起源问题。前人的研究结果显示,在家养山羊中存在至少4个主要的线粒体DNA世系A、B、C及D(Joshi et al., 2004; Luikart et al., 2001; Sultana et al., 2003)。世系A的多样性最为丰富,广泛分布于所有大陆;世系B仅分布于东亚和南亚,包括蒙古、老挝、马来西亚、巴基斯坦和印度;世系C以较低的频率出现在蒙古、瑞士、斯洛文尼亚、巴基斯坦和印度;世系D的频率最低,仅在巴基斯坦和印度的地方山羊品种中检测到。

这4个世系间的分歧时间都在20年以上,远早于家养动物驯化的时间1万年左右(Joshi et al., 2004; Luikart et al., 2001; Sultana et al., 2003)。结合以前的考古学发现,这些结果提示家养山羊有多个母系起源(Porter, 1996; Pringle, 1998; Zeder and Hesse, 2000)。然而,进一步的取样,特别是那些可能包含小的mtDNA世系的地区,将会增加我们对山羊起源地的认识(MacHugh and Bradley, 2001)。中国所

处的地方就是这样的地区之一，而且尚未对中国家养山羊开展遗传学方面的详尽研究。中国拥有许多地方类型的、多用途的土著山羊品种，分布范围从干旱、寒冷及环境恶劣的高海拔地区（青藏高原）到潮湿、温暖及环境温和的低海拔地区（中国南方）。由于环境条件相差极大，因而不同土著品种间呈现出许多表型差异。中国古代广大的西南山区曾被认为是山羊的一个驯化中心（谢成侠，1985）。可以预期的是，对中国山羊土著品种开展分子水平上的研究将有助于了解山羊的起源与驯化过程，并促进全世界山羊系统发育关系的研究。由 Luikart 等（2001）所开展的全世界范围的研究仅包含来自中国的两个山羊品种的两个个体。虽然过去对中国家养山羊的遗传多样性方面的研究也有所开展（李祥龙等，2000a；2000b），但对它们的起源与进化仍然知之甚少。在本研究中，我们测定了来自 18 个土著山羊品种共 368 个个体的 mtDNA 控制区第一高变区（HVSI）的 DNA 序列，用它们来研究中国家养山羊的遗传多样性、起源与系统发育地理结构。

### 一、样品采集与 PCR 扩增及测序

我们从 10 个省、自治区的偏远山区共收集了没有直接血缘关系的 368 份血液样品

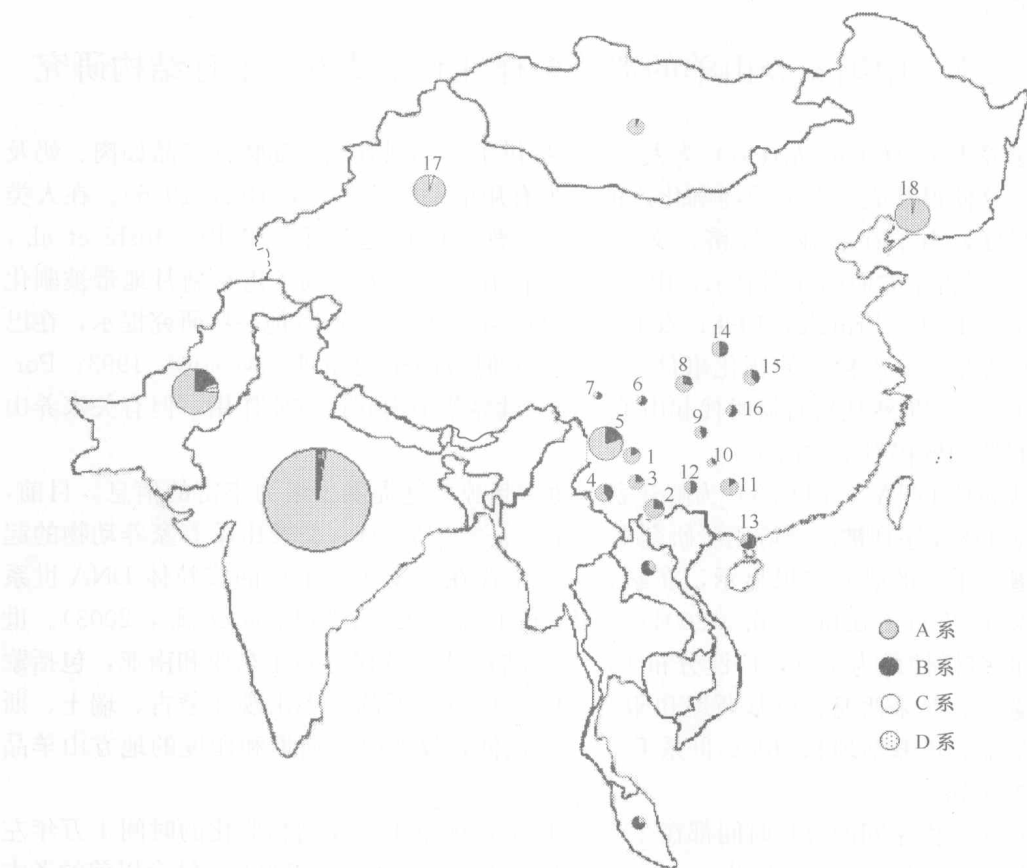


图 1-3 本研究中的中国山羊样本的地理分布与每个世系在各品种中的分布频率，以及世系 B 在亚洲国家中的分布频率图。图上编号对应的品种名称见表 1-2，圆圈区域大小与样本大小成比例。

(图 1-3), 代表了 18 个中国土著山羊品种, 详细信息见表 1-2。我们采用标准酚/氯仿法从血液提取总基因组 DNA 后, 对 mtDNA 控制区第一高变区 (HVSI) 一个长 481bp 的 DNA 片段进行了 PCR 扩增和序列测定, 并选取了来自世系 A、B 和 C 中任意具有不同控制区 HVSI 序列的 2 个个体和来自世系 D 的 1 个个体共 7 个山羊个体, 扩增并测定了 mtDNA 1140bp 的细胞色素 b (cyt b) 基因全序列。

表 1-2 样本信息及每个品种的单倍型数、单倍型多样度和核苷酸多样性

地图 编号	品种名称	地理分布	样本 大小	单倍 型数	单倍型多样性 ( $\pm$ SE)	核苷酸多样性 ( $\pm$ SE)
1	黑头山羊	西南地区, 云南	28	16	0.9418 $\pm$ 0.0250	0.0279 $\pm$ 0.0144
2	无角山羊	西南地区, 云南	26	14	0.9323 $\pm$ 0.0280	0.0351 $\pm$ 0.0180
3	圭山奶山羊	西南地区, 云南	18	15	0.9804 $\pm$ 0.0243	0.0315 $\pm$ 0.0165
4	龙陵黄山羊	西南地区, 云南	24	10	0.9130 $\pm$ 0.0285	0.0352 $\pm$ 0.0181
5	云岭黑山羊	西南地区, 云南	53	25	0.9398 $\pm$ 0.0179	0.0329 $\pm$ 0.0165
6	成都麻羊	西南地区, 四川	8	5	0.8571 $\pm$ 0.1083	0.0355 $\pm$ 0.0202
7	藏山羊	西南地区, 四川	7	5	0.8571 $\pm$ 0.1371	0.0490 $\pm$ 0.0282
8	川东白山羊	西南地区, 四川	14	6	0.7473 $\pm$ 0.1114	0.0228 $\pm$ 0.0124
9	贵州白山羊	西南地区, 贵州	14	7	0.9011 $\pm$ 0.0465	0.0346 $\pm$ 0.0184
10	贵州黑山羊	西南地区, 贵州	9	6	0.9167 $\pm$ 0.0725	0.0293 $\pm$ 0.0165
11	都安山羊	南部地区, 广西	14	11	0.9670 $\pm$ 0.0366	0.0292 $\pm$ 0.0156
12	隆林山羊	南部地区, 广西	10	7	0.9111 $\pm$ 0.0773	0.0391 $\pm$ 0.0214
13	雷州山羊	南部地区, 广东	15	8	0.8762 $\pm$ 0.0595	0.0295 $\pm$ 0.0157
14	陕南白山羊	北部地区, 陕西	16	7	0.8417 $\pm$ 0.0595	0.0305 $\pm$ 0.0162
15	宜昌白山羊	中部地区, 湖北	13	7	0.8974 $\pm$ 0.0537	0.0329 $\pm$ 0.0177
16	马头山羊	中部地区, 湖南	12	3	0.7121 $\pm$ 0.0691	0.0321 $\pm$ 0.0174
17	新疆山羊	西北地区, 新疆	40	24	0.9744 $\pm$ 0.0099	0.0243 $\pm$ 0.0125
18	辽宁绒山羊	东北地区, 辽宁	47	21	0.9204 $\pm$ 0.0230	0.0159 $\pm$ 0.0084

## 二、中国家养山羊的 mtDNA 核苷酸多样性分析

我们对中国山羊的 368 条 HVSI 区段序列、7 条来自野生山羊物种的同源序列 (GenBank 收录号: AJ317864、AJ317870 ~ AJ317872、AJ317874、AJ317875 和 AB110590), 以及 3 条来自巴基斯坦山羊中属于世系 D 的序列 (GenBank 收录号: AB110587 ~ AB110589) 用 Clustal W (Thompson et al., 1994) 程序进行序列比对。为了研究以前确定的家养山羊世系 B 的起源问题, 已发表属于世系 B 的所有序列 (表 1-3) 都包括在我们的分析里面。

结果显示, 所有 368 条 mtDNA 控制区 HVSI 的 481bp 序列中没有碱基插入和缺失。根据这 368 条序列的 119 个变异位点 (其中只有 1 个是颠换) 确定了 146 个单倍型。其中, 最大的一个单倍型由 20 个个体组成, 有 3 个单倍型包括十余个个体, 有 2 个单倍型包括 10 个个体。每个品种中检测到的单倍型数目为 3 ~ 25 个; 单倍型多样性值从马头山羊的 0.7121  $\pm$  0.0691 到圭山奶山羊的 0.9804  $\pm$  0.0243; 核苷酸多样性值 ( $\pi$ ) 在藏山羊中最高 (0.0490  $\pm$  0.0282), 而在辽宁绒山羊中最低 (0.0159  $\pm$  0.0084) (表 1-2)。



表 1-3 世系 B 在亚洲国家山羊样本中的个体数目及其比例、单倍型个数、独有类型个数和数据来源信息

国家	个体数 (占有所有样本的比例 %)	单倍型数 (独有类型)	数据来源
中国	92 (25)	25 (23)	本研究
蒙古	1 (6.7)	1 (0)	Luikart et al. (2001)
老挝	8 (66.7)	5 (4)	Mannen et al. (2001)
马来西亚	12 (75)	4 (3)	Luikart et al. (2001)
巴基斯坦	15 (19.2)	5 (4)	Luikart et al. (2001)和 Sultana et al. (2003)
印度	6 (1.6)	5 (4)	Luikart et al. (2001)和 Joshi et al. (2004)
总计	134	45 (38)	

当把中国家养山羊的 mtDNA 变异水平同印度和巴基斯坦的家养山羊进行比较时,发现中国家养山羊 ( $\pi = 0.0352 \pm 0.0174$ ) 的多样性水平高于印度山羊 ( $\pi = 0.0212 \pm 0.0108$ ), 但稍低于巴基斯坦山羊 ( $\pi = 0.0365 \pm 0.0184$ )。印度山羊和巴基斯坦山羊的核苷酸多样性值 ( $\pi$ ) 是基于对前人数据的重新分析得到的 (Joshi et al., 2004; Sultana et al., 2003)。有趣的是, 对 Luikart 等 (2001) 数据的重新分析的结果表明, 中国家养山羊的核苷酸多样性值还高于来自旧大陆的所有山羊样本的核苷酸多样性值 ( $\pi = 0.0278 \pm 0.0139$ )。

### 三、系统发育关系树的构建

基于似然比检验, 用 Modeltest 3.06 (Posada and Crandall, 1998) 挑选出适合于我们数据的最优模型是 HKY+I+ $\Gamma$ 。用 PAUP\* 4.0 b10 (Swofford, 2003) 构建了 1 个邻接树 (neighbor-joining tree), 树节点的支持率采用 1000 次 Bootstrap 重抽样 (Felsenstein, 1985) 获得。贝叶斯分析由 MrBayes 3.0 b4 (Huelsenbeck and Ronquist, 2001) 程序完成。从两个不同的随机树开始独立运行 MCMC 两次, 每次运行 500 万代, 同时运行 4 条链。在达到稳定之前的所有抽样全部丢弃, 剩余的抽样用于构建贝叶斯主要一致树和节点的后验概率。

同以前确定的世系 A、B、C (Luikart et al., 2001) 和 D (Joshi et al., 2004; Sultana et al., 2003) 的单倍型相比较, 基于中国山羊的 146 个单倍型、7 条野生物种的序列和 3 条属于世系 D 的序列构建的邻接树显示, 中国家养山羊可分为 4 个明显 mtDNA 世系 A~D (图 1-4)。除了内部分枝的一些小的地方不一致外, 贝叶斯树的拓扑结构与邻接树的拓扑结构基本一致。世系 A、B 和 C 分别包含 117、25 和 3 个单倍型, 并分别由 269、92 和 6 个个体组成; 而世系 D 仅有 1 个单倍型来自 1 个个体。

为了详细了解中国家养山羊各单倍型之间的可能关系, 我们用 Network 4.0 (www.fluxus-engineering.com) 构建了中国山羊 146 个单倍型之间的中介网络图 (Bandelt et al., 1999) (图 1-5)。单倍型多样性及其标准误、核苷酸多样性、误配分析 (Schneider and Excoffier, 1999) 和分子变异分析 (AMOVA) (Excoffier et al., 1992) 都用 Arlequin 2.0 (<http://anthropologie.unige.ch/arlequin>) (Schneider et al., 2000)