

## 独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得 浙江大学 或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：曾长军

签字日期： 2007 年 8 月 27 日

## 学位论文版权使用授权书

本学位论文作者完全了解 浙江大学 有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权 浙江大学 可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后适用本授权书）

学位论文作者签名：

导师签名：

签字日期：      年   月   日

签字日期：      年   月   日

学位论文作者毕业后去向：

工作单位：

电话：

通讯地址：

邮编：

## 中文摘要

大熊猫 (*Ailuropoda melanoleuca*) 是我国特有的珍稀濒危野生动物。长期以来, 栖息地片段化和种群数量下降, 致使其野外仅分布于秦岭、岷山、邛崃、大相岭、小相岭和凉山等 6 个相互隔离的山系, 并分化为两个独立的亚种——秦岭亚种和指名亚种。

为抢救性的保护大熊猫这一珍稀濒危物种, 我国政府于 2003 年启动了大熊猫的离体保护工程。该工程的主要内容之一, 就是通过建立大熊猫的基因组 DNA 文库, 达到保护其遗传资源, 并为相关的基础科学研究提供基因材料平台的目的。

本论文以细菌人工染色体 (BAC) 为载体, 建立了大容量的大熊猫基因组 DNA 文库。进一步地, 通过对大熊猫 II 类主要组织相容性复合体 (MHC) 基因组物理图谱的构建, 既验证了该文库的有效性, 又对文库进行了基础研究的应用, 达到了立题目的。其主要研究结果如下:

1. 以新鲜外周血为材料, 利用 pCC1BAC 载体, 成功构建了大熊猫的基因组细菌人工染色体 (BAC) 文库。文库包含的 205,800 个 BAC 克隆, 保存于 2,100 块 96 孔培养板上。经对随机挑取的 174 个 BAC 克隆进行脉冲电泳检测, 计算出文库的平均插入片段为 97kb, 覆盖大熊猫基因组约 6.8 倍 (除去 5.74% 的空载体)。同时, 从文库中获得任一单拷贝或多拷贝大熊猫基因的概率为 99.93%。
2. 16 对 PCR 引物平均每对能够在文库中成功扩增到 6.4 个阳性克隆, 表明了所建文库具有较高的覆盖率。
3. 运用 4D-PCR 文库筛选技术, 建立了大熊猫基因组 BAC 文库的高效 PCR 筛选文库。该文库包括 43 个超级库和 1462 (43×34) 个 1D、2D、3D、4D 超级亚库。任一引物或探针只需 77 个 PCR 反应, 即可从 BAC 文库中筛选到所需的阳性 BAC 克隆。
4. 以大熊猫皮肤成纤维细胞为材料, 运用染色体荧光原位杂交 (FISH) 技术, 成功地将大熊猫的 MHC II 类基因组定位于第 9 号染色体的长臂上。

5. 通过 GenBank 数据库中的同源序列，设计了 10 对大熊猫 MHC II 类基因的特异性引物，并成功扩增到了大熊猫 II 类 MHC 的 BTL2、DRA、DQA、DRB1、DOB、LMP2、DMA、DMB、COL11A2 和 DAXX 基因（基因序列已提交 GenBank 数据库）。
6. 运用 BAC 克隆的末端测序方法，并利用上述 10 对引物，构建了大熊猫 MHC II 类基因组的物理图谱。进一步地，利用限制性酶切指纹图谱法，证明了所构建的物理图谱中的 BAC 克隆是有效重叠的，其 MHC II 类基因组的长度约为 650kb。

**关键词：**大熊猫；基因组 BAC 文库；染色体荧光原位杂交；MHC 基因组染色体定位； MHC II 类基因组物理图谱；限制性酶切指纹图检验

## Abstract

The giant panda (*Ailuropoda melanoleuca*) panda is rare and endangered species endemic to China. Unfortunately, due to habitat fragmentation, the population size of giant panda decline rapidly. Now, the wild giant pandas distribute six completely isolated mountain ranges such as Qinling Mountain, Minshan Mountain, Qionglai Mountain, Xiangling Mountain and Liangshan Mountain. Moreover, the giant panda population shows significant genetic differentiation and yield two evolutionarily significant units (ESU), a new Qinling subspecies from the nominate Sichuan subspecies.

In order to protect this rare and endangered species, the Chinese government implements the gene resources conservation project followed by the *in-situ* and *ex-situ* conservation projects of giant panda in 2003. One goal of the project is conserves the genetic resources by constructing the genomic DNA library and provide the gene resources platform for related scientific research.

As the main part of giant panda gene resources conservation project, the present study is to construction and characterization of high coverage genomic DNA library. Then further researchs show that this genomic BAC library is high quality and effective by constructing physical map of MHC class II of giant panda. The main results are as follows:

1. The giant panda bacterial artificial chromosome (BAC) library consisting of 205,800 clones has been constructed using the whole blood and pCC1BAC vector. The average insert size was calculated to be 97 kb based on the examination of 164 randomly selected clones, indicating that the giant panda library contained 6.8-fold genome equivalents. About 5.74% non-insert clones were observed in the 174 clones tested (10 out of 174). The probability of any giant panda gene being found in this library should be about 99.93%.
2. Two or more positive clones (2 ~ 16) were found for these 16 loci, with an average of 6.4 positive clones per locus. The PCR screening results were thus

- in good agreement with an expected 6.8-fold genomic coverage of the library.
3. Construction of high-effective rapid screening libraries by 4D-PCR. The 205,800 BAC clones were divided into 43 superpools and 1462 (43×34) sub-superpools of 1D, 2D, 3D, 4D. We can obtain the desired positive BAC clones by any primers or probes through 77 PCR reactions.
  4. The giant panda metaphase chromosomes were prepared from fibroblast cells of giant panda. The results indicate that the MHC is located on the chromosome 9q by fluorescence *in situ* hybridization.
  5. We have designed 10 giant panda MHC class II gene PCR primers by the homologs counterpart of the MHC class II in the GenBank database. All the primers were amplified successfully and the PCR products of BTL2, DRA, DQA, DRB1, DOB, LMP2, DMA, DMB, COL11A2 and DAXX were submitted to GenBank.
  6. The physical map of MHC class II of giant panda was constructed by screening the genomic BAC library using 10 MHC class II gene primers. The end sequencing primers were designed according to positive BAC clones and screen further the giant panda genomic BAC library to cover the gaps of physical map. According to the restriction fingerprint analysis and test, the whole physical map of MHC class II region was spanning about 650 kb.

**Keyword:** giant panda; genomic BAC library; fluorescence *in situ* hybridization; chromosome mapping of MHC; physical map of MHC class II; restriction fingerprint analysis

# 第一章 文献综述

## 第一节：基因组学研究的主要内容和方法

基因组 (Genome) 是 Winkler 于 1920 年首次提出的 (Winkler, 1920)，它表示的是一个物种全部的遗传物质的总和, 包括了所有的基因和基因与基因之间的间隔区域。

### 1. 基因组学的定义及其内容

一个物种的基因组蕴藏着它的生命和 (或) 繁殖所需的全部生物学信息。而基因组学 (Genomics) 则是现代分子生物学中解析基因组结构、功能及其相关信息分析的一门前沿分支学科 (Passarge, 2001; Hieter & Boguski, 1997)。1987 年 “*Genomics*” 专业学术期刊的创刊, 标志着这一新兴前沿学科的诞生。

基因组学与遗传学 (genetics) 的区别在于, 前者包含了与分子和细胞生物学有关的诸多方面, 如基因组图谱、核酸测序、组装、储存、数据管理、基因识别、功能预测、基因组进化和众多生物基因组之间的比较等等 (Passarge, 2001)。而后者则主要研究遗传 (Heredity) 及其相关的机制和结果等。因此, 基因组研究的主要内容, 就包括了以全基因组测序为目标的结构基因组学 (structural genomics) 和以功能鉴定为目标的功能基因组学 (functional genomics) 两大方面。其中, 结构基因组学是基因组研究的早期阶段, 主要以建立物种高分辨率的遗传、物理和转录图谱, 以及蛋白质或生物大分子的三维结构为主 (Teichmann et al., 1999; Blundell & Mizuguchi, 2000), 而功能基因组学则是利用结构基因组学提供的信息, 系统地研究基因的功能, 它以高通量、大规模实验以及计算机辅助分析等为其主要特征 (Hieter & Hieter, 1997; 赵剑华等, 2000)。

## 2. 结构基因组学研究的主要内容和方法

### 2.1 遗传图谱

遗传图谱 (Genetic map) 又称为连锁图谱或遗传连锁图谱, 是以具有遗传多态性的遗传标记为“路标”, 以基因座之间的重组率 (1%的重组率为1cM) 为“图距”的基因组图 (方福德等,1998)。遗传图谱的绘制是基因组学研究的第一步, 它以染色体上某一位点为遗传标记, 并辅以与之相对应的遗传性状, 经过连锁分析, 将编码该性状的基因定位于特定的染色体上。

自Botstein等人根据RFLP (限制性酶切片长度多态性) 的特点提出用RFLP等DNA标记来构建遗传图谱 (Botstein et al.,1980) 以来, 遗传图便逐渐成为了基因克隆与定位等相关研究的重要工具, 并在研究基因的功能、定位及分离克隆新基因、排列DNA片段和研究染色体上基因的顺序等方面发挥着重要作用 (Lander & Botstein, 1985; Gebhardt et al.,1995)。如Saito等利用RFLP构建了人6号染色体长臂的遗传连锁图, 对于分离克隆潜在的肿瘤抑制基因以及其它遗传疾病的研究具有重要的意义 (Saito & Nakamura, 1993)。但是, 由于RFLP在染色体组中的分布极不均匀, 且提供的多态信息十分有限, 因此, 要绘制高密度的遗传图是非常困难的 (Botstein et al.,1980)。于是, 在RFLP的基础上, 结合如AFLP、RAPD、生化标记、性状标记位点及以简单序列长度多态性(SSLP)标记如微卫星、小卫星等为基础的第二代遗传图谱 (Dobuvsky et al.,1995; Yuan et al.,1997), 绘制出了更加精细的遗传图 (Menz et al.,2002; Ouedraogo et al.,2002; Ferreira et al.,2000; Whisson et al.,1995; Kiss et al.,1993)。如Ouédraogo等就利用AFLP、RFLP、RAPD、生化标记以及一些生物性状位点构建了豇豆的遗传连锁图, 对于豇豆育种中的标记辅助选择、图位克隆等研究具有重要的指导价值 (Ouedraogo et al.,2002), 由此使遗传图谱显示出了更高的应用价值。

随着基因组研究的进一步深入, 该领域又相继出现了以SNPs (单核苷酸多态性) 为标记的第三代遗传图谱 (Garg et al.,1999)。SNP是指在染色体基因组水平上由于单个核苷酸的变异而引起的DNA序列多态性, 包括单碱基的转换、颠换、插入以及缺失等 (Vignal et al., 2002)。由于SNPs可以在基因组的编码区和非编码区广泛存在, 且其标记数目多, 覆盖密度大, 因而在基因定位等方面的

研究中，第三代遗传图谱就有着其他标记系统无可比拟的优越性和发展潜力 (Davis & Hammarlund, 2006)。

## 2.2 物理图谱

物理图谱 (physical map) 是以特定的DNA序列为标记所构建的染色体图，其中最粗略的物理图谱是染色体的组型图，如Tian等通过对小熊猫(red panda)核型分析，表明小熊猫具有 $2n=36$ 条染色体 (Tian et al., 2002)。而最精确的物理图谱便是核苷酸的序列图，人类基因组全序列等。限制性内切酶的发现，促进了第一个物理图谱SV40的完成 (Danna & Nathans, 1971)。继后，随着YAC、BAC/PAC等新克隆技术的建立，以及克隆片段长度STS (序列标签位点) 成为有力的物理图标，使得相关物理图谱的制作水平得到了很大的提高，从而保证了物理图谱的连贯性 (Olson et al., 1989; 于军, 1998)。目前，常见的物理图谱有STS图谱、YAC (酵母人工染色体) 图谱、PAC (噬菌体人工染色体) 图谱，以及BAC (细菌人工染色体) 图谱等。迄今，构建物理图谱的方法主要有以下几种：

**a) 限制性内切酶作图：**该方法是最经典的物理图谱制作法之一。当一个克隆或DNA片段被限制性内切酶酶切成不同长度的片段时，便形成了通常所称的DNA指纹 (DNA Fingerprint)，但这种指纹图谱是没有顺序的。这类比较有代表性的图谱是*E.coli*、*C.elegans* 和酵母的物理图谱 (Kohara et al., 1987; Coulson et al., 1986; Olson et al., 1986)。继后，在单酶切法的基础上发展起来的多酶完全酶切法 (Multiple-Complete-Digest, MCD)，则比上述的DNA指纹图更具有准确、可自动化和规模化的特点，已被广泛地应用于酵母、人等的大规模测序和精细物理图谱的研究工作中 (Gillett et al., 1996; Yu & Wong, 1997; Wong et al., 1997)。

**b) 序列标签位点 (STS) 作图：**Olson等人根据物理图谱的制作需要，提出并建立了序列标签位点 (STS) 的作图技术 (Olson et al., 1989)。STS的来源包括EST、已知遗传标记、随机测序序列和末端测序等，从而使得低分辨率物理图谱的制作水平发生了新的飞跃。1997年人类第7号染色体和X染色体物理图谱的制作，就是基于该技术完成的，其分辨率达到了100kb (Bouffard et al., 1997; Nagaraja et al., 1997)。其后，愈来愈多的基于STS的高密度和高分辨率的物理图谱得以顺利完成 (Baysal et al., 1997; Mumm et al., 1998; Blanco et al., 1998;

Srivastava et al., 1999), 从而使STS作图技术成为了物种遗传物理图谱制作研究的首选技术。

c) 其它物理图谱作图法: 除了以上2种作图法以外, 以原位杂交(Klockars et al., 1996)为标记的光学作图或染色体显微切割及辐射杂交作图(Radiation-hybrid mapping, RH), 也是制作物理图的常用方法(Robic et al., 2001; Martins-Wess et al., 2003)。其中, RH作图法由于可以产生高分辨率的物理图, 因而可以用于多个物种之间共线性区域保守(Synteny conservation)程度的研究(Goldammer et al., 2002)。

## 2.3 基因组测序

随着遗传图谱和物理图谱的完成, 以及荧光自动分析仪的发明(Hunkapiller et al., 1991), 使得物种大规模基因组测序的工作成为了可能。1977年在英国Sanger实验室首先完成了第一个生命基因组(Phi X174)序列的测定(Sanger et al., 1977)。基因组测序的最终目标, 是获得目的生物体基因组的全部DNA序列。对于较小的基因组如噬菌体、BAC、Cosmid、Fosmid、细菌和质粒, 通常采用的方法是随机鸟枪法测序。它通过序列的拼接和组装, 完成基因组的完整序列(Anderson et al., 1981; Baer et al., 1984; Akanuma, 2002)。对于较大的基因组而言, 则可以采用全基因组鸟枪法测序(Whole Genome Sequencing, WGS)(Fleischmann et al., 1995; Moxon, 1995; Webber & Myers, 1997)、克隆重叠群(Clone by clone)基础上的鸟枪法测序(Cai et al., 2001; Chen et al., 2004; Khorasani et al., 2004), 以及随机定向测序(Directed Sequencing)战略。定向测序技术是充分利用鸟枪法快速测序和物理图谱的有序排列优点, 并在克服了上述两种测序方法的不足后发展起来的。目前, 该方法已得到了广泛的应用(Malykh et al., 2004)。

## 3 功能基因组学研究的主要内容和方法

近年来, 随着人类和一些动物、植物及微生物基因组计划的完成, 基因组学的研究便步入了后基因组时代——功能基因组学(Functional Genomics)

(Woychik et al., 1998)。功能基因组学的主要研究内容,就是在结构基因组学丰富信息资源的基础上,利用大通量的试验方法并结合统计与计算机分析,研究生物在其生长、发育和繁殖等生命过程中相关基因的功能、表达调控、基因与蛋白质之间的相互作用等(Hieter & Boguski, 1997)。

### 3.1 基因的功能、表达及调控研究

在不同的组织和发育阶段,基因的表达模式是不同的。由于常规的RT-PCR技术、RNA印迹杂交技术等方法一次只能研究一个或几个基因的功能和表达情况,因此,高通量的基因表达分析技术——微阵列(DNA Microarray)及基因表达系列分析(Serial analysis of gene expression, SAGE)便应运而生。DNA微阵列技术始于1995年Schena等用于拟南芥45个基因表达模式的定量分析研究(Schena et al., 1995),并多用于基因表达研究及点突变 / SNP分析(Heller, 2002)。尽管所有的DNA微阵列都是基于核酸序列间的杂交互补,但是,其技术选择在不同的微阵列平台间却存在着广泛的差异(Hardiman, 2004)。这些平台之间最主要的差别在于探针的长度(Draghici et al., 2006)。微阵列可分为两种:其一是cDNA微阵列,它的探针通常来自于PCR扩增的产物,长度可达数千个bp。其二是寡核苷酸微阵列,分为短寡核苷酸微阵列(25-30mer)和长寡核苷酸微阵列(60-70mer)(或更长序列)。此外,DNA微阵列技术除了在分子生物学和基因组学中的应用外,还在药物基因组学、蛋白质组学、传染病、遗传疾病和癌症的诊断、法医学和遗传识别等方面具有重要的作用。

为了大规模测定基因的表达,1992年Okubo等人提出了基因表达的物理图谱的概念(Okubo et al., 1992)。其主要的策略就是测定cDNA 3'端的部分序列,比较不同组织类型的细胞中cDNA的种类和数量。1995年Velculescu等人提出了SAGE——能够定量和同时分析大量的基因表达(转录本)的方法(Velculescu et al., 1995)。它提供了广泛的定量分类和比较不同发育阶段基因表达情况的分析手段。SAGE技术区别于差异显示、消减杂交等方法的主要特点在于,它可寻找低丰度、低表达水平的转录本,最大限度地收集基因组的表达信息。Caron等人运用SAGE技术构建了人类全基因组水平上的mRNA表达谱,该表达谱包含了人类12种类型的组织的245万个转录标签,发现了高表达丰度的基因成簇分布于染色体的特定

区域(Caron et al., 2001), 同时也为在癌症研究中那些过量表达或沉默表达基因的研究提供了有效的分析工具。此外, 通过这些基因表达谱的研究, 也为寻找新基因提供了有效的手段, 例如Xu等人便是利用SAGE技术分析了雄性激素调控基因的不同组织的表达谱后, 发现了PEMAE1新基因(Xu et al., 2000)。

此外, 通过基因敲除和转基因技术建立模式生物的方法, 是研究基因功能的主要手段之一。而RNAi(RNA interference)技术则是通过在RNA水平上的表达抑制来直接研究基因的具体功能(Fire et al., 1998)。进一步的研究表明, 在多种动物体内均存在dsRNA介导的RNAi现象, 这表明RNAi是生物界普遍存在的RNA水平上的基因调节方式。因此RNAi已成为大规模遗传筛选和基因功能鉴定等相关研究的最有效手段。

### 3.2 蛋白质功能及蛋白质的相互作用研究

由于大部分细胞的生命活动发生在蛋白质水平而非RNA水平, 因此, 即使清楚了基因的表达情况, 也难以阐明基因的实际功能, 因为基因在生物体上的功能最终由其编码的蛋白质来体现(赵剑华等, 2000)。为此, 以阐明生物体内的蛋白质表达模式和功能模式为目标的蛋白质组学(Proteomics), 便是功能基因组学研究的重要内容之一。蛋白质组(Proteome)最初是由Wasinger等人于1995年提出来的, 它表示基因编码的全部蛋白质(Wasinger et al., 1995)。蛋白质组是动态的, 具有时空性和动态性, 能够反映基因的表达时间、表达水平以及蛋白质翻译后的加工修饰和亚细胞分布等。

蛋白质组的分析主要包括蛋白质的分离与鉴定。双向凝胶电泳技术(O'Farrell, 1975; Gorg et al., 2000)和质谱技术(mass spectrometry, MS)是公认的蛋白质研究的标准方法(Gorg et al., 2004)。总的说来, 随着基因组计划的进一步深入, 基因组学研究的内容也在不断的拓展并衍生出新的交叉学科和新的技术, 如生物信息学(Bioinformatics)则是以生物大分子为研究对象、以计算机为工具, 运用数学和信息学的相关理论和方法来研究生命现象及海量的序列数据, 其主要研究内容包括序列分析和比对、基因识别、蛋白质结构预测、分子进化和数据库技术等。而比较基因组学(Comparative genomics)则是通过物种之间基因、基因家族及基因组区域的比较分析, 来揭示基因、基因家族的起源、功能及其在进化过程中的复杂

机制(Frazer et al., 2003; Nobrega & Pennacchio, 2004)等等。

## 第二节：基因组细菌人工染色体（BAC）文库研究概述

构建高质量和大插入片段的基因组 DNA 文库，是开展复杂真核生物大规模物理图谱、重要功能基因的定位、分离和大规模基因组测序等研究工作的基础，也是濒危物种基因资源离体保护的主要研究内容。BAC 文库是在早期的粘粒（cosmid）文库、噬菌体（fosmid）文库和 YAC 文库的基础上发展起来的。在早期基因组文库的构建过程中，研究者大多使用  $\lambda$ -噬菌体(Collins & Hohn, 1978)和柯斯质粒（Cosmid）作为克隆载体(De Jong et al., 1989)。由于这两种克隆载体的插入片段较小（15-45kb），所以，尽管其克隆效率很高，但插入片段过小和不稳定等特点（Yokobata et al.,1991；Kim et al.,1992），却致使其在大规模物理图谱构建、染色体步移和基因组测序等的研究中的应用受到限制。继后发明的酵母人工染色体（YAC）载体（Burke et al.,1987）能够克隆达 1Mb 的 DNA 片段，对于构建复杂基因组重叠群及物理图谱非常有用。然而，由于 YAC 系统存在着克隆效率低、嵌合体比率高、插入片段不稳定和提取纯化 YAC DNA 极为困难等诸多缺点（Green et al.,1991；Kouprina et al.,1994；Larinov et al.,1994；Cai et al.,1998；Woo et al.,1994）而使之不太适合于后续的染色体步移、物理图谱构建和基因组的测序研究。继后，Shizuya 等人在质粒 F 因子的基础上，构建了第一个细菌人工染色体(BAC)载体（Shizuya et al.,1992）。

以 BAC 为载体进行基因组文库构建时虽然其克隆插入片段(100-350 kb)较 YAC 要小得多，但其能够在宿主菌中稳定遗传，且无缺失、重组和嵌合等现象、同时其 BAC DNA 的提取纯化操作简单，所以，自该载体被发明以来就一直得到了广泛的应用（Cai et al.,1995；Suzuki et al.,2000；Thorsen et al.,2005）。

### 1. BAC 文库及其载体的发展

BAC 文库构建的基础是 *E.coli* 中的 F 因子（O'Connor et al.,1989）。研究发现 F 因子在 *E.coli* 中的复制是受到严格控制的，通常只有 1-2 个拷贝。此外，F 因子具有携带大插入片段的潜能，插入片段可达 100kb(Hosoda et al.,1990)，这使得构建大插入片段的基因组 BAC 文库成为可能。

O'Connor等人首次利用F因子构建载体pMBO131来克隆大片段的DNA分子(O'Connor et al.,1989),而Shizuya等人则在此基础上将SP6和T7启动子序列、 $\lambda$ 和P1噬菌体的CosN和LoxP位点引入pMBO131载体,构建了第一个BAC载体pBAC108L。该载体的插入片段达300kb以上,且传代100代后仍可稳定遗传,无重组和嵌合现象(Shizuya et al.,1992)。但是,该载体没有重组子选择标记,必须进行杂交验证,于是,Shizuya又将pGEM3Z载体上的*lacZ*基因插入pBAC108L载体的多克隆位点上,构建了第二代BAC载体——pBeloBAC11。该载体可通过*lacZ*的 $\alpha$ -互补所形成的蓝白菌落来筛选重组子,使得重组子的筛选更趋于直观、简便(Kim et al.,1996)。

由于pBeloBAC11载体所形成的非重组子的蓝色菌斑没有多拷贝载体那么明显,新的改进型载体pIndigoBAC在pBeloBAC11载体基础上使得非重组子的蓝白斑更加明显(Birren et al.,1997),同时,在多克隆位点引入*EcoRI*酶切位点,还能够插入经*EcoRI*酶切的DNA片段。Frijter等人通过定点突变方式将pBeloBAC11其中的一个*EcoRI*酶切位点去掉,构建了具一个*EcoRI*克隆位点的pECBAC1和pECSBAC4,这既丰富了多克隆位点,又方便了AFLP和RACE片段的克隆(Frijter et al.,1997)。此外,在基因组的部分酶切中,由于*EcoRI*和*EcoRI*甲基化酶的配合使用能够得到稳定和可重复的结果,这就大大提高了BAC文库的构建效率和质量。总之,这些DNA克隆载体在构建和发展的过程中,其结构、功能和克隆效率都逐渐得到了完善,并广泛应用于基因组文库和物理图谱构建、基因组大规模测序、定位克隆以及重要功能基因等的研究中。

## 2. 基因组 BAC 文库构建中的几个关键问题

尽管基因组 BAC 文库的构建工作看起来比较简单,但要成功构建出高质量大插入片段的基因组文库也并非易事,它需要特别就以下几个关键步骤予以高度重视。

### 2.1 高分子量 DNA (HMW DNA) 的制备

高分子量 DNA 的制备是构建大插入片段 BAC 文库并进而开展复杂基因组

研究的关键之一。与传统的 DNA 提取方法不同，要防止高分子量 DNA 在制备过程中免受剪切等物理损伤，通常需要将动物的组织或细胞等材料用低熔点琼脂糖凝胶（LMP）包埋，形成凝胶块（plugs/blocks）或凝胶球（Microbeads），以利于后续的细胞裂解和酶切等操作。

通常情况下，构建真核动物的基因组文库的材料，应来自单一的动物个体，这样就避免了由于序列的多态性而造成在 BAC Contig 构建和序列的组装过程中出现较大的麻烦（Osoegawa & de Jong, 2004）。与此同时，构建 BAC 文库的 DNA 的材料来源，也是非常重要的一个因素。本文在研究过程中发现，外周血是哺乳动物基因组 BAC 文库构建之大片段 DNA 的最佳材料来源，因为此材料 DNA 丰富、实验操作方便、白细胞的数量极易测算、酶切条件容易控制，并节省实验时间，最重要的是所构建的文库质量好。相比之下，虽然培养的动物细胞系也具有外周血的上述优点，但细胞系在传代的过程中，染色体容易产生重排或缺失，而造成某些 DNA 片段的丢失，使文库的质量在一定程度上受到影响，给后续利用文库开展相关的研究带来无穷的麻烦（Osoegawa et al.,2001）。当然，对于一些珍稀濒危或体型过小的物种来说，如果外周血难以获得或外周血的量太小，其组织细胞如脾脏、肾脏（Suzuki et al.,2000）或大脑（Vaiman et al.,1999）等材料，也不失为构建基因组 DNA 文库的大片段 DNA 的材料来源（Osoegawa et al.,2001）。

## 2.2 适宜 DNA 片段的获得

适宜 DNA 片段的获取是成功构建大插入片段 BAC 文库的最关键环节。尽管脉冲场电泳（PFGE）技术能够有效的分离不同分子量的 DNA 片段（Birren et al.,1993），但由于不同分子量的 DNA 片段存在着“共迁移”现象，导致在后续的 DNA 连接反应中，小分子量 DNA 的连接和转化效率大大超过大片段的 DNA 分子的结果（Leonardo & Sedivy, 1990），从而使 BAC 文库的平均插入片段极度偏小或过小。

## 2.3 DNA 片段的回收方式

一般来说, 加热融化低熔点琼脂糖凝胶, 再辅以琼脂糖凝胶消化酶(GELase)消化凝胶而将 DNA 片段释放出来 (McCormick et al.,1990; Larin et al.,1991; Frijters et al.,1997), 就能够回收构建基因组文库的 DNA。但是, 通过加热及后续的凝胶酶处理, 往往会破坏高分子量 DNA 的完整性 (Woon et al.,1998)。若添加一定浓度 (25-100mM) 的 NaCl 或多胺类 (如精胺和亚精胺等) 虽然能够减少加热过程对 DNA 的损伤, 但大片段 DNA 的回收效果却并不尽如人意 (McCormick et al.,1990; Riley et al.,1992; Amemiya et al.,1996; Riethman et al.,1997; Connelly et al.,1991)。

有研究 (Strong et al.,1997; Osoegawa et al.,1998) 表明, 电洗脱 (electro-elution) 回收 DNA 具有方法简单、回收效率高、使用常规的高熔点琼脂糖凝胶便易于分离大片断 DNA, 且普通的凝胶酶处理便能获得理想的实验效果等优点。

## 2.4 电转化参数的确定

所谓电转化方法, 是利用简短的高压脉冲, 在大肠杆菌细胞上形成暂时性的小孔隙 (Neuman et al.,1989; Chang et al.,1992), 从而使质粒 DNA 在电场力的作用下进入大肠杆菌细胞内 (Klenchin et al.,1991)。电转化的效率受到温度、电压梯度、电阻和电容等诸多因素的影响。同时, 宿主菌的质量、遗传背景、生长条件和电转化后的处理、DNA 样品的拓扑结构等, 都会直接影响电转化的效率及文库构建的质量。因此, 在电转化实验的实际操作过程中, 如果要获得较好的实验转化率, 电转化参数的设定, 是应随不同物种或不同大小的 DNA 转化片段而有所变化的。

## 2.5 嵌合体检测——FISH or not?

在基因组 BAC 文库构建结束后, 对于嵌合体的检测尤其重要, 因为这些嵌合体 BAC 克隆将对物理作图、测序、染色体步移等后续研究产生严重的影响。对于嵌合体的检测, 通常是以 BAC DNA 为材料, 利用荧光原位杂交 (FISH)

和缺刻平移等方法，将荧光探针与中期染色体进行杂交，然后检测克隆的嵌合体状态。由于早期的 YAC 克隆遗传不很稳定，容易发生染色体重排、缺失、基因组片段复制等嵌合现象，因此，运用 FISH 技术检测此类文库的嵌合状态，则成其为一种较优越的方法 (Green et al.,1991)。

然而，与 YAC 克隆不同，BAC 克隆由于在宿主菌内仅有 1-2 个拷贝，且遗传稳定，嵌合体比率低，因此采用 FISH 技术很难检测到 BAC 克隆实际的嵌合状态 (Ioannou et al.,1994; Woon et al.,1998; Korenberg et al.,1999)。所以，针对 BAC 克隆中的小片段重组或缺失的特点，参照 Osoegawa 等人的方法 (Osoegawa et al.,2000)，将一个 BAC 克隆的末端定位到另外一个相同的基因组区域或 BAC Contig 中予以检测，将会获得理想的检测结果。

### 3 基因组 BAC 文库的应用

近百年来，人口的急剧增长、环境污染、野生动物资源的过度利用和外来物种的入侵等诸多因子的胁迫，不断加剧了珍稀动物的濒危和灭绝的速度。因此，基因组 DNA 文库作为长久保存一个生物种的有效方法和手段，已成为了濒危物种离体保护工程的重要组成部分。

目前，一些学者已利用 BAC 克隆技术完成了许多物种的基因组 DNA 文库的构建工作 (Yuan et al., 2006; Luo et al., 2006; Suzuki et al., 2004; Zimmer & Gibbins, 1997; Osoegawa et al., 2000; Woon et al., 1998; Rogel-Gaillard et al., 2001; Vaiman et al., 1999; Suzuki et al., 2000; Kim et al., 1996; Li et al., 1999; Buitkamp et al., 2000)。但在濒危物种基因资源保存方面的应用，却极为鲜见 (Romanov et al., 2006; Luo et al., 2006; Thorsen et al.,2005; Xu et al., 2004; Qian et al., 2004)。因此，综合该领域长期的实践结果和目前相关技术发展的现状，本文认为 BAC 克隆技术是迄今能够完整和久远地保存濒危物种遗传基因资源的一种最为有效和可靠的方法。

### 第三节 主要组织相容性复合体 (MHC) 研究概述

主要组织相容性复合体 (Major Histocompatibility Complex, MHC) 是脊椎动物体内编码与免疫相关的蛋白质分子的一类具有高度多态性的多基因家族

(Snell, 1981; Benacerraf, 1981; The MHC sequencing consortium, 1999)。MHC 首先是在小鼠中发现的, 记为 H-2 (Gorer et al., 1948; Snell & Higgins, 1951), 随后在人 (Dausset, 1959)、鸡 (Schierman & Nordskog, 1961) 及较原始的脊椎动物软骨鱼类如鲨鱼中发现 (Bartl et al., 1997; Bartl & Weissman, 1994)。近年的研究表明, MHC 是仅存于脊椎动物体内的一类免疫适应性基因群 (Danchin et al., 2004; Laird et al., 2000)。MHC 基因编码的抗原蛋白, 能够在不同细胞的表面表达, 它不仅与 T 细胞的分化发育有关, 而且在免疫应答的启动和调节中发挥着重要作用。

MHC 是目前已知的多态性最丰富的基因区域, 拥有极大数量的等位基因, 赋予物种巨大潜力以适应复杂多变的变化环境 (Klein et al., 1986; the MHC sequencing consortium, 1999)。自 Watkins et al. (1988) 首次研究了 MHC 与种群的遗传多样性之间的关系以来, 有关 MHC 多态性的维持机制一直是研究的热点。在综合 MHC 的结构与功能如交配选择和亲缘识别等的基础上, 许多学者认为平衡选择是 MHC 保持其多态性的首要原因之一 (Hedrick & Thomson, 1983; Hedrick, 1999)。也正是因为 MHC 与动物的免疫功能和适应性进化潜力密切相关, 所以近年来它已被认为是濒危动物保护生物学和进化生物学研究的最佳候选标记之一 (Potts & Wakeland, 1990; Hedrick, 1994; Sommer, 2005)。

迄今为止, 已有人 (The MHC sequencing consortium, 1999)、鸡 (Kaufman et al., 1999)、小鼠 (Kumánovics et al., 2003)、猪 (Renard et al., 2006) 等动物的 MHC 基因组的精细图得以完成。其它重要的模式生物如爪蟾 (Ohta et al., 2006) 等动物的 MHC 基因组的精细图研究也取得了一定的进展。另外, 家猫 (Yuhki et al., 2003) 及狗 (Debenham et al., 2005) 的 II 类 MHC 基因组测序也全部完成。MHC 的上述研究积累, 为相关学者开展与野生动物特别是濒危野生动物免疫、进化和性选择机制有关的研究, 奠定了坚实的基础。

## 1. MHC 的组织分布、结构和功能

MHC 所编码的受体与细胞内的多肽结合后被递呈到各类细胞的表面, 通过与 CD4<sup>+</sup> 或 CD8<sup>+</sup>T 细胞之间的识别作用, 使机体产生免疫应答反应 (Ploegh &

Watts, 1998)。

依据基因的功能, MHC 被分为经典 I 类 (class Ia) 基因、非经典 I 类 (class Ib) 基因、经典 II 类 (class IIa) 基因、非经典 II 类 (class IIb) 基因和 III 类基因等 5 个部分。其中, 经典 I、II 类基因所编码的蛋白质经与抗原 (主要是细菌和病毒) 结合并递呈给 CD8<sup>+</sup>T 细胞 (Klein, 1986) 后随即被 NK 细胞识别, 清除细菌或病毒。非经典的 I 类和 II 类 MHC 基因除多态性较低外, 其结构上与经典的 I 类和 II 类 MHC 基因非常相似, 但没有抗原递呈功能, 只是扮演着经典 MHC 基因的分子伴侣角色 (Alfonso & Karlsson, 2000; Afanassief et al., 2001)。

### 1.1 MHC I 类分子的组织分布、结构和功能

经典的 MHC I 类基因由  $\alpha$  链和  $\beta$  链经非共价键连接成异源二聚体 (Cresswell & Turner, 1973)。 $\alpha$  链由 3 个胞外结构域 ( $\alpha 1$ 、 $\alpha 2$  和  $\alpha 3$ )、跨膜区和胞质区三部分组成 (Orr et al., 1979)。 $\alpha 1$ 、 $\alpha 2$  和  $\alpha 3$  分别约包含 90 个氨基酸, 其中  $\alpha 3$  结构域是 T 细胞表面与 CD8 分子相结合的部分。跨膜区由 25 个氨基酸残基组成, 而亲水性胞质区结构域则是由 30-40 个氨基酸残基所组成的。 $\alpha$  链的  $\alpha 1$  和  $\alpha 2$  结构域相互作用, 组成一个抗原结合区。I 类基因的  $\beta$  链为  $\beta_2$  微球蛋白 ( $\beta_2$ -m), 属可溶性蛋白, 其氨基酸序列高度保守, 主要作用是稳定 I 类分子的结构, 并使其能够有效的表达于细胞的表面。 $\alpha$  链是由具有多态性的经典 I 类 MHC 基因 (Ia) 所编码的, 如人的  $\alpha$  链由 HLA-A、HLA-B 和 HLA-C 三个位点所编码。在哺乳动物或其它大多数的脊椎动物中,  $\beta_2$ -m 是由 MHC 外的基因编码的, 它与 MHC Ia 链和 II 类 MHC 基因的进化关系较远, 但编码  $\beta_2$ -m 的基因位点不具有多态性 (Michaelson, 1983; Parham et al., 1988; Hughes & Yeager, 1998)。

MHC I 类分子在几乎所有有核细胞表面表达, 滋养层细胞 (Donaldson et al., 1992)、神经元细胞以及某些分化阶段的精细胞除外, 主要递呈内源性的多肽 (主要是病毒) 到 CD8<sup>+</sup>细胞毒性 T 淋巴细胞 (Bjorkman & Parham, 1990)。MHC I 类分子在各个物种间无论基因组大小还是基因的复杂度差异都很大, 这可能与 MHC 经历时常的扩张或收缩有关 (Vincek et al., 1987; Delarbre et al., 1992)。所有经典的 I 类分子 (Ia) 都具有高度的多态性、表达以及抗原递呈功能这一基本特征, 因此在结构上是非常保守的 (Kaufman et al., 1994)。而非经典的 I 类分子

(Ib) 尽管结构上与经典的 I 类分子相似,但是由于 RNA 可变剪切或终止密码子提前而造成 3'端缺失,因而几乎所有的非经典 I 类分子都是没有功能的,一些基因如 HLA-G 具有抗原递呈功能除外 (Shawar et al.,1994)。同时,非经典的 I 类基因多态性较低,在组织中的分布也比较有限。原因可能与这些基因所编码的 MHC 分子具有对某些特定保守的抗原相对应的结合基序 (binding motif) 有关,从而受到的选择压力小而失去遗传多样性 (Bouwer et al.,1994; Cheroutre et al.,1995)。

## 1.2 MHC II 类分子的组织分布、结构和功能

与 MHC I 类分子的广泛分布相比, MHC II 类分子的分布就较为局限,仅分布在一些特异的抗原递呈细胞如 B 细胞、树突细胞、胸腺上皮细胞或巨噬细胞表面,与外源性多态 (主要是细菌) 结合并递呈到 CD4+T 细胞 (Kappes & Strominger, 1988)。MHC II 类分子是由  $\alpha$  链和  $\beta$  链组成的异源二聚体,  $\alpha$  链和  $\beta$  链以非共价键相互连接。与 I 类分子不同的是, II 类分子的  $\alpha$  链和  $\beta$  链个由两个结构域 ( $\alpha 1$ 、 $\alpha 2$  和  $\beta 1$ 、 $\beta 2$ ) 组成。 $\alpha 2/\beta 2$  结构域能够与 T 细胞表面的 CD4 受体结合,而  $\alpha 1$  和  $\beta 1$  结构域相互结合形成 II 类分子的抗原结合区。II 类分子的另一个结构特点是  $\alpha/\beta$  异源二聚体可形成一个双二聚体 (dimer of dimers), 此结构可能与信号传导有关。

MHC II 类分子编码多态性的  $\alpha$  和  $\beta$  基因,以及编码可诱导  $\gamma$ -干扰素的蛋白补体成分 (LMP2 和 LMP7) 和与 MHC class Ia 抗原递呈相关的 TAP1、TAP2 等。非经典的 II 类分子延伸区大部分是编码非免疫相关的基因,而编码 TAP 结合蛋白 (TAPBP 或 Tapasin) 的基因对 class Ia 的生物合成至关重要;另外一个转录因子 RXRB 主要参与 I 类分子的表达 (Dey et al.,1992)。

一般来说, MHC II 类分子中的 DP、DO、DM、DQ 以及 DR 位点在整个哺乳动物中都是保守的,只是由于种属特异性复制造成这些位点中的  $\alpha$  和  $\beta$  基因的数量不同 (Kumánovics et al.,2003)。猫的 MHC (FLA) 缺乏整个 DQ 区域,只保留一些 DP 位点的假基因印迹,小鼠的 DP 位点也和猫一样,全部是假基因 (Yuhki et al.,2003; Kumánovics et al.,2003)。牛、山羊和绵羊具有反刍动物特有的 II 类基因——DYA 和 DIB (Hess et al.,1999; Ballingall et al.,2001)。除了 DM

位点外, MHC II 类分子在非哺乳动物中变异较大 (Kaufman et al.,1999), 系统发生分析研究表明, 非经典的 II 类 DM 的分化时间早于 II 类分子的进化时间 (Takahashi et al.,2000)。与哺乳动物不同, 鸟类的 II 类分子的连锁是一种种属特异性方式进化的, 鸡的 MHC II 类分子完全缺乏 DP、DQ 以及 DR 位点。

### 1.3 MHC III 类分子的结构和功能

在哺乳动物中, III 类 MHC 基因与机体对某些疾病的易感性有关, 如 I 型糖尿病、风湿性关节炎、红斑狼疮等 (Nishimura et al.,2003; Okomato et al.,2003; Gruen & Weissman, 2001)。III 类 MHC 基因是动物基因组中基因密度最大的区域之一 (Xie et al.,2003)。III 类 MHC 基因主要编码一些与免疫和非免疫相关的补体成分, 如 C4、TNF、HSP70 基因家族等, 主要以可溶性形式存在于血浆中 (Milner et al.,2001), 其基因产物不参与抗原递呈, 亦非膜结合分子。

由于 III 类 MHC 主要是编码一些血清补体成分、炎症相关以及非免疫功能的基因, 同时也仅表现出有限的多态性 (Falus et al., 1987)。所以相关学者在动物性选择机制及抗病力领域方面的研究很少涉及这类基因 (Hughes & Yeager, 1998; Sommer, 2005)。

## 2. MHC 的起源和进化

虽然在一些植物和无脊椎动物的体内, 也发现有组织相容性基因或 MHC 的直系同源基因 (Castro et al.,2004), 但通常认为 MHC 基因是伴随着有颌类脊椎动物的物种分化及适应性免疫系统的出现而产生的 (Danchin et al., 2004; Laird et al., 2000)。根据序列比较分析发现, 哺乳动物与无脊椎动物如黑腹果蝇和线虫等在某些 MHC 位点 (TUBB, TNXB, PBX2, NOTCH4, RXRB 及 RPS18) 存在着共线性 (synteny) 关系, 表明这些 MHC 位点可能先于脊椎动物适应性免疫系统基因出现的 (Trachtulec et al.,1997; The MHC sequencing consortium, 1999)。

MHC 在各个物种间是非常保守的, 包括 MHC 的整个功能以及多肽递呈位点的结构, 甚至于整个 MHC 的组织结构在各个物种间都是非常保守的 (Madden, 1995; Kulski et al.,2002), 其直接的证据, 源于人与小鼠 III 类 MHC 基因组织结构和功能的比较而获得的“III 类 MHC 基因在大小、外显子数量和基因间距非常

相似 (Xie et al.,2003) 的结果。迄今为止, 从所有有颌类脊椎动物的机体中都能分离到 I 类和 II 类 MHC 基因, 其中软骨鱼类如鲨鱼是最古老的具有 MHC 基因的脊椎动物 (Kasahara et al.,1995; Flajnik et al.,1999)。尽管人和大多数脊椎动物如小鼠、鸡、爪蟾及鲨鱼等的 I 类和 II 类 MHC 基因是紧密连锁的, 但人与所有已知硬骨鱼类 (鲑鱼、斑马鱼、鲤鱼及青鳉鱼等) 的直系同源基因, 却分布在不同的染色体位点上, 且其编码 I 类基因  $\alpha$  链和 II 类基因  $\alpha$  或  $\beta$  链的 MHC 基因并不是连锁的 (Matsuo et al.,2003; Kulski et al.,2002)。

MHC 基因具有众多源于基因复制所产生的旁系同源基因。由于只有脊椎动物才具有两次染色体的复制现象 (两次染色体复制事件发生在头索动物与有头类动物分化之后的约 766 百万年前, 以及有颌类动物辐射进化之前的约 528 百万年前) (Abi-Rached et al.,2002; Vienne et al.,2003), 且脊椎动物才具有适应性免疫系统, 所以, 在 MHC 出现之前, 一定存在其原始的祖先基因 (proto-MHC)。

关于 MHC 及其 3 个主要的旁系同源 (Paralogous) 区域的起源和进化主要有两种观点: 第一种观点认为它们来自于相同的祖先, 是由于模块复制 (block duplication) 或染色体复制而产生的 (Kasahara et al.,1996; Abi-Rached et al.,1999); 第二种观点则认为这些独立复制的基因, 是在选择压力的作用下聚集 (功能性聚集) 在一起的 (Hughes, 1998)。此外, 也有学者利用基因转换假说来解释 MHC 的序列进化模式 (Klein, 1984; Martinsohn et al.,1999), 即基因转换主要发生在 II 类 MHC 之  $\beta$  基因的第 2 外显子中, 它是 II 类 MHC 基因 (DQB、DRB 及 DPB) 中的一种重要的进化方式。但是, 基因转换假说用于解释 II 类 MHC 基因中单态的 DOB 位点、 $\alpha$  基因及 I 类 MHC 基因的序列进化模式, 却存在着极大的局限性。

综述文献: MHC 可能是通过 “生-死” (Birth-birth- Death) 模式——以大量重复基因的复制或丢失为特征进化的 (Nei et al.,1997)。因此, 模块复制可能是 MHC 产生众多旁系同源基因的主要进化模式, 而功能性聚集可能只是次要的因素 (Flajnik et al.,2001)。

### 3. MHC 多态性及其维持机制

MHC 是脊椎动物中一类具有免疫防御功能的基因家族, 由紧密连锁和高度多态的基因座位组成, 是基因组中多态性最丰富的区域之一, 如人 HLA 的 I 类

和 II 类基因的某些座位具有 200 个以上的等位基因 (The MHC sequencing consortium, 1999)。脊椎动物免疫系统中的 I 类和 II 类 MHC 基因的高度多态性, 是机体在其漫长的进化历史过程中与内源性和外源性抗原如病毒、细菌和寄生虫等长期斗争的结果 (Potts & Slev, 1995)。由于 MHC 编码的蛋白质广泛分布于各类细胞的表面, 并参与各种免疫识别、免疫调节和免疫应答等复杂功能, 因此, MHC 的高度多态性, 赋予了动物应对复杂的内外界环境的进化潜力。通常认为, MHC 的多态性越高, 机体抵御病原微生物和寄生虫感染的能力就愈强 (Germain & Margulies, 1993; Potts & Slev, 1995; Kaufman et al., 1995)。

MHC 是多态性极强的功能基因家族, 其多态性的维持机制存在着两种解释, 其一是平衡选择压力 (Hedrick & Thomson, 1983)。平衡选择体现在两个水平: 等位基因水平 (即等位基因频率均衡化) 和氨基酸水平 (有义突变频率均衡化)。由于 MHC 是共显性基因, 任一座位的两个等位基因均能有效地编码抗原呈递蛋白, 所以, MHC 位点上的杂合度越高, 机体递呈外源性多肽的能力和对疾病的抵抗力就愈强 (Doherty & Zinkernagel, 1975; Hughes & Nei, 1988)。这种杂合子优势将稀有等位基因的频率不断提升, 从而使各等位基因的频率趋于均衡化 (Thursz et al., 1997; Carrington et al., 1997; Penn et al., 2002; Arkush et al., 2002)。

外界病原与动物机体一直处于竞争之中, 一旦动物机体已经能够有效识别已有病原体时, 病菌将通过自身变异产生新的变异体, 以期逃避动物机体的识别。因此, MHC 基因的抗原结合区的任何一个有义突变, 都可能具备着识别新病原的能力, 且有义突变的种类越多, 识别抗原多肽的能力就越强。在外界病原与动物机体的竞争中, 有义突变得保留和累积, 因此, 在氨基酸水平的杂合优势机制下, 不同的有义突变在物种的进化过程中得以均衡化, 以确保整个物种的抗病潜力。研究表明, 在中性理论原则下, 同义突变 ( $ds$ ) 率通常是远远大于非同义突变 ( $dn$ ) 率的, 即  $ds \gg dn$ 。但是, Hughes 和 Nei 等人 (1988; 1989) 在研究 MHC 的多态性结合位点时, 却发现了  $ds \ll dn$ , 这一结果表明了氨基酸水平的优胜劣汰模式 (即有义突变均衡化) 维持着 MHC 多态性结合位点的遗传变异。上述证据表明, MHC 的多样性是动物机体与外界病原 (parasite or pathogen-driven) 长期相互作用的结果 (Parham & Ohta, 1996; Apanius et al., 1997; Thursz et al., 1997; Edwards & Hedrick, 1998; Hedrick & Kim, 1999)。

维持 MHC 多态性的第二种解释是性选择机制 (Sexual selection), 包括 MHC 依赖的选择交配 (MHC-dependent mate preferences) 和近交避免假说 (Inbreeding avoidance hypothesis)。选择交配理论认为, 通过对亲本的选择, 有利于产生高适合度的杂合子后代, 从而增加对疾病的抵抗力 (Penn, 2002)。而近交避免机制 (Potts & Wakeland, 1990) 则是亲本之间通过对特定 MHC 所决定的动物体味的识别, 来避免亲缘个体之间的交配繁殖 (亲本间的亲缘关系越近, 彼此的体味就愈接近), 从而增加繁殖后代对疾病的免疫能力 (Apanius et al., 1997)。该机制是动物在其漫长的演化过程中于环境适应的需要而形成的 (Penn, 2002)。

值得注意的是, 在某些多态性的 II 类 MHC 基因位点上, 存在着跨物种进化 (Trans-species evolution) 现象。这些等位基因即使是经过了物种的分化事件, 但也能够在相关的物种中长期存在 (Klein, 1987; Figueroa et al., 1988)。跨物种进化现象在很多物种如鱼类 (Ottova et al., 2005)、食肉类 (Hedrick et al., 2000; Seddon & Ellegren, 2002)、有蹄类 (Hedrick et al., 2000) 及灵长类 (Otting et al., 2002) 等动物中都普遍存在。因此, 跨物种进化也是 MHC 多态性起源及其维持的重要因素之一。

#### 4. MHC 在保护生物学中的应用

MHC 是由紧密连锁的、编码细胞表面糖蛋白并参与免疫调节和免疫应答的相关基因所组成的多基因家族。在大多数物种中, MHC 基因所编码的产物都表现出高度的抗原多态性。MHC 的变异水平, 反映了物种应对内外复杂环境及识别外来病原微生物的适应能力或生存潜力, 在一定程度上也反映了物种潜在的病原抵抗力 (Hughes & Nei, 1989; Takahata & Nei, 1990; Yuhki & O'Brien, 1990)。在一定程度上, MHC 的变异水平也可反映整个物种基因组的变异水平 (Hedrick, 2002)。通常地, 小种群和濒危物种 MHC 的多样性都较低, 因此, 高疾病致死率是对其致危甚至导致物种灭绝的主要因素之一 (Laurenson et al., 1998; Murray et al., 1999)。

目前, MHC 作为一种新的功能基因标记, 已被广泛的应用于保护生物学的各个领域。与中性标记相比, MHC 可反映物种进化及物种对环境的适应性进化过程。因此, MHC 的遗传变异分析对于探索物种的遗传多样性、种群遗传结构

和种群动态、近缘种之间的进化关系与进化历史 (Klein et al.,1997; Hedrick & Parker, 1998; Langefors et al.,1998), 以及濒危物种遗传管理及保护策略的制定等等, 都具有极其重要的意义。

## 第二章 论文的立论依据、研究内容和目的

大熊猫 (*Ailuropoda melanoleuca*) 是我国特有的珍稀濒危野生动物。长期以来, 由于受到栖息地的破坏、人类的猎杀, 以及竹子开花所造成的食物短缺等主要因素的影响, 野生种群现仅分布于四川、陕西和甘肃三省毗邻的秦岭、岷山、邛崃、大相岭、小相岭和凉山等 6 个相互隔离的山系, 并分化为两个独立的亚种——秦岭亚种和指名亚种 (四川亚种)。

截至 2002 年, 我国实施的大熊猫保护策略, 是野外栖息地保护与人工圈养繁育保护并重的策略。然而, 鉴于大熊猫物种保护的 task 仍然十分艰巨, 同时随着国家财政进一步加大了对社会公益型项目的投入, 我国政府于 2003 年启动了大熊猫的离体保护工程。该工程的主要内容和目的之一, 就是以细菌人工染色体 (BAC) 为载体, 建立大熊猫的高容量基因组 DNA 文库, 抢救和永久性的保护大熊猫的遗传基因资源, 并为相关的基础科学研究提供基因材料的共享平台。同时在 BAC 基因组文库的基础上, 利用鸟枪法和步移法等技术, 构建大熊猫 MHC II 类基因组的物理图谱, 以验证该文库的有效性, 并为后续 MHC 基因组计划奠定基础。



## BAC Cloning Kit)

Fig.2-1 pCC1BAC vector map & complete restriction information (cited from Manual of CopyControl BAC Cloning Kit)

文库构建所用的大肠杆菌电转化感受态细胞 Transformax EPI300 *E.coli* 购自 Epicentre 公司，该感受态细胞可获得  $1 \times 10^{10}$  cfu/ $\mu$ g DNA 的电转化效率。

### 1.3 主要仪器设备

CHEF DR-III 脉冲场电泳系统 (Bio-Rad); Gene Pulser II 电转仪 (Bio-Rad); 5417R 高速冷冻离心机 (Eppendorf); 5810R 高速冷冻离心机 (Eppendorf); BS 200S 电子天平 (Sartorius); PTC-200 PCR 扩增仪 (MJ Research 公司); PTC-220 PCR 扩增仪 (MJ Research 公司); Ultra Freeze 超低温冰箱 (Heto); Power PAC 300 电泳仪 (Bio-Rad); Sub-cell GT 电泳仪 (Bio-Rad); Gel Doc2000 凝胶成像系统 (Bio-Rad); SBD-50 Bio 恒温水浴摇床 (Heto); 恒温混匀器 Thermomixer comfort 5355 (Eppendorf); Venticell 111R 烘箱 (MMM); Incucell 111R 培养箱 (MMM); 3310 PH Meter (Jenway); SS-325 灭菌柜 (Tomy); DNA mini 真空干燥系统 (Heto); SmartSpec 3000 紫外分光光度计 (Bio-Rad); Purelab uhq 去离子纯水仪 (ELGA); Vortex-2 genie 旋涡混匀器; 12 道电子移液器 (大龙); 凝胶包埋用 disposable DNA plug molds (Bio-Rad), 等等。

### 1.4 主要试剂和耗材

Copycontrol BAC cloning kit、Transformax EPI300 electrocompetent *E.coli*、Plasmid-safe ATP-dependant DNase、GELase 和 BAC-Tracker supercoiled DNA ladder 均购自 Epicentre 公司。Large-construct kit 购自 Qiagen 公司。MidRange II PFG marker (New Engl& Biolabs)、Lambda Ladder (Bio-Rad)、Lambda mix marker-19 (Fermentas)、 $\lambda$ DNA-PFGE marker (Amersham)、Dialysis membrane (Spectrum)、0.025mm filter (Millipore)、Chloramphenicol、Spermine、Spermidine (Sigma)、96 孔培养板 (Costar)、Ex-Taq DNA polymerase 和 *Hind*III 购自宝生物工程 (大连) 有限公司。PEG8000、IPTG、X-gal、*Not*I 和其它常规试剂购自上海生工生物工程公司。日常质粒快速抽提试剂盒 (Axygen)、脉冲场用琼

脂糖和低熔点琼脂糖 (LMP) 购自 Bio-Rad 公司。PCR 引物由上海生工生物工程有限公司合成。BAC 末端测序由华大基因 (杭州) 公司完成。

## 2. 方法和步骤

大熊猫基因组 BAC 文库的构建方法和步骤, 主要参考 Osoegawa 等人(2001) 和 Osoegawa & de Jong 等人 (2003) 的方法, 但实验中对部分方法和步骤进行了适当的修改 (见如下描述)。

### 2.1 高分子量 DNA (HWM DNA) 的制备

#### 2.1.1 血液样品处理及红细胞裂解

- (1) 将 20ml -80℃ 冷冻保存的血液样品解冻至室温, 分装于 2 个 15ml 无菌离心管; 加入 5ml 冰冷的 PBS 缓冲液, 颠倒混匀;
- (2) 冷冻离心机 4℃、1876g 离心 5 分钟; 移液器小心吸去约 3-4ml 上清液;
- (3) 重新加入 3-4ml PBS 缓冲液, 按步骤 (2) 离心; 反复用冷 PBS 缓冲液重复上一步操作 10 次以上, 直至溶液基本清亮为止;
- (4) 尽量去除上清液, 将沉淀轻轻混匀; 加入 10ml 1×RBC ( ) 裂解液后置于漩涡混匀器中室温保持约 20-30 分钟;
- (5) 注意观察细胞悬液的裂解程度, 当悬液颜色突然加深时, 将离心管取出, 置 4℃ 冷冻离心机、208g 离心 10 分钟;
- (6) 将离心管颠倒去除上清液; 2ml 冰冷的 PBS 清洗沉淀 1-2 次; 微量移液器小心去除上清液 (尽量不要搅动沉淀);
- (7) 加入 5ml 冷 PBS 后置 4℃ 冷冻离心机、208g 离心 10 分钟; 重复前述步骤直至细胞悬液基本变清亮为止;
- (8) 加入适量 PBS 将细胞浓度调至约  $5 \times 10^7$  个/ml 后置冰上备用。

#### 2.1.2 细胞包埋处理

- (1) 超声波清洗器处理凝胶 plug molds 模具, 最后用无菌去离子水冲洗后晾干, 并置紫外灯下照射 30min 备用;
- (2) 取脉冲场专用低熔点琼脂糖 (LMP) 0.2g, 加入 10ml PBS 缓冲液并微波炉熔解后置于 50℃ 水浴中保温备用;

- (3) 上述调整好浓度的细胞悬液与低熔点琼脂糖按 1: 1 比例混匀后置 50℃ 水浴中保温;
- (4) 移液器将凝胶-细胞混合液快速加入 plug molds 模具中, 并置 4℃ 冰箱中 20-30min 使凝胶 plug 凝固。

### 2.1.3 凝胶包埋块 (plugs) 的处理

- (1) 将 DNA plugs 从模具中挤出, 并置于 50ml 离心管中; 加入 45ml 蛋白酶 K 裂解液;
- (2) 将离心管放入 50℃ 水浴锅孵育 48h, 且每隔 2-3h 混匀一次; 24h 更换一次新鲜蛋白酶 K 裂解液;
- (3) 裂解完毕, 小心倒去裂解液, 加入无菌去离子水反复清洗 DNA plugs 数次;
- (4) 加入 50ml TE<sub>50</sub> 溶液置 4℃ 恒温振荡仪中轻轻震荡 24h, 但每隔 8h 需更换一次新鲜配置的 TE<sub>50</sub> 溶液;
- (5) 弃去 TE<sub>50</sub> 溶液, 加入 50ml 0.1mM PMSF 溶液以去除蛋白酶 K 活性; 每隔 2h 更换一次新鲜的 0.1mM PMSF 溶液后于 4℃ 恒温振荡仪中震荡 4h;
- (6) 弃去 PMSF 溶液, 加入 50ml TE<sub>50</sub> 溶液于 4℃ 恒温振荡仪中震荡清洗 24h, 但每 12h 更换一次震荡液;
- (7) 弃去 TE<sub>50</sub> 溶液, 加入 0.5M EDTA 于 4℃ 保存或直接进行下一步操作。

## 2.2 预电泳

- (1) 从 0.5M EDTA 溶液中取出适量的 DNA plugs, 加入 50ml 4℃ 预冷的无菌 TE 或 0.5×TBE 溶液, 浸泡 2-3h;
- (2) 制备 1% 的脉冲场电泳凝胶, 将 4-5 个梳孔用无菌透明胶带封住, 形成一个较大的凝胶槽; 启动冷却循环装置, 将电泳缓冲液预冷至 12℃;
- (3) 将 DNA plugs 小心加入凝胶槽中 (分子量标记为 Lambda Ladder);
- (4) 调节脉冲场电泳参数至: 4V/cm、脉冲时间 5s、12℃ 电泳 8-10h;
- (5) 将 DNA plugs 从凝胶槽中取出, 加入 50ml 无菌 TE 溶液透析 2-3h, 但需每 30 分钟更换一次 TE 溶液或直接将 DNA plugs 放入 0.5M EDTA 中保存;
- (6) 将电泳完的凝胶用 EBtr 染色 20-30min, 清水漂洗 15-20min; 凝胶成像系统检测制备的 DNA plugs 的质量。

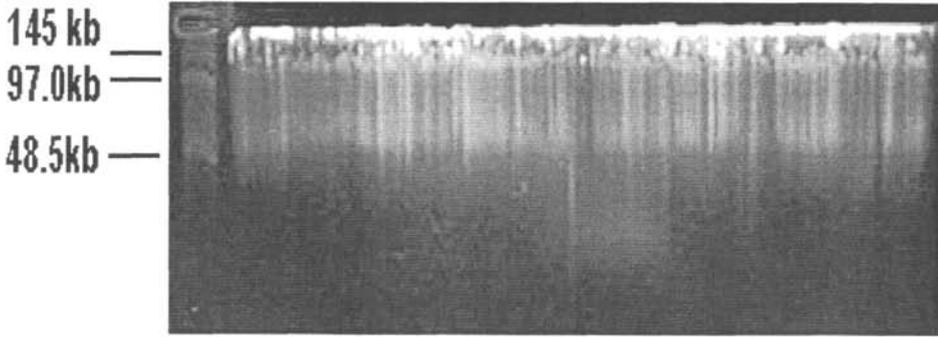


图 2-2: 预电泳图: 左侧为分子量标记 Lambda Ladder。

Fig.2-2: Pre-electrophoresis, size standard is Lambda Ladder.

### 2.3 高分子量 DNA 的不完全酶切

- (1) 将预电泳后的 DNA plugs 切割分成 4 小块, 分别加入 400 $\mu$ l 1 $\times$ M 酶切缓冲液, 置冰上透析 30min;
- (2) 更换新鲜配制的 1 $\times$ M 酶切缓冲液, 每 plug 加入 400 $\mu$ l 缓冲液。为探索最佳的反映酶量, 分别加入 0.3 $\mu$ l、0.5 $\mu$ l、0.7 $\mu$ l、0.9 $\mu$ l、1.1 $\mu$ l、1.3 $\mu$ l、1.5 $\mu$ l *Hind*III (15U/ $\mu$ l), 置冰上保持 20 分钟, 使 *Hind*III 进入 DNA plug 内;
- (3) 将酶切反应管立即放进 37 $^{\circ}$ C 水浴锅内, 酶切反应 20min;
- (4) 酶切反应结束后, 立即加入 1/5 体积的 0.5M EDTA 终止酶切反应;
- (5) 将酶切反应后的 DNA plugs 加入 1% 脉冲场琼脂糖凝胶, 于 6V/cm、脉冲时间 1-40s、12 $^{\circ}$ C 电泳 16h (分子量标记为 Lambda Ladder);
- (6) 将电泳后凝胶用 EBtr 染色 30min 并清水漂洗 15-20min 后, 用凝胶成像系统检测电泳条带的分布情况, 从而判断最佳的酶切反应用量;
- (7) 选定最佳酶切用量后, 重复步骤 1-5 进行 DNA 片段的选择回收。

### 2.4 适宜 DNA 片段的选择

- (1) 选定 5-6 块 DNA plugs 进行酶切反应。为利于充分酶切, 每块 plug 需切割成 6-8 个小块;
- (2) 在制备 1% 脉冲场琼脂糖凝胶时, 将 4-5 个梳孔封住形成一排凝胶槽。在将酶切后的 DNA plugs 塞进凝胶槽后, 上面需用 1% 琼脂糖凝胶封住, 并在一侧加入分子量标记 MidRange II PFG marker 或  $\lambda$ DNA-PFGE marker;

- (3) 选择 DNA 片段需分 3 个连续的步骤进行：第一步，将凝胶点样孔一端靠近正极（即反转 180 度）放入电泳装置内，5.1V/cm、脉冲时间 15s、12℃ 电泳 6h；第二步，更换电泳缓冲液，将凝胶反转回正常电泳位置，以同样的脉冲场电泳条件电泳 6h，使 DNA 回复到原始点样孔位置；第三步，重新在另外一侧加入分子量标记，将脉冲场电泳条件调整为 6V/cm、脉冲时间 1-40s、12℃ 电泳 16h）；
- (4) 将分子量标记连同 1-2mm 含 DNA 条带的凝胶切割下来，EBTr 染色 30min，凝胶成像系统检测 DNA 片段的分布范围；
- (5) 用刻度尺标记 120-350kb 范围的 DNA 片段，在未染色的琼脂糖凝胶中将 120-350kb 范围的凝胶切割下来，水平分成 3 份，每块大约 0.5cm 左右；
- (6) 将切割下来的 DNA 凝胶块分别置于 50ml 无菌的 0.5×TBE 缓冲液中透析 1-2h 或放入 0.5M EDTA 中保存备用。

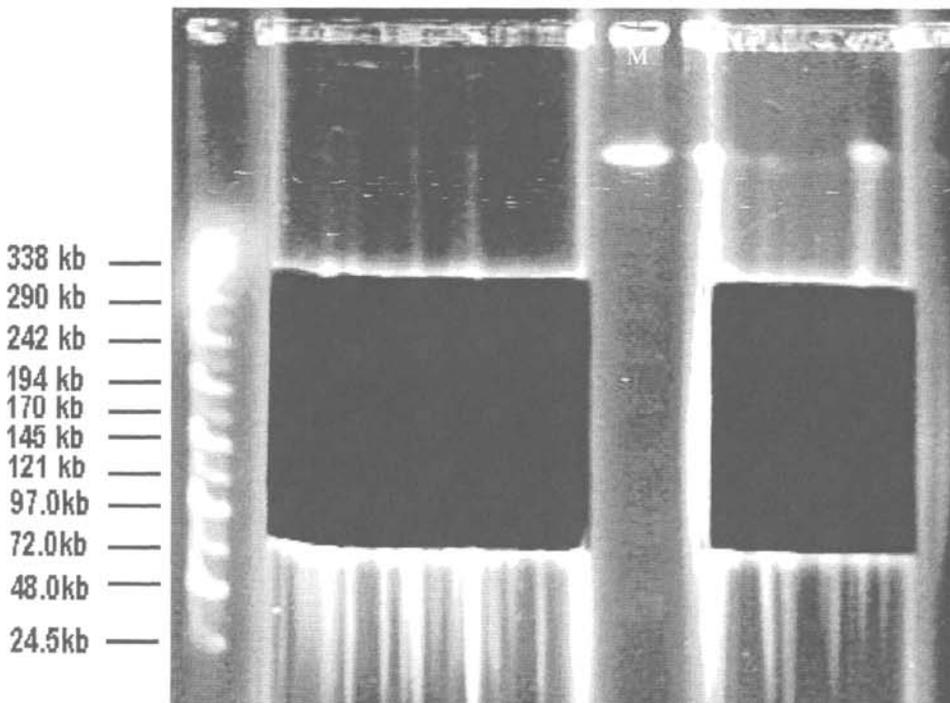


图 2-3: 二次片段选择回收 DNA 示意图，其中分子量标记为 MidRange II PFG marker，中间的 M 为  $\lambda$ DNA-PFGE marker 标记。

Fig.2-3: Slice & recover of desired DNA by second round of size fraction. The size standard is MidRange II PFG marker. M indicate for  $\lambda$ DNA-PFGE marker.

## 2.5 DNA 片段的回收

### 2.5.1 凝胶酶消化回收 DNA 片段

- (1) 待制备的 1%低熔点琼脂糖凝胶凝固后，用无菌刀片将点样孔切成大小约 0.5cm 厚的孔槽，将上一步切割回收的 DNA 凝胶块置入孔槽中，上面及边缘用 1%琼脂糖凝胶封住；
- (2) 设定脉冲场电泳条件为：4V/cm、脉冲时间 5s、12℃电泳 10-12h；
- (3) 从一侧将 1-2mm 含 DNA 片段的凝胶切割下来，EBTr 染色 30min；凝胶成像系统检测 DNA 条带的电泳位置，并用刻度尺做好标记；
- (4) 按刻度尺所标记的 DNA 条带范围将未染色的凝胶切割下来，置 200ml 无菌 TE 或去离子水中透析 2-3h，每 30min 更换一次新鲜的 TE 或去离子水；
- (5) 将透析后的含 DNA 片段的低熔点琼脂糖凝胶放入 1.5ml 无菌离心管内，电子天平称其重量，按凝胶酶（GELase）的说明书加入酶切缓冲液，置冰上透析 1h；
- (6) 酶切缓冲液完全吸出后，将离心管置入 70℃水浴锅内融化凝胶，并轻轻颠倒离心管直至凝胶完全融化；
- (7) 取出离心管置入 45℃水浴锅内孵育 2-3min，同时按说明书加入适量的 GELase（消化时间决定于使用的酶量）；
- (8) 取 5 $\mu$ l 消化液于 1%普通琼脂糖凝胶中电泳，并根据已知浓度的分子量标记判断消化液中的 DNA 浓度；
- (9) 凝胶酶消化后的 DNA 片段可直接用于下一步的连接反应，或加样到含 30%PEG 溶液的 0.025mm 滤膜上进行浓缩。

### 2.5.2 电洗脱回收 DNA 片段

电洗脱回收采用两种方法进行：一是直接用上述 2.3 中步骤 5 切割下来的 DNA 片段；二是进一步浓缩 DNA 条带，具体方法如下：

- (1) 将含有 DNA 片段的琼脂糖凝胶块用无菌 0.5 $\times$ TBE 清洗数次。
- (2) 将透析袋剪成合适长度，置于无菌去离子水中漂洗 30min，取出透析袋加入无菌 0.5 $\times$ TBE 冲洗数次；
- (3) 将透析袋一端用夹子或尼龙线封住，加入适量的 0.5 $\times$ TBE，然后将含 DNA 片段的凝胶置入透析袋内，尽量挤去多余的 TBE 溶液，然后将透析袋封

住；

- (4) 将透析袋放入脉冲场电泳装置内的 0.5×TBE 缓冲液中，6V/cm、脉冲时间 30s、12℃电泳 3h；
- (5) 透析袋水平旋转 180 度，以相同的脉冲电泳条件电泳 1min 后，将透析袋取出置于预冷的无菌 TE 溶液中透析 2h，其间每 30min 更换一次 TE 溶液；
- (6) 将透析袋内的 DNA 溶液吸出，取 5 $\mu$ l 溶液进行电泳以确定回收 DNA 的浓度。若浓度过低，可加样到含 30%PEG 溶液的 0.025mm 滤膜上进行浓缩即可。

## 2.6 DNA 连接

- (1) 用去头的移液枪头取 43 $\mu$ l 回收的 DNA 溶液于 1.5ml 无菌离心管内，加入 0.5 $\mu$ L pCC1BAC 载体 (15U/ $\mu$ l) (DNA 片段与载体的摩尔比约为 10: 1)，轻轻混匀，置 55℃水浴锅内孵育 10min；
- (2) 取出离心管约 15min，使之冷却至室温；
- (3) 加入 5 $\mu$ l 10×Fast-link 连接缓冲液、0.5 $\mu$ l 100mM ATP 和 1 $\mu$ l Fast-link DNA Ligase，轻轻混匀；
- (4) 将连接反应的离心管置于 16℃恒温混匀器中连接 4h (必要时也可延长连接时间或过夜；离心管置于 65℃水浴锅内孵育 15min 终止连接反应；
- (5) 将连接溶液加入制备好的 agarose cone 内置冰上脱盐 30min；
- (6) 吸出去盐连接产物置于含 30%PEG 和 0.5mM Spermidine 的 TE 溶液中的 0.025mm 滤膜上浓缩至 10 $\mu$ l；吸出浓缩产物，4℃保存或直接用于转化。

## 2.7 确定适宜的电转化参数

### 2.7.1 电转化

- (1) 将 100 $\mu$ l 电转化大肠杆菌感受态细胞 (Electrocompetent EPI300 *E.coli*) 于冰上解冻后分装离心管 (20 $\mu$ l/离心管)；
- (2) 每只离心管内加入 2 $\mu$ l 去盐连接产物，轻轻吹打混匀；
- (3) 混合物加入预冷的 0.1cm 电击杯内 (防止产生气泡)，将杯外壁用吸水纸擦干，置于电转仪中准备转化；

- (4) 调节电转化参数为 0.9、1.1、1.3、1.5、1.7kV、100 $\Omega$  和 25 $\mu$ F 进行电击；
- (5) 转化结束后在各离心管内立即加入 1ml 37 $^{\circ}$ C 预热的 SOC 培养基，混匀后吸入新的无菌离心管内；
- (6) 37 $^{\circ}$ C、200rpm 震荡培养 60min。

#### 2.7.2 涂板及蓝白斑筛选

- (1) 制备加入了 12.5 $\mu$ g/ml 氯霉素的 LB+Agar 平板培养基；
- (2) 加入 100 $\mu$ l IPTG (25mg/ml) 和 20 $\mu$ l X-gal (50mg/ml) 后涂布于平板上；
- (3) 分别吸取 100 $\mu$ l 的各电转化参数的电转化产物，均匀涂布于 LB 平板培养基上，37 $^{\circ}$ C 培养箱过夜；
- (4) 检查蓝白斑比例，并就每个平板随机挑取 5 个白色菌落，加入 1ml LB+氯霉素培养液；37 $^{\circ}$ C、280rpm 震荡培养过夜。

#### 2.7.3 检测插入片段大小

- (1) 培养过夜后的菌落于 8000rpm 离心 2min，完全弃去上清液；
- (2) 加入 25 $\mu$ l EpiLysis 溶液，旋涡震荡仪剧烈震荡至沉淀溶解悬浮；
- (3) 加入 10 $\mu$ l EpiBlue 溶液，旋涡震荡仪剧烈震荡后于 12000rpm 离心 2min；
- (4) 0.8-1%琼脂糖凝胶置于含 1%TAE 电泳缓冲液的 Sub-cell GT 电泳仪中；加入 20 $\mu$ l 裂解后的上清液，再加入 15 $\mu$ l Supercoiled DNA Ladder 作为分子量标记于 3.5V/cm 电泳 3h；
- (5) EBtr 染色 30min 后用凝胶成像系统检测不同电转化参数的插入片段分布；
- (6) 根据插入片段大小和转化效率判断最适宜的电转化参数。

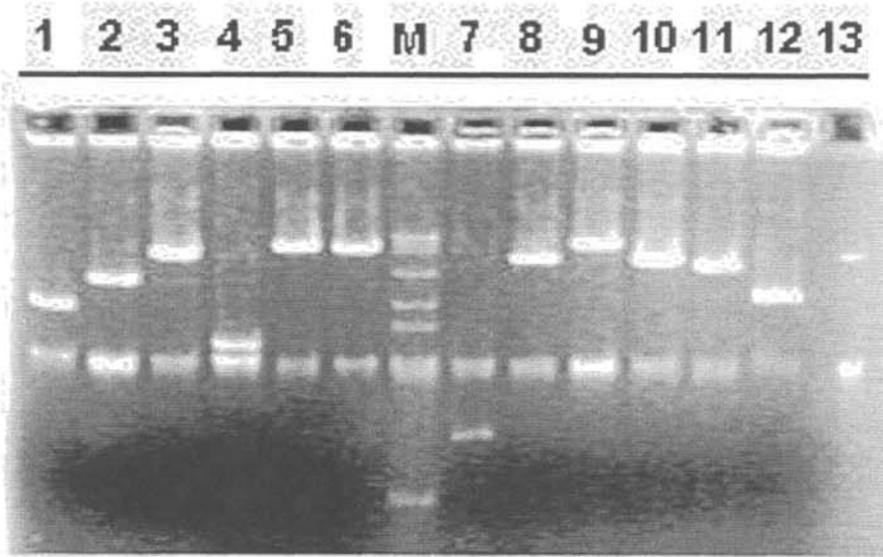


图 2-4: 1%普通琼脂糖凝胶电泳检测 BAC 克隆插入片段大小。其中 M 为 Supercoiled DNA Ladder; 从上至下依次为 165、120、95、55、38、28、8kb。

Fig.2-4: Size analysis of BAC clones by 1% agarose electrophoresis, M is indicated for supercoiled DNA ladder contains supercoiled DNA of 8, 28, 38, 55, 95, 120, 160 kb.

## 2.8 BAC 克隆挑取及保存

- (1) 确定合适的电转化参数后, 重复步骤 2.6.1 进行大规模转化;
- (2) 在 96 孔培养板内加入 200 $\mu$ l LB 冷冻培养液, 用无菌牙签挑取平板上的白色菌落于 96 孔板中;
- (3) 将挑好的 96 孔板置于 37 $^{\circ}$ C 培养箱中静止培养 18-24h;
- (4) 编号记录 96 孔培养板后于 -80 $^{\circ}$ C 超低温冰箱中冷冻保存。

## 2.9 大熊猫 BAC 文库的检测和鉴定

### 2.9.1 BAC 克隆的稳定性检测

- (1) 随机挑取 5 个 BAC 克隆, 加入 1ml 含 12.5 $\mu$ g/ml 氯霉素的 LB 培养液 37 $^{\circ}$ C、280 rpm 震荡培养过夜 (记为 0);
- (2) 分别取出 300 $\mu$ l 菌液, 加入 700 $\mu$ l LB 培养液、1 $\mu$ l Copycontrol Induction 高拷贝质粒诱导液, 37 $^{\circ}$ C、280 rpm 震荡培养 2h (记为 D<sub>0</sub>);
- (3) 将上述诱导成高拷贝的 BAC 菌液和未诱导的 BAC 菌液各取 500 $\mu$ l, 加入

5ml LB 培养液，37℃、280 rpm 震荡培养过夜（记为 D<sub>1</sub>）；

- (4) 从 5ml 菌液中各取 500 $\mu$ l，加入 5ml LB 培养液 37℃、280 rpm 震荡培养过夜（记为 D<sub>2</sub>），并重复 5 天；
- (5) 将 D<sub>0</sub>、D<sub>1</sub>、D<sub>2</sub>、D<sub>3</sub>、D<sub>4</sub> 和 D<sub>5</sub> 培养的 BAC 克隆用常规碱裂解法提取质粒 DNA；
- (6) 制备 0.8%琼脂糖凝胶，取 20 $\mu$ l 质粒 DNA 按 3.5V/cm 电泳 3h（分子量标记为 15 $\mu$ l Supercoiled DNA Ladder）；
- (7) EBtr 染色 30min，凝胶成像系统检测电泳后质粒 DNA 的条带情况。

### 2.9.2 平均插入片段大小的检测

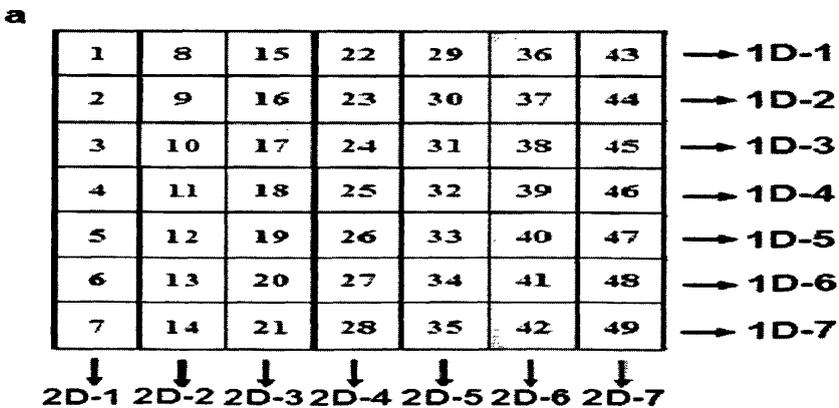
- (1) 从随机挑取的 174 个 BAC 克隆中各吸出 5 $\mu$ l BAC 菌液，加入 1ml LB 培养液（12.5 $\mu$ g/ml 氯霉素）于 96 孔无菌培养板中；37℃ 静止培养过夜；
- (2) 吸出 600 $\mu$ l 菌液弃去，另加入 600 $\mu$ l 新鲜 LB 培养液（12.5 $\mu$ g/ml 氯霉素）于 37℃、250 rpm 震荡培养 30min；
- (3) 各加入 1 $\mu$ l Copycontrol Induction 诱导液将 BAC 克隆诱导为高拷贝，并于 37℃、300 rpm 震荡培养 120min；
- (4) 将诱导好的 BAC 克隆加入 10ml 新鲜 LB 培养液（12.5 $\mu$ g/ml 氯霉素）于 37℃、280 rpm 震荡培养过夜；
- (5) 按 Sambrook 等人的常规碱裂解方法(2001)提取 BAC DNA，并于 80 $\mu$ l TE 中溶解备用；
- (6) 在 80 $\mu$ l BAC DNA 溶液中加入 1 $\mu$ l Plasmid-safe DNase（10U/ $\mu$ l）、10 $\mu$ l 10 $\times$  缓冲液、5 $\mu$ l 25mM ATP 和 4 $\mu$ l 无菌去离子水共 100 $\mu$ l 反应体系，于 37℃ 恒温培养箱内酶切反应 2h；
- (7) 酶切反应结束后于 70℃ 水浴锅内孵育 30min 终止 DNase 活性；
- (8) 加入 0.7 倍体积的异丙醇于 4℃、12000 rpm 离心 20min，弃去上清液；
- (9) 加入 500 $\mu$ l 75%乙醇漂洗沉淀，12000 rpm 离心 10min，弃去上清液；
- (10) 真空干燥仪干燥沉淀，加入 50 $\mu$ l 去离子水或 TE 溶解沉淀；
- (11) 取 25 $\mu$ l DNA 溶液，加入 0.5 $\mu$ l *NotI*（10U/ $\mu$ l）、3 $\mu$ l Buffer O<sup>+</sup>酶切缓冲液和 1.5 $\mu$ l 去离子水于 37℃ 水浴酶切 5h；
- (12) *NotI* 酶切结束后加入 3 $\mu$ l 10 $\times$ Loading buffer 终止酶切反应；

- (13) 1%脉冲场琼脂糖凝胶电泳 (6V/cm、脉冲时间 0.1-40s 和 14°C) 16h;
- (14) EBtr 染色 30min 凝胶成像系统检测;
- (15) 将凝胶图象输入 TotalLab 2.0 软件, 计算每个 BAC 克隆的插入片段大小;
- (16) 统计插入片段大小分布、平均插入片段数、无插入片段 BAC 克隆数等。

### 2.9.3 大熊猫 BAC 文库筛选系统的建立

基因组 BAC 文库通常由几万到几十万个 BAC 克隆组成, 采用传统的方法筛选目的基因既费力又耗时。因此, 建立高效简便的文库筛选系统, 将有利于降低实验成本和缩短实验周期。本文基于 Asakawa 等人 (1997) 所建立的 4D-PCR 方法, 并运用 PCR 技术建立了大熊猫 BAC 文库的筛选系统。以下是筛选系统的主要内容:

- (1) 每 49 块 96 孔板 (7×7) 组成一个大熊猫 BAC 文库的超级库 (superpool);
- (2) 49 块 96 孔板按 7×7 方式排列, 第 1、8、15、22、29、36 组成 SP<sub>1D-1</sub> 亚库。依次类推, 第 7、14、21、28、35、42、49 组成 SP<sub>1D-7</sub> 亚库。第 1-7 块 96 孔板组成 SP<sub>2D-1</sub>, ……第 43-49 块 96 孔板组成 SP<sub>2D-7</sub> 亚库(图 2-5a)。将 49 块 96 孔板排列成一列, 其中 49 块 96 孔板的 A<sub>1</sub>、A<sub>2</sub>、……A<sub>12</sub> 组成 SP<sub>3D-1</sub>。依次类推, 将 49 块 96 孔板的 H<sub>1</sub>、H<sub>2</sub>、……H<sub>12</sub> 组成 SP<sub>3D-8</sub>; 49 块 96 孔板的 A<sub>1</sub>、B<sub>1</sub>、……H<sub>1</sub> 组成 SP<sub>4D-1</sub>, ……A<sub>12</sub>、B<sub>12</sub>、……H<sub>12</sub> 组成 SP<sub>4D-12</sub> (图 2-5b);



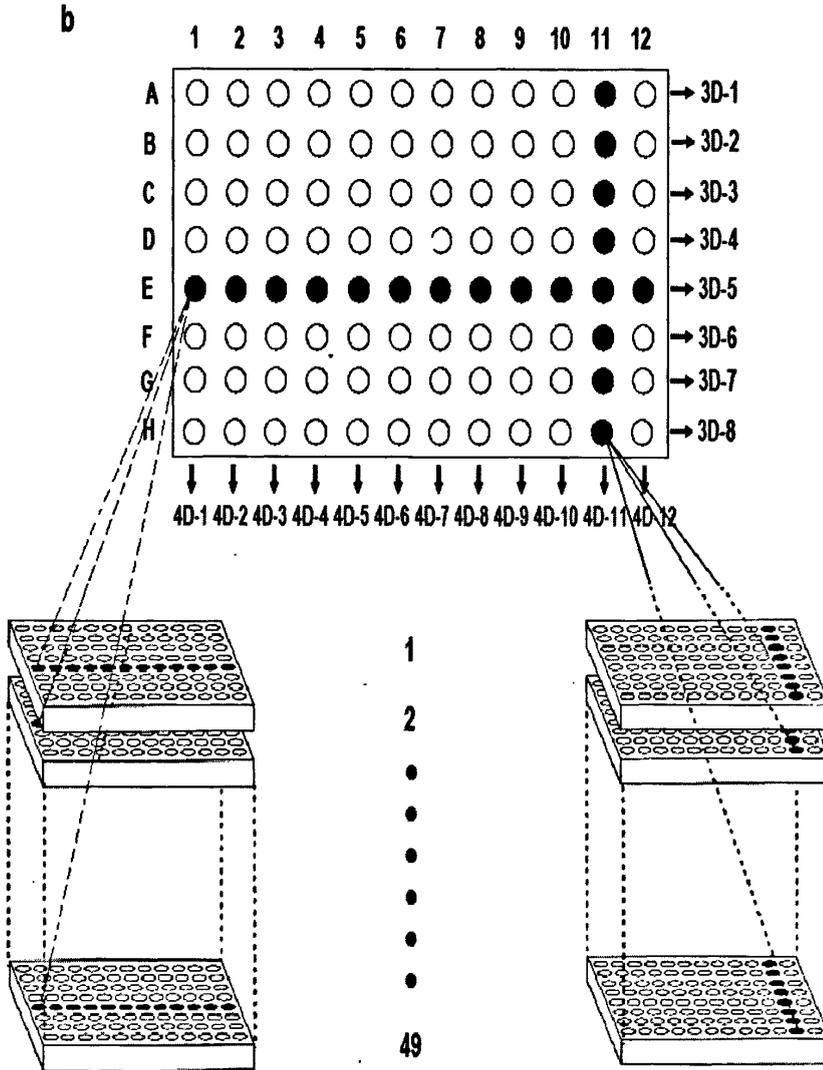


图 2-5：大熊猫基因组 BAC 文库 PCR 筛选系统示意图。(a) 大熊猫基因组 BAC 文库 1D、2D 亚库组成示意图；(b) 3D、4D 亚库组成及筛选示意图。(Based on Asakawa et al., 1997)

Figure 2-5: Construction and characterization of super-pools (SP) based on PCR screening.(a) SP-1D,SP-2D for plates 1-49; (b) SP-3D,SP-4D superpools. (Based on Asakawa et al., 1997)

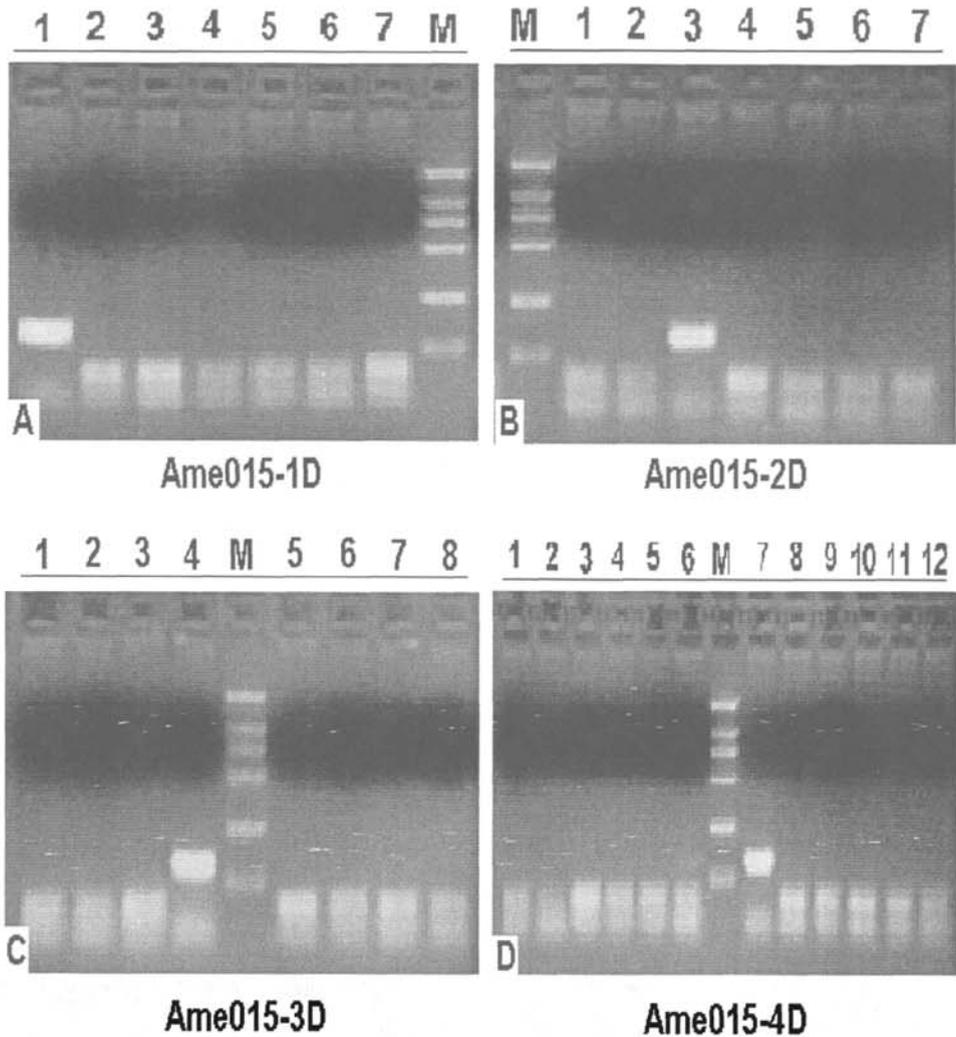


图 2-6: 运用 4D-PCR 方法筛选微卫星引物 Ame015 示意图 (筛选 31 个超级库得到阳性的 SP27, 其中 A、B、C、D 分别代表超级库 SP27 的 1D、2D、3D、4D 亚库; 根据阳性克隆的亚库序号 SP27-1D-1、SP27-2D-3、SP27-3D-4 和 SP27-4D-7, 找到 BAC 克隆 1289D7 为最终的阳性克隆)。

**Figure 2-6:** 4D-PCR screening of the giant panda BAC library using microsatellite primer Ame015. First, screen 31 superpools of the giant panda BAC library and get the positive superpools; then screen further the sub-superpools 1D, 2D, 3D, 4D of positive SP27; finally, according to positive sub-superpools SP27-1D-1, SP27-2D-3, SP27-3D-4 and SP27-4D-7, we find the desired positive BAC clone 1289D7 of primer Ame015.

(3) 每个超级库和亚库中的 BAC 菌液混匀后转入 2ml 无菌离心管, 并加入 10%

甘油置-80℃冷冻保存备用；

- (4) 每个超级库和亚库各取 100 $\mu$ l 加入 0.9ml LB 培养液 (12.5 $\mu$ g/ml 氯霉素) 于 250 rpm、37℃震荡培养 1h；
- (5) 加入 1 $\mu$ l Copycontrol Induction 诱导液于 280 rpm、37℃震荡培养 2h 后将 BAC 克隆诱导为高拷贝；加入 2ml LB 培养液震荡培养过夜；
- (6) 用 Axygen 日常质粒抽提试剂盒抽提 BAC DNA，加入 100 $\mu$ l TE 溶液后置 -20℃保存备用。

#### 2.9.4 PCR 法筛选大熊猫 BAC 文库

- (1) 将设计的 PCR 引物加无菌去离子水稀释，-20℃保存备用；
- (2) 用 1 对 PCR 引物筛选若干个超级库，得到阳性的超级库序号；再用该 PCR 引物筛选阳性超级库所在的 BAC 克隆亚库 (SP<sub>1D</sub>、SP<sub>2D</sub>、SP<sub>3D</sub>、SP<sub>4D</sub> 共 34 个样品)，由此据筛选结果的地址序号，得到阳性的单个 BAC 克隆；
- (3) 通过 PCR 验证筛选出来的单个 BAC 阳性克隆以排除假阳性克隆 (筛选过程见图 2-6)。

#### 2.9.5 BAC 文库覆盖率的计算

基因组 BAC 文库的质量，是由所构建的 BAC 文库的覆盖率和阳性克隆率决定的。相关的计算公式如下：

(1) 基因组覆盖率=BAC 克隆数 $\times$ 平均插入片段/基因组大小

(2)  $N = \ln(1-P) / \ln(1-[I/GS])$  (Clarke & Carbon,1976) (N 为文库中的 BAC 克隆数；P 为期望概率；I 为 BAC 克隆平均插入片段大小；GS 为基因组大小)

## 第二节 结果和讨论

### 1. 大熊猫基因组 BAC 文库的特征

本文构建的大熊猫基因组 BAC 文库共包含 205,800 个 BAC 克隆，保存于 2,100 块 96 孔平板上。经对随机挑取的 174 个 BAC 克隆进行 *NotI* 酶切、脉冲场凝胶电泳，并利用 TotalLab 2.0 软件计算其插入片段长度的结果表明：10 个 BAC 克隆没有检测到插入片段，空载体比率为 5.74% (10/174)；164 个 BAC 克隆的

平均插入片段为 97kb (图 2-7)。

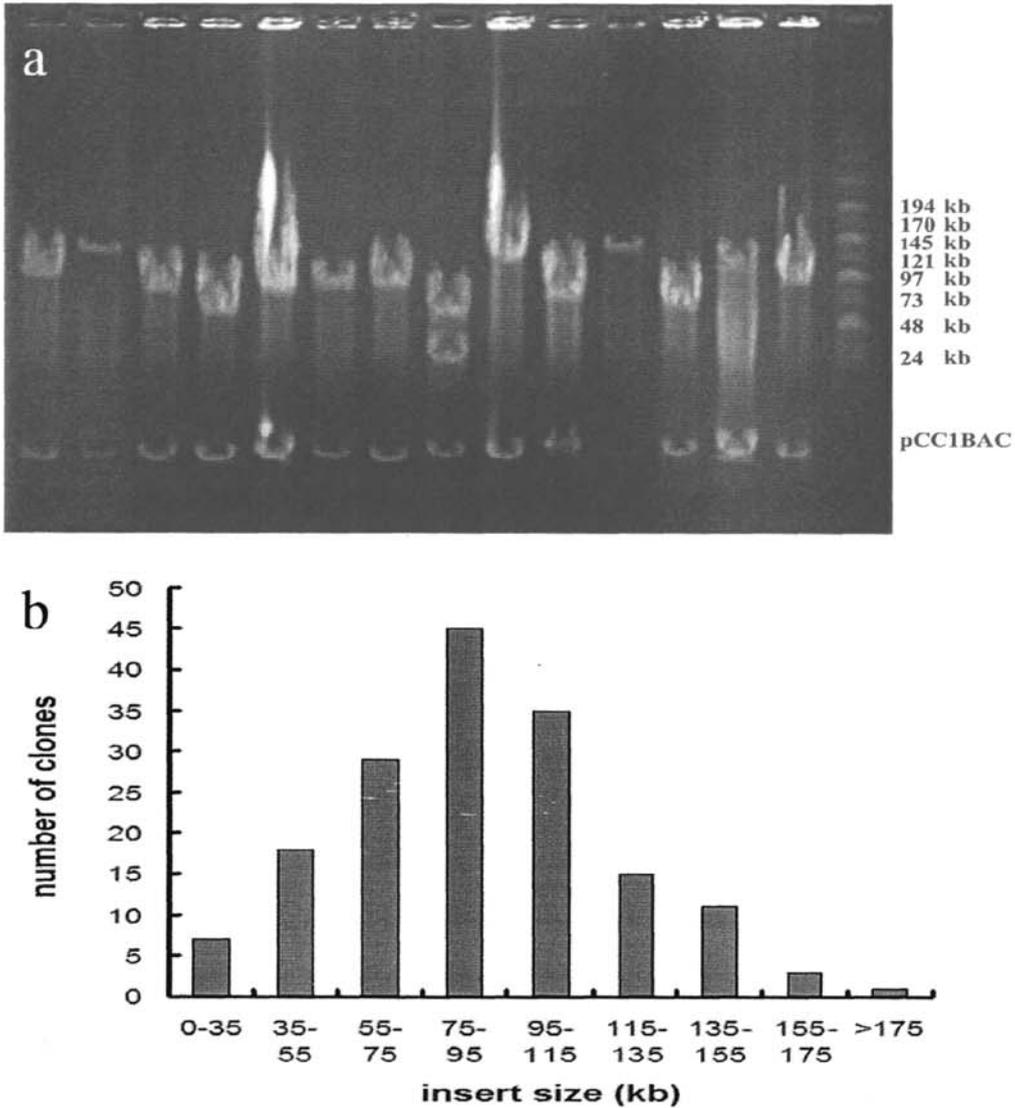


图 2-7: 大熊猫基因组 BAC 文库插入片段长度及片段分布的检测结果图: (a) 部分 BAC 克隆经 NotI 酶切后的脉冲场凝胶电泳图 (载体为 8.1kb; 分子量标记为 MidRange II PFG marker)。 (b) 随机挑取的 164 个 BAC 克隆插入片段的长度分布图。

Fig 2-7. The insert size analysis and distribution of giant panda genomic BAC library. (a) Insert size analysis of some randomly selected BAC clones. The MidRange II PFG markers were used as size standard in the right flanking lane. The size of pCC1BAC vector is 8.1 kb. (b) Distribution of insert size in the giant panda BAC library from randomly selected 164 BAC clones.

据动物基因组的数据库 (<http://www.genomesize.com>) 信息可知, 食肉目熊科 (*Ursidae*) 中的棕熊 (Brown bear) 和喜马拉雅黑熊 (Himalayan black bear) 基因组大小的 C-value 值约为 2.75 (图 2-8)。而大熊猫与食肉目熊科 (*Ursidae*)

中的棕熊 (Brown bear) 和喜马拉雅黑熊 (Himalayan black bear) 同属相同的熊科, 因此我们推测大熊猫的基因组大小与棕熊和喜马拉雅黑熊的基因组大小基本相似。因此, 依据公式  $N = \ln(1-P) / \ln(1-[I/GS])$ , 本文所构建的 BAC 文库约覆盖了 6.8 倍 ( $97000 \times (205800 - 205800 \times 5.74) / 2.75 \times 10^9 = 6.84$ ) 的大熊猫基因组, 且从该文库中获得任一单拷贝或多拷贝大熊猫基因的概率为 99.93%。

计算结果和过程如下:

$$205800 = \ln(1-P) / \ln(1-[97000/2.75 \times 10^9])$$

$$P = 0.9993$$

Class	Order	Family	Species	Common Name	C.value	Ch. Num	Method	Cell Type	Std Sp.	Refs
Phylum Chordata: Subphylum Vertebrata										
Mammalia	Carnivora	Ursidae	<i>Ursus arctos</i>	Brown bear	2.75	74	FCM	SP	HS, RT	542
Mammalia	Carnivora	Ursidae	<i>Ursus hibetarnus</i>	Himalayan black bear	2.75		FCM	LK	HS, OM	598

图 2-8: 动物基因组数据库 (Animal Genome Size Database) 中食肉目熊科的棕熊和喜马拉雅黑熊的基因组大小值 (C-value) (引自 <http://www.genomesize.com>)。

Figure 2-8: The C-value of Brown bear (*Ursus arctos*) and Himalayan black bear (*Ursus hibetarnus*) of Family Ursidae in the Animal Genome Size Database (see <http://www.genomesize.com>).

## 2. 大熊猫 BAC 文库的 PCR 筛选

根据 Asakawa 等人的方法 (1997), 本文构建了大熊猫基因组 BAC 文库的高效筛选系统: 文库共分 43 个超级库, 每个超级库包含  $7 \times 7 = 49$  块 96 孔培养板。每个超级库又分成 34 个亚库 (SP<sub>1D-1</sub>, SP<sub>1D-2</sub>, ..., SP<sub>1D-7</sub>; SP<sub>2D-1</sub>, ..., SP<sub>2D-7</sub>; SP<sub>3D-1</sub>, ..., SP<sub>3D-8</sub>; SP<sub>4D-1</sub>, ..., SP<sub>4D-12</sub>)。

本文共使用 16 对大熊猫特异性引物进行文库的筛选 (引物序列特征及筛选结果见表 2-1)。在 16 对 PCR 引物中, DQA 引物得到了 16 个阳性 BAC 克隆, 其它 15 对引物分别扩增到 3-12 个阳性的 BAC 克隆, 每个 PCR 位点平均获得阳性 BAC 克隆 6.4 个。结合前述之文库的 6.8 倍基因组覆盖率, 表明本文所构建的

表 2-1: 筛选大熊猫基因组 BAC 文库 16 对引物序列特征及结果

Table 2-1: The characterization and results from screening the giant panda BAC library by PCR with 16 pairs of primers

位点	类型	正向引物	反向引物	产物大小 (bp)	阳性 BAC 数	来源
<i>Notch4</i>	MHC class III	CGCTGTGCCTGCCTCAAT	CATCACCTATGCCCTGAC	217	8	GenBank:AY973811
<i>DR4</i>	MHC class II	AGACAAAGTTCGTGCCACC	CTCCACCTITGCAGTCATAGA	172	2	GenBank:AY973816
<i>IGHC</i>	Genomic sequence	CAGCCCTCTTCCTCCTCACTC	CGCCCAGCACTGTGTTCCA	244	5	GenBank:AY818395
<i>MC1R</i>	Genomic sequence	GCGGGCCATCTCCACTAT	TTGAGGGCAGAGGACCCAT	322	6	GenBank:AY884206
<i>DRB</i>	MHC class II	TTCAACCAACGGGACGGAG	GCTGTGCTCCATGAAGTC	166	12	Wan et al., 2005
<i>DQA</i>	MHC class II	CTCACCCCTACTGACCCAGC	TAAAGAGGCAGAGCAIT	297	16	Unpublished data
<i>MHC13</i>	MHC class I	GCTCCCACTCSMTGAGGTATT	CTGCCCTCGCTCTGGTTGTAG	293	6	GenBank:DQ190937 .1
<i>HSP70</i>	MHC class III	TGCTGATCCAGGTGTACGAG	CGTTGGTGAITGGTGAITCTTG	204	8	GenBank:M69100
<i>G001</i>	Microsatellite	ACGGGAAAGCCTGTCTTCTACACTC	AGACACCCAAACCGACTAAACCAC	164-172	6	Zhang et al., 1995
<i>G905</i>	Microsatellite	CTGGCTTCAAACCTGCTTTGAGAG	GCACTGGATTGTGATGCTATC	169-177	5	Zhang et al., 1995
<i>Ameu.11</i>	Microsatellite	TATGCCACCTGCCCCAGAC	GATGGAAAGAGTAGAGCCCAAGG	248-256	4	LÜ et al., 2001
<i>Ameu.15</i>	Microsatellite	AAGCAGTTGTTTTGTCTTAGTG	TGTCAAAGTAITTTGCCCTCACA	142-150	3	LÜ et al., 2001
<i>Ameu.19</i>	Microsatellite	CAGGCAGCACAGCTATACCA	CCACCTGATACCTATGACACAT	174-182	6	LÜ et al., 2001
<i>Ameu.22</i>	Microsatellite	AGGAAAACATGTTGGCCTTTTCA	AGAGGGCAAATAGGAGGGAA	147-149	4	LÜ et al., 2001
<i>Ameu.25</i>	Microsatellite	CATAATCCCTGGCAATGCT	TGCCCGCAITTGAAAAATG	239-257	3	LÜ et al., 2001
<i>Ameu.26</i>	Microsatellite	TTTTCAGGCCTCCGAAAAC	ATTCCCAATAAAGCAAATCAGA	134-140	5	LÜ et al., 2001

大熊猫的基因组 BAC 文库具有较高的质量。

### 3. BAC 克隆稳定性检测结果

经过连续 5 天（约 100 代）的培养，发现文库中的 DNA 插入片段依然能够在大肠杆菌内稳定遗传，未出现插入、缺失和重组现象，与文献（Shizuya et al.,1992）所述的“由于 BAC 载体来源于 F-因子，因此在大肠杆菌宿主内可以保持 1-2 个拷贝，并能够稳定遗传”情况完全一致。

与此同时，本文库采用了含有一个 *oriV* 来源的单拷贝和高拷贝复制子的载体 pCC1BAC。根据其特性，该单拷贝载体在大肠杆菌一个突变的 *trfA* 基因产物的作用下，可诱导成高拷贝数量。本研究中经过高拷贝数的诱导，BAC 克隆依然能够稳定遗传，从而证明了所建的大熊猫基因组 BAC 文库是稳定的。

## 4. 大熊猫基因组 BAC 文库构建中的几个关键问题

### 4.1 片段选择与平均插入片段的关系

高分子量 DNA 不完全酶切结束后，去除小分子量的 DNA 片段就成为基因组 BAC 文库构建中的关键步骤之一。虽然本文所构建的大熊猫基因组 BAC 文库的平均插入片段的长度与其它学者（Shan & Hardham, 2004; Gill et al.,1999; Le Paslier et al.,2000）的文库相当（约 97kb），但比另外一些学者的平均插入片段却要小一些（Suzuki et al.,2000; Rogel-Gaillard et al.,2001）。

在动物基因组 BAC 文库的构建工作中，小片段 DNA 的去除是必须面对的问题。虽然脉冲场电泳（PFGE）技术能够有效的分离不同分子量的 DNA 片段（Birren et al.,1993; Ganal & Tanksly, 1989），但由于不同分子量的 DNA 片段在脉冲场电泳过程中存在着“共迁移”现象，因而使得小片段的 DNA 难以被完全的去除掉（Cheng et al., 1995; Cai et al., 1995）。为此，一些学者采用预电泳的方法来去除部分小片段的 DNA 和一些抑制因子，但获得的结果却并不理想（Edwards et al.,1992）。再者，一些学者为了在文库构建中获得较高的转化率，便采用了一次性片段选择的方法（Frijters et al.,1997）。该方法虽然可以提高平均插入片段的长度，但由于一次性片段选择的方法，不能够完全去除小分子量 DNA 的污染，

所以其实际效果也不够理想 (Ioannou et al.,1994)。有鉴于此, 一些学者设计了两次片段选择的实验方法来去除“共迁移”小片段 DNA 的污染, 从而通过提高文库的平均插入片段长度和插入片段的均一度 (Woo et al.,1994; Cai et al.,1995; Osoegawa et al.,1998; Frijters et al.,1997)使插入片段达到了 120-180kb, 甚至 300kb (Zimmer et al.,1997), 但该方法的负面结果是转化率极低。针对上述情况, 有学者则在两次片段选择的基础上, 通过脉冲场电泳的进一步浓缩, 去除部分夹杂的小分子量 DNA。然而, 本文在建库过程中运用了这一方法, 但获得的效果却并不明显。

总之, 在酶切、电泳和透析等步骤的过程中, 核酸酶和机械力作用等因素, 均会造成高分子量 DNA 的断裂。同时, 脉冲场的电泳参数也是影响高分子量 DNA 分离的一个不可忽略的重要因素。正是由于众多因素的存在, 才给动物高质量基因组 BAC 文库的构建带来了不小的困难。因此, 在构建文库的过程中, 小心仔细的实验操作, 以及针对不同的物种, 反复进行脉冲场电泳参数的设定与对比实验的分析摸索, 即是基因组 BAC 文库构建中的关键步骤之一。

## 4.2 DNA 回收方式的选择

在基因组 BAC 文库的构建过程中, 琼脂糖酶 (GELase) 的消化处理和电洗脱 (electroelution) 方法, 是回收大片段 DNA 的最常用的两种方法。前者通过琼脂糖凝胶酶消化低熔点的琼脂糖凝胶, 可将大片段的 DNA 分子释放出来 (McCormick et al.,1990; Larin et al.,1991; Frijters et al.,1997)。然而, 经过凝胶酶处理后的高分子量 DNA 的完整性会在很大程度上遭到破坏 (Woon et al.,1998)。鉴于此, Strong 等人通过对加热、凝胶酶处理和电洗脱回收 DNA 的方法进行了比较研究, 发现利用电洗脱的方法回收 DNA 不仅效率高, 而且 DNA 的片段长, 连接效果好 (Strong et al.,1997; Osoegawa et al.,1998)。因此, 本文在研究中选择电洗脱的方法来回收大片段的 DNA 所取得的良好结果, 表明了选择适宜的大片段 DNA 的回收方法, 是保证成功构建高质量基因组 BAC 文库的关键因素之一。

## 4.3 电转化参数与插入片段大小的关系

所谓电转化方法, 就是通过简短的高压脉冲, 在大肠杆菌细胞上形成暂时

性的小孔隙，从而使质粒 DNA 在电场力的作用下进入到宿主细胞内 (Neuman et al.,1989; Klenchin et al.,1991; Chang et al.,1992)。电转化的效率通常受到很多因素的影响，如温度、电压梯度、电阻及电容等。另外，宿主菌的质量、遗传背景、生长条件，以及电转化后的处理和 DNA 样品的拓扑结构等，都会直接影响到电转化的效率和文库构建的质量。Leonardo 等人在研究了质粒大小与大肠杆菌转化效率之间的关系 (Leonardo & Sedivy, 1990) 后发现：质粒越大，转化效率越低。Sheng 等人的研究 (Sheng et al.,1995) 表明，低电压梯度更有利于大质粒的转化。因此，针对不同大小的质粒，选择不同的电转化条件 (电压梯度和电阻值等)，即可获得较为优化的实验结果 (Zimmer et al.,1997)。

本文在对上述方法进行比较实验后，发现：(1) 采用低电压梯度 (0.9-1.1 kV) 的转化，不仅效率低，而且插入片段也未能得到增加；(2) 采用高电压梯度 (1.7-2.5kV) 的转化，虽然效率极高，但插入片段较小；(3) 选择 1.3 kV 或 1.5 kV、100 $\Omega$  和 25 $\mu$ F 的电转化条件，既获得了较高的转化效率，又保证了较大插入片段的实验效果。因此，在动物高质量基因组 BAC 文库的构建中，摸索出适宜的电转化条件，是其最为关键的环节之一。

#### 4.4 基因组 BAC 文库的筛选

在基因组 BAC 文库的早期筛选方法中，菌落原位杂交、Southern 杂交、滤膜杂交，以及由此衍生的二维 PCR 筛选等方法 (Maier et al.,1992; Ross et al.,1992; Cai et al.,1995; Foster, 1995)，都在一定的程度上取得了良好的研究结果，但方法耗时费力，往往需要几天的时间才能获得最终的阳性克隆。四维 PCR (4D-PCR) 方法的建立，则可只需要快速简便的 PCR 反应，就能得到所需的阳性克隆 (Asakawa et al.,1997)。

本文运用 4D-PCR 方法建立的大熊猫基因组 BAC 文库的筛选系统，只需  $43+14+20=77$  个 PCR 反应，即可确定阳性的 BAC 克隆数。这对于提高实际工作效率，保证研究方法的可靠性，具有十分重要的意义。

### 第三节 结 论

本文成功的构建了大熊猫基因组的 BAC 文库。文库包含 205,800 个 BAC 克

隆,保存于2100块96孔的培养板上。文库中大熊猫DNA的平均插入片段为97kb,约覆盖了6.8倍的大熊猫基因组。文库中BAC克隆能够在大肠杆菌宿主细胞内稳定遗传。经16对PCR引物的检测结果表明,每1对引物平均能够在大熊猫BAC文库中成功的扩增到6.4个阳性克隆,证实了该文库具有较高的覆盖率。因此,本文所构建的大熊猫的基因组BAC文库,达到了保护大熊猫的遗传基因资源,并为相关的基础科学研究提供基因材料的共享平台的目的。

## 第四章 大熊猫主要组织相容性复合体（MHC）基因的 染色体定位

荧光原位杂交（Fluorescence In Situ Hybridization, FISH）技术是 20 世纪 80 年代初期由 Langer 等人建立起来的（Langer et al.,1981）。该技术的原理是利用荧光标记好的 DNA 或 RNA 与细胞核染色体中的核苷酸杂交，以检测经荧光标记的 DNA 或 RNA 在细胞或染色体上的位置、数量以及结构改变等。FISH 技术通常使用 3 种类型的探针：克隆化探针、酶促扩增探针和化学合成探针。酶促扩增探针是指任何带有特定引物的 DNA 都可以进行探针序列扩增（如 PCR 扩增或制备探针文库），探针序列可以作为一个扩增片段在扩增时标记。化学合成探针包括简单和特殊的寡核苷酸探针，主要用于检测一些高度重复的序列或具有特殊结构的序列（Beatty et al., 2002）。克隆化探针主要包括粘粒探针、噬菌体探针、质粒探针、YAC 探针、PAC 探针和 BAC 探针等。克隆化的 DNA 探针能够提供很好的 FISH 检测信号，而且不需要二次扩增，是目前 FISH 技术中最常用的探针类型之一。

目前，FISH 技术是在真核细胞染色体区带上定位某一特定 DNA 片段或部分基因组之快速而又有效的技术方法之一。荧光原位杂交技术与相关的其他原位杂交技术相比的优点在于：不需要放射性同位素标记、实验周期短、非特异性杂交信号污染较少、标记的探针稳定性好、信号检测方便快捷。也正因为如此，FISH 技术才在近年来的基因组学和分子细胞遗传学的相关研究中得到广泛的应用（Beatty et al.,2002; Geigl et al., 2006; Di Meo et al., 2007）。

近年来，有不少学者已成功的运用 FISH 技术将多种动物的 MHC 基因组定位于所在的染色体上，如 Rabin 等人将猪的 I 类 MHC 基因组定位于 7 号染色体短臂上（Rabin et al.,1985）；Mäkinen 以及 Gustafson 等人将马的 I、II、III 类 MHC 基因组定位于 20 号染色体长臂上（Mäkinen et al.,1989; Gustafson et al., 2003）；Larsson 以及 Helou 等人将大鼠 MHC 基因组定位于 20 号染色体短臂（20p）上（Larsson et al., 1996; Helou et al.,1998）；而 Huber 等人则将恒河猴（rhesus

macaque) 包含 7 对 MHC 基因定位于 6 号染色体上的 6q24 (Huber et al.,2003)。然而, 有关大熊猫的 MHC 基因组究竟位于哪号染色体上这一问题, 除 Nash 等人 (Nash et al., 1998) 利用染色体杂交的方法将其初步框定在第 5 号染色体短臂、第 10 号染色体长臂 (实际为 9 号染色体) 和第 19 号染色体上之外, 其进一步精确定位的研究, 却迄今未见报道。有鉴于此, 本文将运用 FISH 技术来进行大熊猫 MHC 基因组的染色体精确定位。

## 第一节 材料和方法

### 1. 材 料

#### 1.1 样品来源

大熊猫成纤维细胞由中国科学院动物学研究所陈大元研究员惠赠。成纤维细胞的培养及中期染色体的制备工作在中国科学院动物研究所生殖生物学国家重点实验室完成。

#### 1.2 BAC 克隆

BAC 克隆 59G10 筛选自本文所构建的大熊猫基因组 BAC 文库。BAC 克隆 59G10 所包含的大熊猫 II 类 MHC 基因的 DQA 位点, 经 PCR 扩增并测序予以验证。

#### 1.3 主要仪器设备

CHEF DR-III 脉冲场电泳系统 (Bio-Rad); 5417R 高速冷冻离心机 (Eppendorf); 5810R 高速冷冻离心机 (Eppendorf); PTC-200 PCR 扩增仪 (MJ Research 公司); Ultra Freeze 超低温冰箱 (Heto); Power PAC 300 电泳仪 (Bio-Rad); Gel Doc2000 凝胶成像系统 (Bio-Rad); 恒温混匀器 Thermomixer comfort 5355 (Eppendorf); Venticell 111R 烘箱 (MMM); SS-325 灭菌柜 (Tomy); DNA mini 真空干燥系统 (Heto); SmartSpec 3000 紫外分光光度计 (Bio-Rad); Purelab UHQ 去离子纯水仪 (ELGA); Vortex-2 genie 旋涡混匀器; Olympus BX61 荧光显微镜; Li-Cor4200 测序仪。

## 1.4 主要试剂和耗材

Large-construct kit 购自 Qiagen; Lambda Ladder (Bio-Rad); SequiTherm EXCEL II DNA Sequencing Kit 购自 Epicentre 公司; Digoxigenin-dUTP 购自 Roche 公司; Lambda mix marker-19 购自 Fermentas 公司、Chloramphenicol 购自 Sigma 公司、Ex-Taq DNA polymerase、T4 连接酶、pUC19 载体、Mung Bean Nuclease 和 *HindIII* 购自宝生物工程(大连)有限公司; IPTG、X-gal、*NotI* 及其它常规试剂, 均购自上海生工生物工程公司; 日常质粒快速抽提试剂盒 (Axygen) 和脉冲场用琼脂糖购自 Bio-Rad 公司; PCR 引物由上海生工生物工程公司合成。

## 2. 实验方法和步骤

### 2.1 大熊猫成纤维细胞中期染色体制备

大熊猫成纤维细胞中期染色体的制备, 主要参照王宗仁等人的方法(1990), 并作了部分调整和修改:

- (1) 对液氮保存的大熊猫成纤维细胞解冻后, 加入 5ml 新配制的 F10 培养液, 并置 37°C 的 CO<sub>2</sub> 培养箱中培养传代;
- (2) 每天观察细胞的生长情况: 待细胞生长至 70-80% 汇合时, 加入 0.5 $\mu$ g/ml 秋水仙素, 继续于 37°C 的 CO<sub>2</sub> 培养箱中培养 3-5h;
- (3) 移液器吸出培养液, 加入 2ml 0.25% 的胰蛋白酶消化贴壁的成纤维细胞(保持 30-60s), 并于显微镜下观察细胞的消化情况; 待大多数细胞成圆形时, 立即加入 3ml 含 BSA 或血清的 F10 培养液终止消化;
- (4) 于 1000 rpm 离心 10min; 弃上清后加入 4ml 0.075M 的 KCL 于 37°C 水浴孵育 30min;
- (5) 加入 1ml 甲醇与冰醋酸 (3: 1) 的固定液, 混匀后于 1000rpm 离心 10min;
- (6) 弃上清后加入 4 ml 甲醇与冰醋酸 (3: 1) 的固定液, 室温保持 10-15min;
- (7) 1000rpm 离心 10min 弃上清; 重复步骤 (6) 3 次;
- (8) 加入 300 $\mu$ l 新鲜固定液, 室温保存备用;
- (9) 将载玻片用浓硫酸/重铬酸钾浸泡 72h; 自来水冲洗和三蒸水漂洗后, 加入超纯水于 4°C 保存备用;

- (10) 取出 4℃ 保存的载玻片，甩掉多余的水分后立即用移液器吸取 30 $\mu$ l (约 3 滴) 已固定好的细胞悬液于 30-40cm 高度滴打于载玻片上；
- (11) 将载玻片置于室温或 50℃ 恒温板上，待固定液挥发直至出现颗粒样时，置 70℃ 水蒸气中保持 2-3s；
- (12) 室温干燥后置显微镜下观察中期染色体的形态及分散程度并保存备用。

## 2.2 BAC 克隆的筛选及 BAC DNA 的制备

根据牛的 II 类 MHC 基因的 DQA 引物序列 (Heb et al., 1999)，本文合成了如下筛选大熊猫 BAC 克隆的一对 II 类 MHC 基因 DQA 位点的引物：

Forward primer: 5'-CGACTCAGATGACCACATTG-3'；

Reverse primer: 5'-GGGGACACATACTG TTGGTAG-3'。

利用上述引物，以及本文基于 4D-PCR 的文库筛选方法 (Asakawa et al., 1997) 所建立的大熊猫 BAC 克隆的高效筛选系统，作者进行了包含大熊猫 II 类 MHC 基因 DQA 位点的阳性克隆的筛选工作：

- (1) 利用上述引物进行如下反应体系 (2 $\mu$ l EX-Taq buffer, 1.5 $\mu$ l dNTP, 1.5 $\mu$ l Forward primer, 1.5 $\mu$ l Reverse primer, 0.1ul EX-Taq, 2 $\mu$ l BAC DNA, ddH<sub>2</sub>O up to 20 $\mu$ l) 的 PCR 扩增：94℃ 5min；35 个循环 (94℃ 45sec、56℃ 45sec、72℃ 45sec)；72℃ 5min；
- (2) 筛选得到的阳性 BAC 克隆重新进行 PCR 验证；
- (3) 阳性 BAC 克隆划线接种于 LB+氯霉素 (12.5 $\mu$ g/ml) 平板上，37℃ 培养过夜；
- (4) 挑取单个 BAC 菌落，加入 1ml LB+氯霉素培养液于 280 rpm、37℃ 震荡培养过夜；
- (5) 吸取 1 $\mu$ l BAC 菌液重复 PCR 反应，验证 BAC 克隆；
- (6) 200ml LB 培养液大量培养阳性的 BAC 克隆；
- (7) 常规碱裂解的方法制备 BAC DNA 。

## 2.3 阳性 BAC 克隆的验证

### 2.3.1. PCR 产物验证

- (1) 将上述 II 类 MHC 基因 DQA 位点的引物按如下反应体系进行 PCR: 2.5 $\mu$ l

10×Ex-Taq buffer、2 $\mu$ l dNTP、10nmol/L DQA 正反向引物各 1.5 $\mu$ l、BAC DNA 2 $\mu$ l、Ex-Taq 酶 0.15 $\mu$ l 并加水至 25 $\mu$ l。PCR 反应条件为：94 $^{\circ}$ C 5min、94 $^{\circ}$ C 1min、56 $^{\circ}$ C 1min 和 72 $^{\circ}$ C 1min，共 35 个循环；

- (2) PCR 产物置 1%的琼脂糖凝胶中电泳，回收目的条带（操作程序按厂家说明书进行）；
- (3) 将回收的 PCR 产物连接入 pMD18-T 载体。

### 2.3.2. Shotgun 测序验证

- (4) 将 BAC DNA 用细胞粉碎机随机断裂为 1-3kb 长度的 DNA 片段，加入 0.6 倍体积的无水乙醇沉淀，25 $\mu$ l TE 溶解；
- (5) 随机断裂的 DNA 溶液中加入 1 $\mu$ l Mungbean Nuclease 于 37 $^{\circ}$ C 孵育 10min 后立即放入 65 $^{\circ}$ C 水浴中终止酶切反应；
- (6) 酶切修饰后的 DNA 溶液置 1%琼脂糖凝胶中电泳，切割回收 1.5-3kb 大小范围的 DNA 片段（回收程序按厂家说明书进行）；
- (7) 在 7 $\mu$ l 回收溶液中加入 1 $\mu$ l 经 *Sma*I 酶切、CIAP 去磷酸化的 pUC19 载体，1 $\mu$ l T4 连接酶（15U/ $\mu$ l），16 $^{\circ}$ C 连接过夜；
- (8) 将连接好的产物用 DH5a 大肠杆菌感受态细胞转化（Sambrook et al.,2001）；挑取白色菌落培养并测序；
- (9) 用 BLASTN 程序对测序结果进行 GenBank 同源性比对，验证该 BAC 克隆是否包含有大熊猫的 II 类 MHC 基因序列。

## 2.4 荧光探针 BAC DNA 的标记及纯化

运用缺口平移法（Langer et al.,1981）将 Digoxigenin-11-dUTP 标记到 BAC DNA 上后：

- (1) 在 25 $\mu$ l 反应体系中，加入 2.5 $\mu$ l 10×NT 缓冲液、2.5 $\mu$ l Dntp、2.5 $\mu$ l Digoxigenin-11-dUTP、0.4 $\mu$ l DNase、10 $\mu$ l BAC DNA（约 5 $\mu$ g/ml）、0.6 $\mu$ l DNA Polymerase I 和 6.5 $\mu$ l 去离子水；15 $^{\circ}$ C 恒温箱中孵育 2h，加入 5 $\mu$ l 0.5M EDTA 终止反应；
- (2) 制备葡聚糖 G-50 纯化柱，加入 TE 溶液，8000 rpm 反复离心，直至纯化柱达到进出平衡为止；
- (3) 将步骤（1）中缺口平移法标记好的混合液加入 G-50 纯化柱，8000 rpm

离心 5min;

- (4) 取 2 $\mu$ l 纯化后的探针 DNA 于 1.5% 琼脂糖凝胶电泳; 凝胶成像检测, 判断探针的片段范围及质量 (探针闭光-20 $^{\circ}$ C 保存备用)。

## 2.5 探针与中期染色体的杂交

### 2.5.1 中期染色体载玻片的预处理

- (1) 将载玻片放入 37 $^{\circ}$ C 含有蛋白酶 K 溶液的染色缸中 7min;
- (2) 2 $\times$ SSC 浸泡 5min, 重复 3 次;
- (3) 70%、90%、100% 乙醇各浸泡 3-5min, 取出载玻片晾干;

### 2.5.2 变性和杂交

- (1) 在载玻片上加入 100 $\mu$ l 70% 的甲酰胺溶液, 盖上盖玻片; 92-94 $^{\circ}$ C 下充分变性 2.5-3.5min;
- (2) 取出载玻片并迅速揭去盖玻片后立即浸入-20 $^{\circ}$ C 预冷的 70% 乙醇中浸泡 5min;
- (3) 将探针混合物 (3 $\mu$ l 标记好的探针 DNA、4 $\mu$ l 鲑鱼精 DNA、2 $\mu$ l 20 $\times$ SSC、10 $\mu$ l D-F 溶液、4 $\mu$ l 50%DX 溶液) 混匀后快速离心; 91 $^{\circ}$ C 水浴变性 5min 后迅速取出置于 0 $^{\circ}$ C 冰水中保持 15min;
- (4) 将 20 $\mu$ l 探针 DNA 混合液加到载玻片中, 盖上盖玻片 (小心气泡);
- (5) 将载玻片放入密闭的湿盒于 37 $^{\circ}$ C 下孵育过夜 (8h 以上)。

## 2.6 杂交后处理

### 2.6.1 杂交后的洗涤

- (1) 从密闭湿盒中取出载玻片, 浸入 2 $\times$ SSC 溶液中直至盖玻片自行脱落 (约 5min);
- (2) 取出载玻片, 浸入 42 $^{\circ}$ C 2 $\times$ SSC 溶液中 10min;
- (3) 2 $\times$ SSC 溶液室温中浸泡 5min;
- (4) 1 $\times$ PBS 溶液中浸泡 5min, 取出晾干备用。

### 2.6.2 染色

- (1) 加入 100 $\mu$ l 抗体溶液 (1 $\mu$ l Dig-Rod、20 $\mu$ l 5 $\times$ block buffer 和 79 $\mu$ l ddH<sub>2</sub>O) 于载玻片上;

- (2) 用深色塑料薄膜将载玻片表面封住后放入湿盒于 37℃ 保温 30min;
- (3) 取出载玻片, 浸泡于 1×PBS 溶液中; 去除塑料薄膜;
- (4) 用 1×PBS 溶液漂洗 3 次 (每次 5min), 取出晾干;
- (5) 加入 15μl 抗褪色剂 (含 DAPI 荧光染料), 盖上盖玻片, 放入密闭湿盒中于 4℃ 保存备用。

## 2.7 镜检及图象处理

在 Olympus BX61 荧光显微镜下观察 II 类 MHC 基因探针在大熊猫中期染色体上的杂交信号, 并记录观察的细胞数和杂交信号的位置 (图象处理采用 Adobe photoshop 6.01 软件)。

## 第二节 结果和讨论

根据牛的 II 类 MHC 基因的 DQA 引物序列 (Heb et al., 1999), 本文合成了如下筛选大熊猫 BAC 克隆的一对 II 类 MHC 基因 DQA 位点的引物:

Forward primer: 5'-CGACTCAGATGACCACATTG-3';

Reverse primer: 5'-GGGGACACATACTG TTGGTAG-3'。

利用上述引物, 以及本文基于 4D-PCR 的文库筛选方法 (Asakawa et al., 1997) 所建立的大熊猫 BAC 克隆的高效筛选系统, 作者进行了包含大熊猫 II 类 MHC 基因 DQA 位点的阳性克隆的筛选工作:

本文利用 DQA 引物, 并采用 4D-PCR 的方法在所构建的大熊猫基因组 BAC 文库中, 筛选到了 3 个阳性 BAC 克隆 (59G10, 71F2, 77F8)。经鸟枪法 (shotgun 测序) 随机对 59G10 进行测序, 发现其序列结果与 GenBank 中其它动物的 MHC 基因序列高度同源, 证实了所筛选的 BAC 克隆 59G10 中, 包含着大熊猫的 II 类 MHC 基因。

在大熊猫中期染色体固定铺片过程中, 我们获得了分散均匀, 且染色体形态良好, 并有利于 FISH 杂交和信号检测的玻片 (图 3-1)。从图 3-1 中可以知道: 大熊猫共有 42 (2n=42) 条染色体, 其中 20 对常染色体, 1 对性染色体, 与文献报道的结果 (O'Brien et al., 1985) 相一致。

运用缺口平移法对 BAC 克隆 59G10 进行 Digoxigenin 荧光标后，再与大熊猫中期染色体进行 FISH 杂交，共观察了 15 个细胞中期染色体的 FISH 杂交图谱。结果发现 MHC 的 DQA 位点均成功的在 9 号染色体长臂靠近着丝粒的位置实现了杂交，表明大熊猫的 MHC 基因组位于 9 号染色体的长臂上（见图 3-2），完成了大熊猫 MHC 基因组在其染色体上的精确定位。

研究中我们注意到，适当的控制染色体铺片的干燥速度，是获得分散均匀和无重叠中期染色体的关键。本文的体会条件是：若室温较高且比较干燥，就应在载玻片刚开始出现“颗粒”状物质时，立即将其置于 70℃ 水蒸汽中约 2-3s，以利于染色体的伸展；但当室温高湿度大时，就应缩短载玻片在水蒸气中的时间或放弃此步骤。

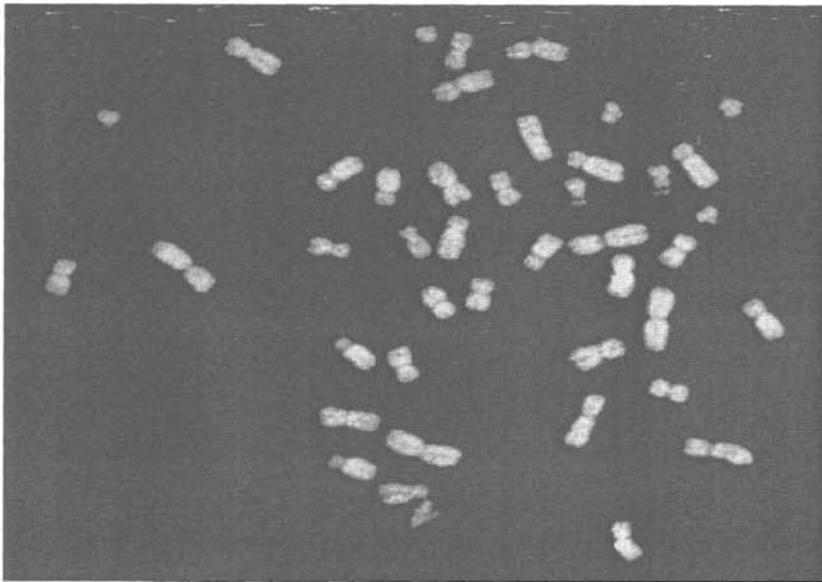
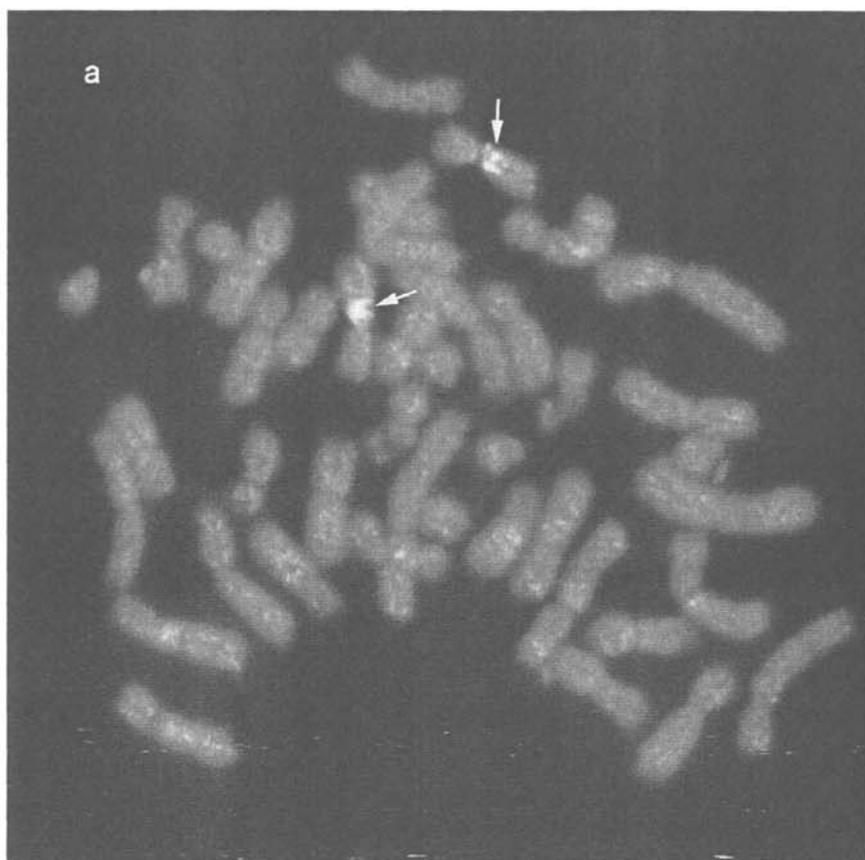


图 3-1 大熊猫中期染色体铺片。

Fig.3-1. the metaphase chromosomes of giant panda



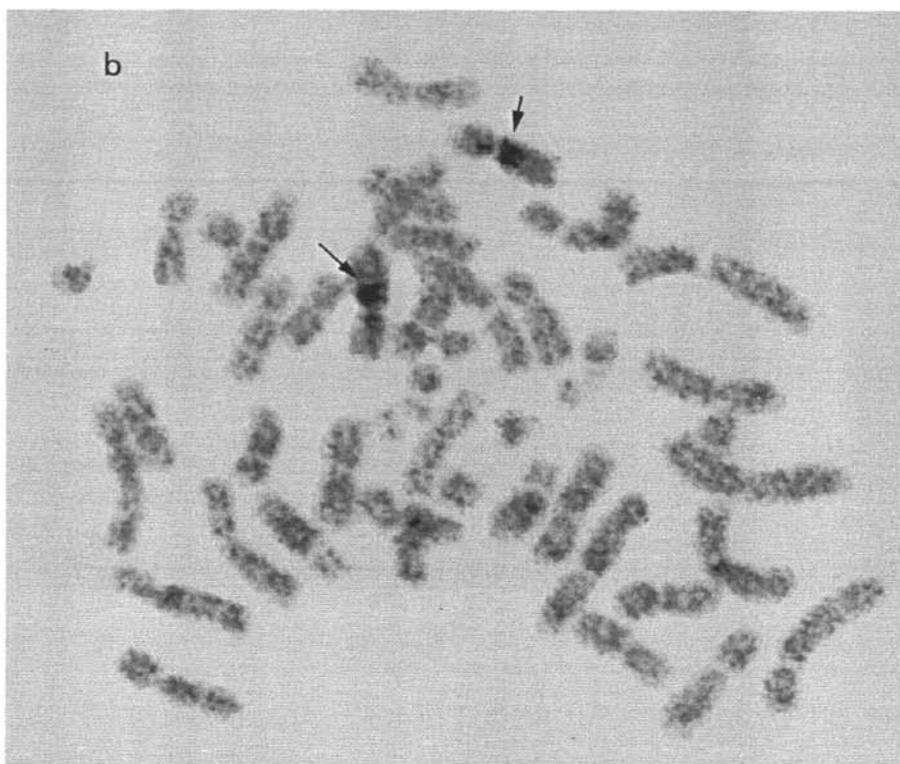


图 3-2: 大熊猫 II 类 MHC 基因组的 FISH 杂交结果(标记的 DNA 探针来自于包含 II 类 MHC 基因 DQA 位点的部分基因序列的大熊猫 BAC 克隆 59G10。图中箭头表示杂交位点): (a) 地高辛标记的 BAC DNA 探针与中期染色体的 FISH 杂交图; (b) DAPI 染色体带型图。

Fig.3-2 FISH mapping of the MHC class II of the giant panda. The DNA probe prepared from the 59G10 clone containing the DQA locus and MCH class II partial sequences. Arrows indicate chromosomes with hybridization sites. (a) The FISH map of BAC DNA probes with metaphase chromosomes of giant panda. (b) The same metaphase chromosome spreads with the hybridization signals superimposed on reverse DAPI-banded chromosomes

### 第三节 结论

本文根据牛的 II 类 MHC 基因的 DQA 引物序列, 合成了如下筛选大熊猫 BAC 克隆的一对 II 类 MHC 基因 DQA 位点的引物:

Forward primer: 5'-CGACTCAGATGACCACATTG-3';

Reverse primer: 5'-GGGGACACATACTG TTGGTAG-3'。

利用上述引物, 以及本文基于 4D-PCR 的文库筛选方法所建立的大熊猫 BAC

克隆的高效筛选系统，筛选出了 3 个包含大熊猫 II 类 MHC 部分基因的阳性克隆（59G10, 71F2, 77F8）。经鸟枪法（shotgun 测序）随机对 59G10 进行测序，发现其序列结果与 GenBank 中其它动物的 MHC 基因序列高度同源，证实了所筛选的 BAC 克隆 59G10 中，包含着大熊猫的 II 类 MHC 基因。

在大熊猫中期染色体固定铺片过程中，本文获得了分散均匀，且染色体形态良好，并有利于 FISH 杂交和信号检测的玻片。

运用缺口平移法对 BAC 克隆 59G10 进行 Digoxigenin 荧光标后，再与上述大熊猫中期染色体进行 FISH 杂交，结果发现 MHC 的 DQA 位点均成功的在 9 号染色体长臂靠近着丝粒的位置实现了杂交，表明大熊猫的 MHC 基因组位于 9 号染色体的长臂上，完成了大熊猫 MHC 基因组在其染色体上的精确定位。

## 第五章 大熊猫 MHC II 类基因组的物理图谱构建

由于脊椎动物的 MHC 既是机体免疫应答的“开关”基因，又是亲缘避免之性选择机制的主要决定因子之一，所以，有关 MHC 基因组的结构和功能的研究，受到了相关学者的极大关注，并使之在短短的时间内，就相继完成了人(The MHC sequencing consortium, 1999)、鸡(Kaufman et al., 1999)、小鼠(Kumánovics et al., 2003)、猪(Renard et al., 2006)等多种动物的全部 MHC 基因组以及狗(Debenham et al., 2005)、猫(Yuhki et al., 2003)等多种动物的 MHC II 类基因组的精细图研究，从而为相关学者开展野生动物特别是濒危野生动物免疫遗传、疾病防御、系统演化、性选择机制和繁殖策略等方面的研究奠定了的基础。

然而，在诸如性选择机制的研究中，要想从根本上揭示其分子机理或规律性问题，获得群体水平上全部 MHC 位点扫描检测的基础数据乃是其关键中的关键。但是，由于基于上述人或动物的 MHC 基因组的数据库仍难以通过跨物种 PCR 扩增的方法，完全分离出目的物种 MHC 基因组中的每一个功能位点，所以，近期相关的报道，都仅仅限于从庞大的 MHC I 类或 II 类的基因组中挑选出一个或少数几个位点，进行多样性或杂合度与动物择偶取向之间的相关性分析(Wan et al., 2006; Zhu et al., 2007; Brown et al., 2004; Ziegler et al., 2005)，从而难以解决动物性选择机制的分子机理问题。因此，精确的解析目的物种 MHC 基因组中的每一个功能位点，即逐一分离出 MHC 基因组中的全部功能位点，为群体水平上所有 MHC 位点的扫描检测提供可靠的系统支撑标记，就成其为诸如性选择机制、免疫防御能力和繁殖策略等多方面研究的关键。

文献表明，通过构建物理图谱的研究方案和相关的技术路线来开展 MHC 的基因组计划(MHC 基因组的全序列测定)，是分离解析 MHC 基因组全部功能位点的最为有效的方法之一(Gillett et al., 1996; Yu & Wong, 1997; Wong et al., 1997)。因此，本文拟在已精确的将大熊猫的 MHC 基因组定位于第 9 号染色体长臂上(Zeng et al., 2005)的工作基础上，继续开展“大熊猫 MHC II 类基因组物理图谱的构建”的研究，旨在为本实验室后续开展大熊猫 MHC 的基因组计划，并从 MHC 的基因组中分离鉴定 MHC 的全部功能位点，从而深化大熊猫的保护遗传学研究，提供必要的前期工作积累。

## 第一节 材料和方法

### 1. 材料

#### 1.1 材料来源

大熊猫 MHC II 类基因组物理图谱构建的材料，来自于本文上述所构建的大熊猫基因组 BAC 文库。

#### 1.2 引物设计

利用自 Genebank 中获得的人、大熊猫、家猫和狗的 II 类 MHC 基因序列 (the MHC sequencing consortium, 1999; Debenham et al., 2005; Yuhki et al., 2003), 以及 Primer Premier 5.0 软件, 设计了从大熊猫的基因组 BAC 文库中筛选 MHC II 类基因阳性克隆的 PCR 引物 (表 5-1)。实验中 BAC 末端引物的设计, 则是在上述实验的基础上根据 BAC 克隆的测序结果而设计的 (表 5-2)。其中, DRB<sub>1</sub> 和 DQA 引物, 均源于本实验室已发表的论文数据 (群体验证表明了二者只能扩增到大熊猫的 MHC II 类基因的 DRB 和 DQA 两个位点) (Wan et al., 2006; Zhu et al., 2007)。

为防止重复序列对末端引物设计的影响, 本文或通过 RepeatMasker (RepeatMasker web server) 的软件扫描来确认引物中不包括重复序列, 或在设计引物之前, 使用 RepeatMasker 软件来屏蔽掉末端序列中的重复序列, 从而保证实验中 PCR 产物的准确性和 BAC 克隆重叠的可靠性。而对于不能设计引物的末端序列, 本文则首先对该克隆进行随机的鸟枪法测序后, 再据所获得的靠近 BAC 末端的非重复序列来设计引物 (表 5-1)。

#### 1.3 主要仪器设备

LI-COR 4200 DNA 测序仪、CHEF DR-III 脉冲场电泳系统 (Bio-Rad)、5417R 高速冷冻离心机 (Eppendorf)、5810R 高速冷冻离心机 (Eppendorf)、BS 200S 电子天平 (Sartorius)、PTC-200 PCR 扩增仪 (MJ Research 公司)、PTC-220 PCR 扩增仪 (MJ Research 公司)、Ultra Freeze 超低温冰箱 (Heto)、Power PAC 300 电泳仪 (Bio-Rad)、Sub-cell GT 电泳仪 (Bio-Rad)、Gel Doc2000 凝胶成像系统

(Bio-Rad)、SBD-50 Bio 恒温水浴摇床(Heto)、恒温混匀器 Thermomixer comfort 5355 (Eppendorf)、Venticell 111R 烘箱(MMM)、Incucell 111R 培养箱(MMM)、3310 PH Meter (Jenway)、SS-325 灭菌柜 (Tomy)；DNA mini 真空干燥系统 (Heto)；SmartSpec 3000 紫外分光光度计 (Bio-Rad)；Purelab uhq 去离子纯水仪(ELGA)；Vortex-2 genie 旋涡混匀器；真空负压装置。

## 1.4 主要试剂和耗材

LI-COR 4200 测序试剂盒 SequiTherm EXCEL II DNA Sequencing Kit-LC、long ranger gel solution 购自 Epicentre 公司；Ex-Taq DNA polymerase、dNTP、HindIII、pMD18-T 载体和 DH5a 大肠杆菌感受态细胞，购自宝生物工程（大连）有限公司；PEG8000、IPTG、X-gal、NotI、矿物油 (mineral oil) 及其它常规试剂等，购自上海生工生物工程公司；日常质粒快速抽提试剂盒、与琼脂糖凝胶回收试剂盒购自 Axygen 公司；超纯质粒抽提试剂盒 Large-construct kit 购自 Qiagen 公司；PCR 引物由上海生工生物工程公司合成；PCR 产物测序由本实验室 LI-COR 4200 测序仪自行完成；BAC 末端测序由华大基因（杭州）公司完成。

## 2. 方法及步骤

### 2.1 BAC 文库的筛选

基于大熊猫 MHC II 类基因引物和 BAC 末端引物自大熊猫基因组 BAC 文库中筛选阳性克隆的方法和流程，参见上述第三章 2.8.4 的描述。其中，筛选的阳性 BAC 克隆经划线接种，挑取单个 BAC 克隆进行培养后，再经 PCR 扩增验证，排除假阳性克隆。

### 2.2 PCR 产物转化

为避免假阳性克隆的产生，对每个筛选出来的阳性克隆进行 PCR 产物测序，以验证 BAC 克隆的可靠性：

- (1) 将 PCR 产物置 1.5% 琼脂糖凝胶中电泳，割胶回收目的片段（具体操作参见 Axygen 公司凝胶回收试剂盒说明书）；回收 PCR 产物溶于 25 $\mu$ l 无菌去离子水中备用；

- (2) 取 5 $\mu$ l 回收的 PCR 产物，加入 0.5 $\mu$ l pMD18-T 载体和 5 $\mu$ l Ligation I 轻轻混匀后，置 16 $^{\circ}$ C 恒温连接仪中连接 1.5-2h；
- (3) 取出连接反应离心管，加入 30 $\mu$ l DH5a 大肠杆菌感受态细胞，置冰上保持 30min；
- (4) 将离心管从冰上取出后立即置于 42 $^{\circ}$ C 恒温混匀器中热击 90s，并取出置冰上保持 1-2min；
- (5) 加入 1ml 37 $^{\circ}$ C 预热的 SOC 溶液混匀；37 $^{\circ}$ C、220rpm 震荡培养 60min；
- (6) 4000rpm 离心沉淀转化后的大肠杆菌细胞，去上清；
- (7) 沉淀混匀后加到 LB+氨卞青霉素（1mg/ml）平板上涂抹均匀，37 $^{\circ}$ C 恒温培养箱中培养过夜；
- (8) 挑取白色菌落置于 1ml LB 培养液中，并于 37 $^{\circ}$ C、280rpm 剧烈震荡培养 6-8h；
- (9) 取 1-1.5 $\mu$ l 菌液进行 PCR 反应，验证插入片段是否正确；
- (10) 取验证好的菌液 300 $\mu$ l，加入 3-5ml LB 溶液中于 37 $^{\circ}$ C、280rpm 剧烈震荡培养过夜。

## 2.3 PCR 产物测序

### 2.3.1 测序 PCR 反应制备

- (1) Axygen 超纯质粒提取试剂盒自 3-5ml 过夜培养的转化菌液中提取超纯质粒（具体操作参见 Axygen 公司超纯质粒提取试剂盒说明书），并将其溶解于 30 $\mu$ l 无菌去离子水中备用；
- (2) 按下列配方制备测序 PCR 反应：

3.5 $\times$ SequiTherm EXCEL II sequencing buffer	3.6 $\mu$ l
SequiTherm EXCEL II DNA polymerase	0.8 $\mu$ l
M13 forward IRD-700 tagged primer	0.8 $\mu$ l
(OR M13 reverse IRD-800 tagged primer)	(0.8 $\mu$ l)
Deionized water	1.2 $\mu$ l
Ultrapured plasmid DNA	4.0 $\mu$ l
10.4 $\mu$ l	

测序引物序列如下：

表 5-1 大熊猫 MHC II 类基因组 BAC 克隆物理图谱构建之引物表

Table 5-1: Primers of MHC class II gene locus used in this study to construct the BAC contig covering the giant panda MHC class II region

引物名称	正向引物(5'-3')	反向引物(5'-3')	Tm 值	产物大小	引物来源
BTL2	GGGAAGATGCTCTACTCA	GATGTCGAAATGTTGTTT	50.0°C	223 bp	GenBank:AY152836.1 GenBank:AJ630362
DRA	AGACAAGTTCTCGCCACC	CTCCACCTTGCAATCATA	57.8°C	172 bp	GenBank:AY973816
DRB	TTCACCAACGGGACGGAG	GCTGCTGCTCCATGAAGTC	55.0°C	166 bp	GenBank:AF190132 GenBank:AF061038
DQA	CTCACCCCTACTGACCAGC	TAAAGAGGCAGAGCATT	56.0°C	297 bp	GenBank:AJ630363.1 GenBank:AY375906.1
DOB	GGGACAGTGATCCAGAG	AGGAGTCAITTCAGCATC	53.8°C	202 bp	GenBank:AY973815
LMP2	CCCCTAGAGCCATCCCG	CAGCCCTTTGCCATCGGT	53.4°C	700 bp	GenBank:AY152832.1 GenBank:AJ630364.1
DMA	GAAGCCCTGGAGTTTG	GGCCTGGAAGCTGAGTTTCAT	54.4°C	157 bp	GenBank:AY973813
DMB	GCTTCTATCCAGCTGATG	GGTGTAAAGTGTCCCCCGTA	51.5°C	156 bp	GenBank:AY973814
COL11A2	CTCAAACACTTCTTCATCCAG	GACCAGGTGGCACCCGTGT	53.8°C	169 bp	GenBank:AY957499.1 GenBank:AJ630366.1
DAXX	TCGCTCCTGTAACTGTGATG	CTCTTTGGGGGCTTTGT	55.7°C	279 bp	GenBank:AJ630366.1 GenBank:AY152827.1
692B2-F7R7	TCTATCACTACTCGGCTGGAA	CAATCTGGGCATAAGAAAACAC	51.9°C	265bp	Shotgun sequence
1262B6-F1R1	CCTCCACATTTCTCTGCTT	GCCATGGGTATAGTTACCT	50.3°C	273 bp	Shotgun sequence

表 5-2 大熊猫 MHC II 类基因组 BAC 克隆物理图谱构建之 BAC 末端引物表

Table 5-2: Primers based on the BAC end sequence used in this study to construct the BAC contig covering the giant panda MHC class II region

826G2-RP2	TGTGACCCGAGTTGGAGTA	GAAAGAGAGATGGGGATG	48.7°C	442 bp	End sequence
206G5-T7	AGAGGTCACGACCAGCAC	TTTGGAAAGGTTGTTTATCAC	49.0°C	429 bp	End sequence
900H8-T7	GAAAGTGAGGGTAGAGGAC	AAGCTGCTGTTAGAAAATG	47.7°C	249 bp	End sequence
692B2-T7	GTTGAGCCTGAGGGAAGC	AGGGTGGTCTATCACTACTCG	51.4°C	357 bp	End sequence
186A7-T7	GCCTCTTGTCCAGTTCCCT	GGTCTCCCTCTGTTGCTC	51.4°C	198 bp	End sequence
P1P2	GTCAATGGAAAATGGGTGCT	CAAAGACTGGGCCCAGAACTT	49.0°C	569 bp	End sequences of 206G5 and 900H8
P3P4	CAGCCCTTTGCCATCGGT	AAGCCTGAGATCCAGAAAG	53.0°C	1615 bp	LMP2 gene and end sequence of 900H8

M13 forward IRD-700 tagged primer: 5'-CAC GAC GTT GTA AAA CGA C--3' (稀释后浓度为 1.0 nmol) ;

M13 reverse IRD-800 tagged primer sequence: 5'-GGA TAA CAA TTT CAC ACA GG-3' (稀释后浓度为 1.0 nmol) ;

- (3) 将以上反应混合物按 2.5 $\mu$ l/管的规格分成 4 管, 并置冰上保持备用;
- (4) 在 4 个反应管中分别加入 1.2 $\mu$ l SequiTherm EXCEL II-LC termination mix A、C、G 和 T, 并混合均匀后加入 2 $\mu$ l 无菌矿物油覆盖;
- (5) 反应管置 PCR 仪中进行如下程序的反应: 94 $^{\circ}$ C 4min, 92 $^{\circ}$ C 30s, 51 $^{\circ}$ C 20s, 70 $^{\circ}$ C 1min, 32 个循环;
- (6) 反应结束后加入 1-1.5 $\mu$ l stop/loading buffer 于 4 $^{\circ}$ C 或 -20 $^{\circ}$ C 避光保存备用。

### 2.3.2 测序凝胶的制备

- (1) 清洗 25cm 玻璃板, 按 Li-COR 公司的测序说明书装配备用;
- (2) 按下列配方制备 6% 的测序凝胶:

Long ranger gel solution	3.0 ml
10 $\times$ TBE	2.5 ml
Urea	9.0 g
1% APS	120 $\mu$ l
TEMED	12.0 $\mu$ l
Deionized water	Up to 25 ml

- (3) 制备 0.2mm 厚度的测序胶于室温保持 1.5-2h;
- (4) 按说明书装配测序仪, 进行 10-20min 预电泳。

### 2.3.3 测序 PCR 的产物变性、加样及序列测定

- (1) 将测序 PCR 产物置于 PCR 仪中于 93 $^{\circ}$ C 变性 3-5min 取出, 并立即置于冰上防止复性;
- (2) 用测序加样针取 1.0 $\mu$ l 变性 PCR 产物加样, 并进行 2-3h 的测序反应;
- (3) 测序结束后点击测序程序、读取序列并输出序列。

## 2.4 BAC 克隆随机鸟枪法测序 (shotgun sequencing)

### 2.4.1 单个 BAC 克隆的分离、诱导及培养

- (1) 将含有阳性 BAC 克隆的 96 孔培养板于常温解冻后，用接种环或无菌牙签蘸取少量含 BAC 克隆的培养液，轻轻在含 12.5ug/ml 氯霉素的 LB 培养基表面划线。
- (2) 划线平板置 37°C 培养箱中培养 16-18h;
- (3) 无菌牙签挑取的单个 BAC 克隆加入 1ml LB+12.5ug/ml 氯霉素培养液后于 37°C，280rpm 震荡培养 5-6h;
- (4) 加入 1ul Copycontrol induction 诱导液继续于 37°C，280rpm 震荡培养 1-2h;
- (5) 将 1ml 诱导后的 BAC 菌液加入到 500ml LB+12.5ug/ml 氯霉素培养液的三角烧瓶中于 37°C，280rpm 震荡培养 16-18h。

### 2.4.2 超纯 BAC DNA 的提取

- (1) 培养过夜的 BAC 菌液离心后，用 Qiagen 公司的 Large-construct kit 提取超纯 BAC DNA (具体操作参见厂家说明书);
- (2) 将提取的超纯 BAC DNA 用 2ml TE 溶液溶解后于 -20°C 保存备用;
- (3) 取 5ul DNA 溶液于 0.8% 的琼脂糖凝胶中电泳 (标准浓度的 DNA marker 估计 BAC DNA 的浓度)。

### 2.4.3 Shotgun 文库构建

- (1) 取 300 $\mu$ l 超纯 BAC DNA 溶液，加入 60 $\mu$ l 10 $\times$ MB buffer 和 240 $\mu$ l 去离子水混匀; 细胞粉碎机随机断裂 BAC DNA 为 1.5-3.5kb 的 DNA 片段 (功率 50W、工作 1s 间隔 1s、总时间为 1.5min);
- (2) 随机断裂的 DNA 片段加入 40 $\mu$ l 5M NaCl 溶液和 2 倍体积的无水乙醇后于 12000rpm、4°C 离心 20min;
- (3) 弃上清液，加入 500 $\mu$ l 75% 酒精后于 12000rpm、4°C 离心 5min;
- (4) 弃上清液，将离心管置真空干燥仪中干燥 10min，取出加入 25 $\mu$ l TE 溶解备用;
- (5) 在 25 $\mu$ l 随机断裂的 DNA 溶液中加入 1.5 $\mu$ l Mung Bean Nuclease 和 3 $\mu$ l 10 $\times$ buffer，置于 30°C 水浴中 15min; 取出于 65°C、5min 终止酶活性;
- (6) 将上述末端修饰的 DNA 加入 1% 琼脂糖凝胶中，3V/cm、30min 切割

1.5-3.0kb 范围的 DNA 片段（凝胶回收程序参见产品说明书）；凝胶回收的 DNA 溶液用 25 $\mu$ l 去离子水洗脱后备用；

- (7) 连接反应：加入 7 $\mu$ l 凝胶回收 DNA 溶液、1 $\mu$ l pUC19-*Sma*I 载体、1 $\mu$ l 10 $\times$ buffer 和 1 $\mu$ l T4 连接酶（15U/ $\mu$ l）于 16 $^{\circ}$ C 连接过夜；
- (8) 加入 30 $\mu$ l DH5a 大肠杆菌感受态细胞置冰上 30min，42 $^{\circ}$ C 热击 90s 后置冰上 1-2min；加入 1ml 37 $^{\circ}$ C 预热的 SOC 溶液于 37 $^{\circ}$ C、220rpm 震荡培养 60min；
- (9) 4000rpm 离心 1min 去除多余的上清液，将转化好的菌液均匀涂布于 LB+氨卞青霉素+IPTG+X-gal 平板上并 37 $^{\circ}$ C 培养过夜；
- (10) 挑取白色单个菌落置 96 孔板中，加入 1ml LB+氨卞青霉素培养液于 37 $^{\circ}$ C、280rpm 震荡培养过夜；
- (11) 将 96 孔平板置于 4 $^{\circ}$ C 保存数天以备用。

#### 2.4.4 超纯质粒 DNA 的大量制备

第一种方法是利用 Axygen 公司的 AxyPrep-96 质粒 DNA 制备试剂盒，提取大量的超纯质粒 DNA（具体操作见说明手册）。

第二种方法亦可使用 Axygen 公司的日常质粒快速提取试剂盒进行。

#### 2.4.5 质粒 DNA 的测序

质粒 DNA 的测序方法与 PCR 产物的测序方法相同。但这里需将测序胶的浓度调整为 3.75%，并使用 66cm 测序板进行单向或双向测序（测序时间 14h）。测序完毕后，选择 Q>20 阅读长 600-800bp 的序列。

### 2.5 BAC 末端测序

单个 BAC 克隆的分离和培养方法见 2.4.1（Axygen 公司的日常质粒提取试剂盒提取 BAC DNA）。

测序反应由杭州华大基因公司 ABI377 DNA 测序仪完成。其末端测序引物为 T7（5'-TAA TAC GAC TCA CTA TAG-3'）和 pCC1BAC 载体来源的 pCC1/pEpiFOS RP-2 反向测序引物（5'-TAC GCC AAG CTA TTT AGG TGA GA-3'）。

## 2.6 序列分析

将所有的测序序列按 FASTA 格式输入 VecScreen 在线载体去除程序 (<http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html>) 以去除载体序列和污染的细菌基因组序列。对于需要设计引物的序列，测序序列都经 RepeatMasker (RepeatMasker web server) 软件扫描来确认该序列不包括重复序列 (其它序列则与 GenBank 中的序列进行 BLASTN 比对)。

## 2.7 大熊猫 MHC II 类基因组物理图谱 (BAC contig) 组装

以人、家猫和狗等的 MHC II 类结构特征作为参考背景，并结合本文筛选的 II 类 MHC 基因引物的位置和 BAC 克隆的末端位置等，确定了各个 BAC 克隆之间的相互关系和位置，并结合限制性酶切指纹图谱分析，得到准确的、最少 BAC 克隆覆盖 (tiling path) 的大熊猫 MHC II 类基因组物理图谱

## 2.8 限制性酶切指纹图谱 (DNA Fingerprinting)

### 2.8.1 超纯 BAC DNA 的大量制备

按照 2.4.2 的方法提取超纯 BAC DNA 后，调节 BAC DNA 的浓度至 80-100ng/ul 于 -20℃ 冷冻保存备用。

### 2.8.2 限制性内切酶消化及电泳分析

- (1) 50ul 超纯 BAC DNA 分别各加入 4ul 的 6.4ul 10×M buffer、*Hind*III (15U/ul) 和 *Eco*RI (15U/ul) 后混匀，置 37℃ 恒温仪中酶切消化 10-12h，使 BAC DNA 被完全酶切；
- (2) 取 2ul 酶切产物于 1% 琼脂糖凝胶中电泳 (110V、30min)；
- (3) EBTr 染色 20min 后于凝胶成像系统检测酶切消化的程度；
- (4) 于 30cm 长制胶板中制备 0.8% 的琼脂糖凝胶；
- (5) 酶切产物于真空干燥仪中干燥 20min，以便将酶切产物浓缩至 35-40ul 左右；加入 5ul 10×loading buffer 后混匀；
- (6) 将制备好的 0.8% 琼脂糖凝胶放入 4℃ 预冷的 1% TAE 电泳缓冲液中，2.5V/cm 电泳 16-18h；
- (7) 电泳完毕的 0.8% 琼脂糖凝胶用 EBTr 染色 20-30min 后用清水漂洗 15-20min；

(8) 凝胶成像系统检测、拍照和分析。

## 第二节 结果和讨论

本文据人、大熊猫、猫和狗的同源 MHC II 类基因 (the MHC sequencing consortium,1999; Debenham et al.,2005; Yuhki et al., 2003) 所设计的引物 BTL2、DRA、DRB1、DQA、DOB、LMP2、DMA、DMB、COL11A2 和 DAXX 都在大熊猫的基因组中获得了成功扩增。其产物经与 GenBank 数据库中的相关序列进行 BLASTN 比对后,发现与人、狗和猫的 MHC II 类基因具有很高的同源性 (表 5-3),表明大熊猫 MHC II 类分子的基因或结构与这些物种之间存在共线性关系,对于构建大熊猫准确的物理图谱具有重要的参考作用。这些大熊猫特异性的 MHC 序列已提交 GenBank 数据库,登录号(Accession number)为 EF125957~66。

研究结果表明:大熊猫 MHC II 类基因组 (自 DAXX 到 BTL2 基因之间) 的 BAC 克隆重叠群,覆盖了大熊猫基因组约 650kb 的范围 (图 5-3)。

表 5-3: 大熊猫 II 类 MHC 基因的位点信息

Table 5-3: Nucleotide similarity between the coding sequences of giant panda MHC class II and corresponding counterparts of canis, feline and human MHC.

位点 (Locus)	序列特征 (Characteristic of sequence)	同源性比较 (Percentage Similarity)
BTL2	Exon 2 and partial intron	狗 (92%)、猫 (82%)、人 (82%)
DRA	Partial exon 3	狗 (92%)、猫 (92%)、人 (91%)
DRB1	Partial exon 2	狗 (88%)
DQA	Partial exon 2	狗 (83%)、人 (89%)
DOB	Partial exon 3	狗 (95%)、猫 (94%)、人 (88%)
LMP2	Partial exon 4 and 5	狗 (87%)、猫 (85%)、人 (89%)
DMB	Partial exon 3	狗 (91%)、猫 (97%)、人 (99%)
DMA	Partial exon 3	狗 (95%)、猫 (92%)、人 (87%)
COL11A2	exon 4	狗 (96%)、猫 (97%)、人 (94%)
DAXX	Partial exon 6	狗 (95%)、猫 (93%)、人 (92%)

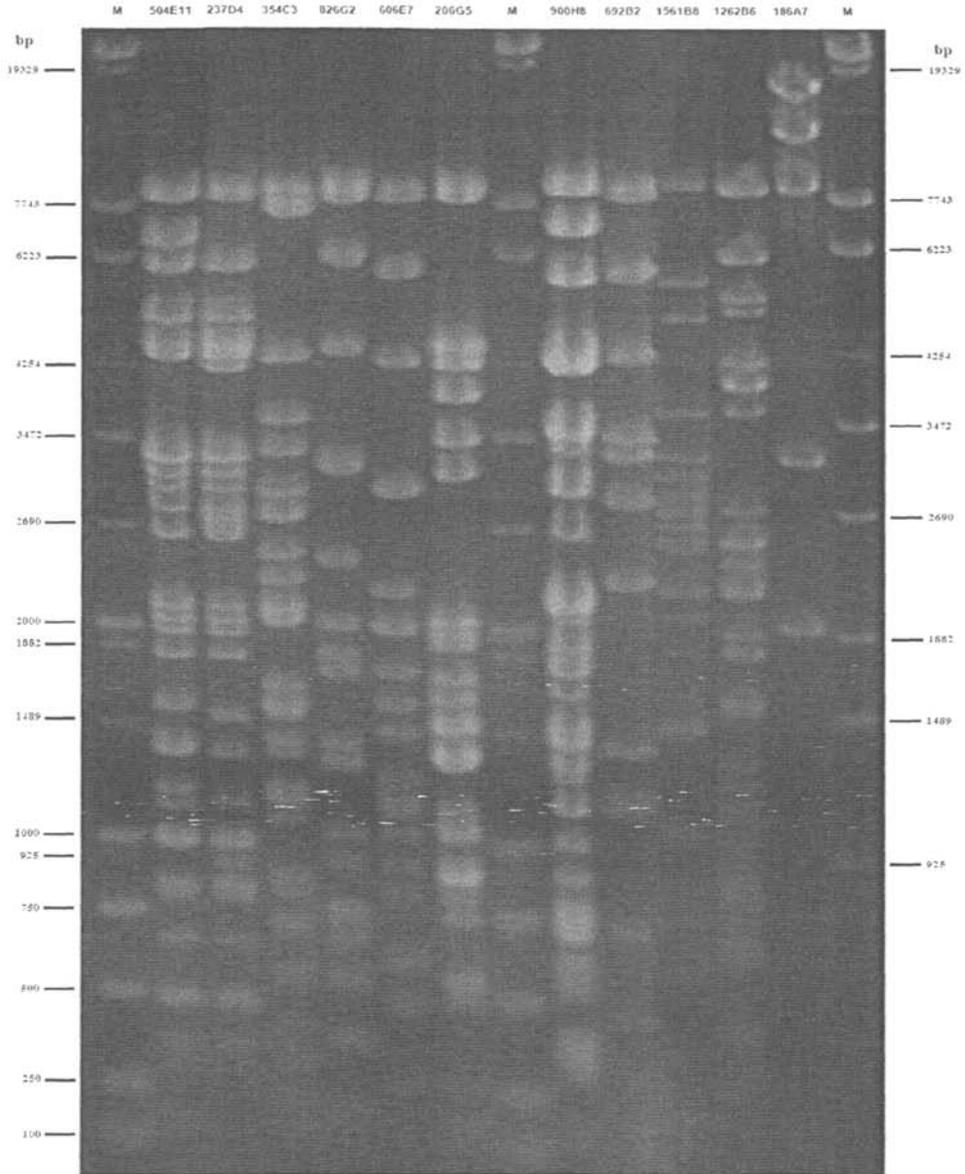


图 5-1: 大熊猫 MHC II 类基因组 BAC 克隆的限制性酶切示意图。

Figure 5-1. A gel image of DNA fingerprints of part of MHC class II positive BAC clones. The positive BAC clones were extracted by Qiagen large-construct kit and digested with *Hind* III and *Eco*R I, followed by separation on 1.0% agarose gel with 2.0 V/cm for 18 hours. M is mixture of  $\lambda$ -*Eco*T14 I digest and DL2000 DNA marker. The sizes in base pairs of markers are indicated on left side.

在已知的人和一些动物的 MHC II 类基因组中，DR 和 DQ 位点均具有高度的序列重组和重复等特征。在人的 MHC 中有 9 个不同的 DRB 基因，且在不同

的单倍型中数量也有所不同 (The MHC sequencing consortium, 1999)。猫有由基因重复和反转所造成的 4 个 DRB 基因 (Yuhki et al., 2003)。狗则仅有一个功能性的 DRB1 基因和 DRB2 假基因 (Debenham et al., 2005)。在 DQ 区域上, 猫的 MHC II 类完全缺失了整个 DQ 区域 (Yuhki et al., 2003)。但在本研究中, 我们成功的扩增到了大熊猫的 DQA 和 DRB1 位点, 表明了大熊猫在 DQ 和 DR 位点上不存在功能基因的丢失问题 (Wan et al., 2005; Zhu et al., 2007)。然而, 在构建大熊猫 MHC II 类基因组的精确物理图谱时, 由于 DQA 和 DRB 基因之间的序列差异很小, 所以在实验中我们需要设计大熊猫 DQA 和 DRB 之 PCR 扩增的特异性引物, 即只能扩增出 DQA 与 DRB 基因的专一性引物。

在大熊猫 BAC 文库质量检测过程中, DRB 和 DQA 引物分别得到了 12 和 16 个阳性的 BAC 克隆。对这些阳性 BAC 克隆的 PCR 产物测序, 结果仅得到 DRB 和 DQA 的 2 个等位基因, 而且这两个等位基因不是新的等位基因, 其序列也与本实验室以前发表的论文中 DRB 和 DQA 的等位基因是重复的。这与以前描述的 DRB 和 DQA 引物只能扩增大熊猫 DRB 和 DQA 的一个位点结果是一致的 (Wan et al., 2005; Zhu et al., 2007)。值得注意的是, 由于本论文所用 DRB 和 DQA 引物只能扩增其相应的 1 个 DRB 和 DQA 位点, 因此不能判断大熊猫 MHC 仅包括 1 个 DRB 和 DQA 位点。事实上, 根据本实验室先前的实验结果, DQA 和 DRB 可能不止 1 个位点, 采用 DRB 和 DQA 第三外显子设计的引物扩增序列结果表明含有 2 个以上的不同序列。因此为了避免在物理图谱构建过程中 BAC 克隆之间造成混乱, 本论文即采用 DRB 和 DQA 的第二外显子设计大熊猫特异性的 PCR 引物, 这些 PCR 引物的结果在本实验室近期发表的论文中亦以描述 (Wan et al., 2005; Zhu et al., 2007)。

在本研究中, 我们经过 BAC 末端测序发现 206G5-RP2、692B2-RP2、1262B6-T7 端都发现了不同程度的重复序列 (SINE、LINE 等)。对于这些克隆, 我们采取随机鸟枪法测序策略, 获得了不同的序列, 经与狗 (*Canis familiar*) 和猫 (*Felis catus*) 的 MHC II 类分子进行同源性比较, 并经 RepeatMasker 程序屏蔽其中的分散重复和简单重复序列, 设计出合适的大熊猫种属特异性引物, 进一步筛选大熊猫基因组 BAC 文库, 构建准确的 BAC contig。因此, 本研究基于 MHC II 类基因特异性引物、末端序列以及限制性酶切 fingerprint 图谱 (图 5-1) 所构建的大熊猫 MHC II 类分子物理图谱是准确可靠的。

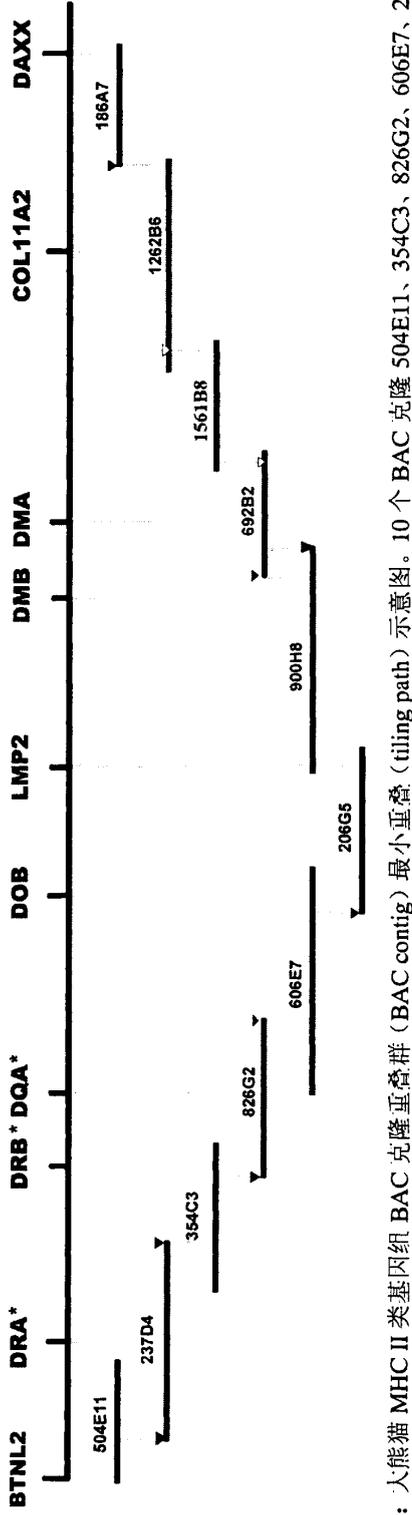


图 5-3: 大熊猫 MHC II 类基因组 BAC 克隆重叠群 (BAC contig) 最小重叠 (tiling path) 示意图。10 个 BAC 克隆 504E11、354C3、826G2、606E7、206G5、900H8、692B2、1561B8 和 186A7 覆盖了大熊猫 MHC II 类基因组中从 BTL2 位点到 DAXX 位点的全部基因组区域 (区域长度约 650kb)。4 个 MHC 基因同源引物位点 (BTL2、LMP2、COL11A2、DAXX) 和 6 个大熊猫引物位点 (DRA、DRB、DQA、DOB、DMB、DMA) 作为标记位点分别锚着于不同的 BAC 克隆上 (黑三角形代表末端引物位点, 白色三角形代表随机鸟枪法测序的引物位点, 星号标记为单一位点引物)。

Figure 5-3: The sequencing ready tiling path of BAC clones MHC class II region. Ten BAC clones 504E11, 354C3, 826G2, 606E7, 206G5, 900H8, 692B2, 1561B8, 1262B6 and 186A7 covered all the region of MHC class II from BTL2 to DAXX gene. The BAC contig map of MHC class II region of giant panda is spanning about 650kb. Four primers (BTL2, LMP2, COL11A2, DAXX) based on MHC homologous counterpart and six species-specific primers (DRA, DRB, DQA, DOB, DMB, DMA) were anchored on the different BAC in the BAC contig map of MHC class II of giant panda. End-sequence primers marked with black triangle, shotgun sequences marked with blank triangle were anchored to the contig and single locus-specific primers marked with star.

此外，与人和其它动物（狗和家猫等）的 MHC II 类相比较，发现大熊猫的 MHC II 类从 BTL2 基因到 DAXX 基因是与其存在着共线性保守（Synteny conservation）关系的（整个区域的结构都比较保守）。但通过扩增，本文未获得大熊猫的 DPA 和 DPB 基因，其原因是大熊猫原本就缺乏这 2 个基因，还是由于“各物种在该区域表现出极大的差异（Kumánovics et al.,2003; Debenham et al.,2005）”而致使引物不适宜进行扩增，则尚待进一步的研究。但是，从 1262B6 的随机鸟枪法测序的序列中我们发现了与人 HLA DP alpha exon 1 序列同源的序列片段。但是这段同源序列与狗 MHC DPA 序列却没有任何相似性。因此，我们推断大熊猫 DP 区域仍然存在，只是序列的变异较大。

在本文的研究中，我们也未能扩增到大熊猫的 DOA 基因，但得到了 DOB 基因。这一结果似乎有悖于“哺乳动物的 II 类 MHC 几乎都存在功能性的 DOA 和 DOB 基因”的预测。在此，本文推测其结果的原因，要么是大熊猫的 DOA 发生了较大变异，从而导致正常扩增的失败；要么是大熊猫的 MHC II 类出现了 DOA 的缺失。但孰是孰非，待后续 II 类 MHC 的基因组计划完成之后就迎刃而解了。

本文的研究表明，大熊猫的 MHC II 类基因组约 650kb 左右，它比家猫（758kb）（Yuhki et al., 2003）、狗（710kb）（Debenham et al., 2005）和人（998kb）（the MHC sequencing consortium, 1999）的 MHC II 类要小得多，但比小鼠（481kb）和鸡（92kb）却又大得多（Kaufman et al., 1999）。由此表明了在大熊猫的 MHC II 类基因组内部，存在着大量的基因重组和种属特异性复制（species-specific duplications）等。

### 第三节 结论

1. 本文基于同源 MHC 序列（包括大熊猫的相关数据），设计了自大熊猫的基因组 BAC 文库中筛选包含大熊猫 MHC II 类基因阳性 BAC 克隆的引物；
2. 成功筛选出含有大熊猫 MHC II 类基因的 BAC 克隆，并据此通过对这些 BAC 克隆的末端测序结果，设计了末端引物；
3. 利用上述引物，进一步自大熊猫的基因组 BAC 文库中，筛选到了覆盖整

个大熊猫 MHC II 类基因组的 BAC 克隆, 并构建了从 BTL2 基因到 DAXX 基因之全部区域的物理图谱;

4. 运用限制性酶切指纹图谱的方法, 验证了所构建的大熊猫 MHC II 类基因组的物理图谱, 是准确可靠的;
5. 据 BAC 克隆间的相互关系, 确定了所构建的 BAC 克隆重叠群 (contig) 是最小的重叠距离 (tiling path) 图谱。大熊猫 MHC II 类基因组的长度约为 650kb, 可直接用于基因组的测序计划。

## 第六章 主要成果及创新点

### 一 主要成果

1. 构建了大熊猫的基因组细菌人工染色体(BAC)文库。文库包含的 205,800 个 BAC 克隆, 保存于 2,100 块 96 孔培养板上。经对随机挑取的 174 个 BAC 克隆进行脉冲电泳检测, 计算出文库的平均插入片段为 97kb, 覆盖大熊猫基因组约 6.8 倍(除去 5.74%的空载体)。同时, 从文库中获得任一单拷贝或多拷贝大熊猫基因的概率为 99.93%。
2. 运用 4D-PCR 文库筛选技术, 建立了大熊猫基因组 BAC 文库的高效 PCR 筛选文库。该文库包括 43 个超级库和 1462 (43×34) 个 1D、2D、3D、4D 超级亚库。任一引物或探针只需 77 个 PCR 反应, 即可从 BAC 文库中筛选到所需的阳性 BAC 克隆。
3. 以大熊猫皮肤成纤维细胞为材料, 运用染色体荧光原位杂交(FISH)技术, 成功地将大熊猫的 MHC 基因组定位于第 9 号染色体的长臂上。
4. 运用 BAC 克隆的末端测序方法, 构建了大熊猫 MHC II 类基因组的物理图谱。进一步地, 利用限制性酶切指纹图谱法, 证明了所构建的物理图谱中的 BAC 克隆是有效重叠的, 其 MHC II 类基因组的长度约为 650kb。
- 5.

### 二 创新点

1. 以细菌人工染色体(BAC)为载体, 成功的构建了大熊猫的基因组 BAC 文库, 从而为相关的后续研究, 提供了高质量的基因资源材料平台;
2. 运用 FISH 技术精确的将大熊猫的 MHC 基因组定位于 9 号染色体的长臂上(9q);
3. 构建了覆盖整个大熊猫 MHC II 类基因组的 BAC 克隆重叠群图谱(物理图谱), 为后续开展大熊猫 MHC II 类基因组的测序计划, 奠定了坚实的基础。

## 参考文献

- 方福德, 杨焕明, 张蕃昌等(1998). 分子生物学前沿技术 (第 1 版). 北京: 北京医科大学中国协和医科大学联合出版社, 1998: 19-21
- 于军(1998). 人类基因组的物理图谱与大规模 DNA 测序. 遗传, 20 (6): 41-47
- 王宗仁, 贾凤兰, 吴鹤龄(1990). 动物遗传学实验方法. 北京大学出版社, 96-127
- 赵剑华, 王秀琴, 刘芝华, 吴旻(2000). 功能基因组学研究的主要内容和方法. 生物化学与生物物理研究进展, 27 (1): 6-8
- Abi-Rached L, Gilles A, Shiina T, Pontarotti P, Inoko H(2002). Evidence of en bloc duplication in vertebrate genomes. *Nature Genetics*, 31: 100-105
- Abi-Rached L, McDermott MF, Pontarotti P(1999). The MHC Big Bang. *Immunol Rev*, 167: 33-44
- Afanassieff M, Goto RM, Ha J, Sherman MA, Zhong LW, Auffray C, Coudert F, Zoorob R, Miller MM(2001). At least one class I gene in restriction fragment pattern-Y (Rfp-Y), the second MHC gene cluster in the chicken, is transcribed, polymorphic, and shows divergent specialization in antigen binding region. *J Immunol*, 166: 3324-3333
- Akanuma J(2002). Whole genome sequencing of human mitochondrial DNA. *Nippon Rinsho*, 60 Suppl 4: 683-687
- Alfonso C & Karlsson L(2000). Nonclassical MHC class II molecules. *Annu Rev Immunol*, 18:113-142
- Amemiya CT, Ota T, Litman GW(1996). In Birren B & Lai E (eds), *Nonmammalian Genome Analysis: A Practical Approach*. Academic Press, New York, 223-256
- Anderson S, Bankier A, Barrell B, de Bruijn M, Coulson A, Drouin J, Eperon I, Nierlich D, Roe B, Sanger F, Schreier P, Smith A, Staden R & Young I(1981). Sequence and organization of the human mitochondrial genome. *Nature*, 290 (5806), 457-65
- Animal Genome Size Database [<http://www.genomesize.com>]
- Apanius V, Penn D, Slev PR, Ruff LR, Potts WK(1997). The nature of selection on the major histocompatibility complex. *Crit Rev Immunol*, 17(2): 179-224
- Arkush KD, Giese AR, Mendoca HL(2002). Resistance to three pathogens in the

- endangered winter-run Chinook salmon (*Oncorhynchus tshawytscha*): effects of inbreeding and major histocompatibility complex genotypes. *Canadian Journal of Fisheries and Aquatic Science*, 59: 966-975
- Asakawa S, Abe I, Kudoh Y, Kishi N, Wang Y, Kubota R, Kudoh J, Kawasaki K, Minoshima S, Shimizu N(1997). Human BAC library: construction and rapid screening. *Gene*, 191(1): 69-79
- Baer R, Bankier AT, Biggin MD, Deininger PL, Farrell PJ, Gibson TJ, Hatfull G, Hudson GS, Satchwell SC, Seguin C, Tuffnell P & Barrell B(1984). DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature*, 310 (5974): 207-11
- Ballingall K, MacHugh N, Taracha E, Mertens B,McKeever D(2001). Transcription of the unique ruminant class II major histocompatibility complex-DYA and DIB genes in dendritic cells. *Eur J Immunol*, 31: 82-86
- Bartl S, Baish MA, Flajnik MF, Ohta Y(1997). Identification of class I genes in cartilaginous fish, the most ancient group of vertebrates displaying an adaptive immune response. *J Immunol*, 159: 6097-6104
- Bartl S & Weissman IL(1994). Isolation and characterization of major histocompatibility complex class IIB genes from the nurse shark. *Proc Natl Acad Sci USA*, 91: 262-266
- Baysal BE, van Schothorst EM, Farr JE, James MR, Devilee P, Richard CW 3rd(1997). A high-resolution STS, EST, and gene-based physical map of the hereditary paraganglioma region on chromosome 11q23. *Genomics*, 44(2): 214-221
- Beatty BG, Mai S, Squire JA(2002). *FISH (Fluorescence in situ hybridization)*. Oxford University Press.
- Benacerraf B(1981). Role of MHC gene products in immune regulation. *Science*, 212: 1229-1238
- Birren B & Lai E(1993). *Pulsed Field Gel Electrophoresis, A Practical Guide*. Academic Press, Inc
- Birren B, Green ED, Klapholz S, Myers RM, Riethman H, Roskams J(1997). Bacterial Artificial Chromosomes. In *Genome Analysis: A laboratory manual*. Vol. 3 Cloning Systems (ed. B. Birren et al.). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 242-295
- Bjorkman PJ, Parham P(1990). Structure, function and diversity of class I major

- histocompatibility molecules. *Ann Rev Biochem*, 59:253-288
- Blanco G, Nikitopoulou A, Kraus M, Mason RM, Coulton GR, Brown SD(1998). A STS content physical and transcription map across the ky, kyphoscoliosis, nonrecombinant region. *Genomics*, 54(3): 415-423
- Blundell TL & Mizuguchi K(2000). Structural genomics: an overview. *Progress in biophysics & molecular biology*, 73: 289-295
- Bouffard GG, Idol JR, Braden VV, Iyer LM, Cunningham AF, Weintraub LA, Touchman JW, Mohr-Tidwell RM, Peluso DC, Fulton RS, Ueltzen MS, Weissenbach J, Magness CL, Green ED(1997). A physical map of human chromosome 7: an integrated YAC contig map with average STS spacing of 79 kb. *Genome Res*, 7(7): 673-692
- Bouwer HGA, Lindahl KF, Baldrige JR, Wagner CR, Barry RA, Hinrichs DJ(1994). An H2-T MHC class Ib molecule presents *Listeria monocytogenes*-derived antigen to immune CD8+ CTL. *J Immunol*, 152: 5352-5360
- Botstein D, White RL, Skolnick M, Davis RW(1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet*, 32(3): 314-331
- Buitkamp J, Kollers S, Durstewitz G, Welzel K, Schäfer K, Kellermann A, Lehrach H, Fries R(2000). Construction and characterization of a gridded cattle BAC library. *Anim Genet*, 31(6):347-351
- Cai WW, Chen R, Gibbs RA, Bradley A(2001). A clone-array pooled shotgun strategy for sequencing large genomes. *Genome Res*, 11(10): 1619-1623
- Cai L, Taylor JF, Wing RA, Gallagher DS, Woo SS, Davis SK(1995). Construction and characterization of a bovine bacterial artificial chromosome library. *Genomics*, 29: 413-425
- Cai WW, Reneker J, Chow CW, Vaishnav M, Bradley A(1998). An anchored framework BAC map of mouse chromosome 11 assembled using multiplex oligonucleotide hybridization. *Genomics*, 54: 387-397
- Caron H, van Schaik B, van der Mee M, Baas F, Riggins G, van Sluis P, Hermus MC, van Asperen R, Boon K, Voute PA, Heisterkamp S, van Kampen A, Versteeg R(2001). The human transcriptome map: clustering of highly expressed genes in chromosomal domains. *Science*, 291(5507): 1289-1292
- Carrington M, Nelson GW, O'Brien SJ(1999). HLA and HIV-1: heterozygote advantage and B\*35-Cw\*04 disadvantage. *Science*, 283: 1748-1752

- Castro LFC, Furlong RF, Holland PWH(2004). An antecedent of the MHC-linked genomic region in *Amphioxus*. *Immunogenetics*, 55: 782-784
- Chang DC, Chassy BM, Saunders JA, Sowers AE(1992). *Guide to Electroporation and Electrofusion*. Academic Press, San Diego, CA
- Chen R, Sodergren E, Weinstock GM, Gibbs RA(2004). Dynamic building of a BAC clone tiling path for the Rat Genome Sequencing Project. *Genome Res*, 14(4): 679-684
- Cheng YL, Mancino V, Birren B(1995). Transformation of *E.coli* with large DNA molecules by electroporation. *Nucleic Acids Res*, 23(11): 1990-1996
- Cheroutre H, Holcombe HR, Tangri S(1995). Antigen-presenting function of the TL antigen and mouse CD1 molecules. *Immunol Rev*, 147: 31-52
- Clarke L & Carbon J(1976). A colony bank containing synthetic Col E1 hybrid plasmids representative of the entire *E.coli* genome. *Cell*, 9(1): 91-99
- Collins J & Hohn B(1978). Cosmids: a type of plasmid gene cloning vector that is packageable in vitro in bacteriophage lambda heads. *Proc Natl Acad Sci USA*, 75: 4242-4246
- Connelly C, McCormik MK, Shero J, Hiether P(1991). Polyamines eliminate an extreme size bias against transformation of large yeast artificial chromosome DNA. *Genomics*, 10: 10-16
- Coulson A, Sulston J, Brenner S, Karn J(1986). Toward a physical map of the genome of the nematode *Caenorhabditis elegans*. *Proc Natl Acad Sci USA*, 83(20): 7821-7825
- Cresswell P, Turner MJ, Strominger JL(1973). Papain-solubilized HL-A antigens from cultured human lymphocytes contain two peptide fragments. *Proc Natl Acad Sci USA*, 70(5): 1603-1607
- Danchin EGJ, Vitiello V, Vienne A, Richard O, Gouret P, McDermott MF, Pontarotti P (2004). The major histocompatibility complex origin. *Immunol Rev*, 198:216-232
- Danna K & Nathans D(1971). Specific cleavage of simian virus 40 DNA by restriction endonuclease of *Hemophilus influenzae*. *Proc Natl Acad Sci, USA*, 68: 2913-2917
- Dausset J(1959). Antiplatelet auto-antibodies & immunological tests. *Rev Prat*, 9(19):2087-2088
- Davis MW & Hammarlund M(2006). Single-nucleotide polymorphism mapping. *Methods Mol Biol*, 351: 75-92

- Debenham SL, Hart EA, Ashurst JL, Howe KL, Quail MA, Ollier WE, Binns MM (2005). Genomic sequence of the class II region of the canine MHC: comparison with the MHC of other mammalian species. *Genomics*, 85(1):48-59
- De Jong PJ, Yokobata K, Chen C, Lohman F, Pederson L, McNinch J, Van Dilla M (1989). Human chromosome-specific partial digest libraries in  $\lambda$  and cosmid vectors. *Cytogenet Cell Genet*, 51: 985
- Delarbre C, Jaulin C, Kourilsky P, Gachelin G(1992). Evolution of the MHC: A hundred-fold amplification of the MHC genes in the African pygmy mouse *Nannomys setulosus*. *Immunogenetics*, 37: 29-38
- Dey A, Thornton AM, Lonergan M, Weissman SM, Chamberlain JW, Ozato K(1992). Occupancy of upstream regulatory sites in vivo coincides with MHC class I gene expression in mouse tissues. *Mol Cell Biol*, 12: 3590–3599
- Di Meo GP, Perucatti A, Floriot S, Hayes H, Schibler L, Rullo R, Incarnato D, Ferretti L, Cockett N, Cribsu E, Williams JL, Eggen A, Iannuzzi L(2007). An advanced sheep (*Ovis aries*,  $2n = 54$ ) cytogenetic map and assignment of 88 new autosomal loci by fluorescence in situ hybridization and R-banding. *Anim Genet*, 38(3): 233-240
- Draghici S, Khatri P, Eklund AC, Szallasi Z(2006). Reliability and reproducibility issues in DNA microarray measurements. *Trends Genet*, 22(2): 101-109
- Dubovsky J, Sheffield VC, Duyk GM, Weber JL(1995). Sets of short tandem repeat polymorphisms for efficient linkage screening of the human genome. *Hum Mol Genet*, 4(3): 449-452
- Doherty PC & Zinkernagel RM(1975). A biological role for the major histocompatibility antigens. *Lancet*, 7922: 1406-1409
- Donaldson WL, Oriol JG, Plavin A, Antczak DF(1992). Developmental regulation of class I major histocompatibility complex antigen expression by equine trophoblastic cells. *Differentiation*, 52: 69-78
- Edwards S & Hedrick PW(1998). Evolution and ecology of MHC molecules: from genomics to sexual selection. *Trends Ecol Evol*, 13: 305-311
- Falus A, Wakeland EK, McConnell TJ, Gitlin J, Whitehead AS, Colten HR(1987). DNA polymorphism of MHC III genes in inbred and wild mouse strains. *Immunogenetics*, 25(5): 290-298
- Ferreira AR, Foutz KR, Keim P(2000). Soybean genetic map of RAPD markers assigned to an existing scaffold RFLP map. *J Hered*, 91(5): 392-396.

- Figuroa F, Gunther E, Klein J(1988). MHC polymorphism predating speciation. *Nature*, 355: 265-267
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC(1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, 391(6669): 806-811
- Flajnik MF, Ohta Y, Namikawa-Yamada C, Nonaka M(1999). Insight into the primordial MHC from studies in ectothermic vertebrates. *Immunol Rev*, 167: 59-67
- Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM(1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*, 269(5223): 496-512
- Foster J(1995). The *Plasmodium falciparum* genome project: a resource for researchers. *Parasitol Today*, 11: 1-4
- Frazer KA, Elnitski L, Church DM, Dubchak I, Hardison RC(2003). Cross-species sequence comparisons: a review of methods and available resources. *Genome Res*, 13(1): 1-12
- Frijters ACJ, Zhang Z, Van Damme M, Wang GW, Ronald PC, Michelmore RW (1997). Construction of a bacterial artificial chromosome library containing large *EcoRI* and *HindIII* genomic fragments of lettuce. *Theor Appl Genet*, 94: 390-399
- Garg K, Green P & Nickerson DA(1999). Identification of candidate coding region single nucleotide polymorphisms in 165 human genes using assembled expressed sequence tags. *Genome Research*, 9: 1087-1092
- Gebhardt C, Eberle B, Leonards-Schippers C, Walkemeier B, Salamini F(1995). Isolation, characterization and RFLP linkage mapping of a DNA repeat family of *Solanum spegazzinii* by which chromosome ends can be localized on the genetic map of potato. *Genet Res*, 65(1): 1-10
- Geigl JB, Uhrig S, Speicher MR(2006). Multiplex-fluorescence in situ hybridization for chromosome karyotyping. *Nat Protoc*, 1(3):1172-1184
- Germain RN & Margulies DH(1993). The biochemistry and cell biology of antigen processing and presentation. *Annu Rev Immunol*, 11: 403-450
- Gill CA, Davis SK, Taylor JF, Cockett NE, Bottema CD(1999). Construction and characterization of an ovine bacterial artificial chromosome library. *Mamm*

- Genome, 10(11): 1108-1111
- Gillett W, Hanks L, Wong GK, Yu J, Lim R, Olson MV(1996). Assembly of high-resolution restriction maps based on multiple complete digests of a redundant set of overlapping clones. *Genomics*, 33(3): 389-408
- GOLD: Genome OnLine Database v2.0 (<http://www.genomesonline.org>)
- Goldammer T, Kata SR, Brunner RM, Dorroch U, Sanftleben H, Schwerin M, Womack JE(2002). A comparative radiation hybrid map of bovine chromosome 18 and homologous chromosomes in human and mice. *Proc Natl Acad Sci USA*, 99: 2106–2111
- Gorg A, Obermaier C, Boguth G, Harder A, Scheibe B, Wildgruber R, Weiss W(2000). The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis*, 21(6): 1037-1053
- Gorg A, Weiss W, Dunn MJ(2004). Current two-dimensional electrophoresis technology for proteomics. *Proteomics*, 4(12): 3665-3685
- Green ED, Riethman HC, Dutchik JE, Olson MV(1991). Detection and characterization of chimeric yeast artificial chromosome clones. *Genomics*, 11: 658-669
- Gruen JR & Weissman SM(2001). Human MHC class III and IV genes and disease associations. *Front Biosci*, 6: D960-D972
- Gustafson AL, Tallmadge RL, Ramlachan N, Miller D, Bird H, Antczak DF, Raudsepp T, Chowdhary BP, Skow LC(2003). An ordered BAC contig map of the equine major histocompatibility complex. *Cytogenet Genome Res*, 102(1-4): 189-195
- Hardiman G(2004). Microarray platforms—comparisons and contrasts. *Pharmacogenomics*, 5: 487–502
- Heb M, Goldammer T, Gelhaus A, Ried K, Rappold G, Eggen A, Bishop MD, Schwerin M, Ho stmann RD(1999). Physical assignment of the bovine MHC class IIa and class IIb genes. *Cytogenet Cell Genet*, 85: 244-247
- Hedrick PW(1994). Evolutionary genetics of the major histocompatibility complex. *Am Nat*, 143: 945-964
- Hedrick PW(2002). Pathogen resistance and genetic variation at MHC loci. *Evol*, 56: 1902-1908
- Hedrick PW & Kim TJ(1999). Genetics of complex polymorphisms: parasites and maintenance of MHC variation. In *Evolutionary Genetics from Molecules to*

- Morphology, edited by Singh RS & Krimbas CK. Cambridge University Press, New York.
- Hedrick PW, Parker KM(1998). MHC variation in the endangered *Gila topminnow*. *Evolution*, 52: 194-199
- Hedrick PW & Thomson G(1983). Evidence for balancing selection at HLA. *Genetics*, 104: 449-456
- Heller MJ(2002). DNA microarray technology: devices, systems, and applications. *Annu Rev Biomed Eng*, 4: 129-153
- Helou K, Walter L, Gunther E, Levan G(1998). Cytogenetic orientation of the rat major histocompatibility complex (MHC) on chromosome 20. *Immunogenetics*, 47: 166-169
- Hess M, Goldammer T, Gelhaus A, Ried K, Rappold G(1999). Physical assignment of the bovine MHC class IIa and class IIb genes. *Cytogenet Cell Genet*, 85: 244-247
- Hieter P & Boguski M(1997). Functional genomics: it's all how you read it. *Science*, 278: 601-602
- Hosoda F, Nishimura S, Uchida H, Ohki M(1990). An F factor based cloning system for large DNA fragments. *Nucleic Acids Res*, 18(13): 3863-9
- Huber I, Walter L, Wimmer R, Pasantes JJ, Gunther E, Schempp W(2003) cytogenetic mapping and orientation of the rhesus macaque MHC. *Cytogenet Genome Res*, 103: 144-149
- Hughes AL(1998). Phylogenetic tests of the hypothesis of block duplication of homologous genes on human chromosomes 6, 9, and 1. *Mol. Biol. Evol.* 15, 854-870
- Hughes AL & Nei M(1988). Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature*, 335: 167-179
- Hughes AL & Nei M(1989). Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc Natl Acad Sci USA*, 86: 958-962
- Hughes AL, Yeager M(1998). Natural selection and the evolutionary history of major histocompatibility complex loci. *Front Biosci*, 3:d509-516
- Hunkapiller T, Kaiser RJ, Koop BF, Hood L(1991). Large-scale and automated DNA sequence determination. *Science*, 254(5028): 59-67
- Ioannou PA, Amemiya CT, Garnes J, Kroisel PM, Shizuya H, Chen C, Batzer MA, de

- Jong PJ(1994). A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nat Genet*, 6: 84-89
- Kappes D & Strominger JL(1988). Human class II major histocompatibility complex genes and proteins. *Ann Rev Biochem*, 57: 991-1028
- Kaufman J, Milne S, Gobel TW, Walker BA, Jacob JP, Auffray C, Zoorob R, Beck S (1999). The chicken B locus is a minimal essential major histocompatibility complex. *Nature*, 401(6756): 923-925.
- Kaufman J, Salomonsen J, Flajnik M(1994). Evolutionary conservation of MHC class I and class II molecules—Different yet the same. *Semin Immunol*, 6: 411-424
- Kaufman J, Volk H, Wallny HJ(1995). A "minimal essential Mhc" and an "unrecognized Mhc": two extremes in selection for polymorphism. *Immunol Rev*, 143: 63-88
- Kasahara M, Flajnik MF, Ishibashi T, Natori T(1995). Evolution of the major histocompatibility complex: a current overview. *Transpl Immunol* 3: 1-20
- Khorasani MZ, Hennig S, Imre G, Asakawa S, Palczewski S, Berger A, Hori H, Naruse K, Mitani H, Shima A, Lehrach H, Wittbrodt J, Kondoh H, Shimizu N, Himmelbauer H(2004). A first generation physical map of the medaka genome in BACs essential for positional cloning and clone-by-clone based genomic sequencing. *Mech Dev*, 121(7-8): 903-913
- Kim UJ, Birren BW, Slepak T, Mancino V, Boysen C, Kang HL, Simon MI, Shizuya H(1996). Construction and characterization of a human bacterial artificial chromosome library. *Genomics*, 34: 213-218
- Kim UJ, Shizuya H, De Jong PJ, Birren B, Simon MI(1992). Stable propagation of cosmid sized human DNA inserts in an F factor based vector. *Nucleic Acids Res*, 20: 1083-1085
- Kiss GB, Csanadi G, Kalman K, Kalo P, Okresz L(1993). Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. *Mol Gen Genet*, 238(1-2): 129-137
- Klein J(1984). Gene conversion in MHC genes. *Transplantation*, 38:327-329
- Klein J(1986). Natural history of the major histocompatibility complex. Wiley, New York
- Klein J(1987). Origin of major histocompatibility complex polymorphism: the trans-species hypothesis. *Hum Immunol*, 19(3): 155-162
- Klein J, Figueroa F, Klein D, Sato A, O'hUigin C(1997). Major histocompatibility

- complex genes in the study of fish phylogeny. In: *Molecular Systematics of Fishes* (eds Kocher TD, Stepien CA), Academic Press, San Diego, 271-283
- Klenchin VA, Sukharev SI, Serov SM, Chernomordik LV, Chizmadzhev Yu A(1991). Electrically induced DNA uptake by cells is a fast process involving DNA electrophoresis. *Biophys J*, 4: 804-811
- Klockars T, Savukoski M, Isosomppi J, Laan M, Jarvela I, Petrukhin K, Palotie A, Peltonen L(1996). Efficient construction of a physical map by fiber-FISH of the CLN5 region: refined assignment and long-range contig covering the critical region on 13q22. *Genomics*, 35(1): 71-78
- Kohara Y, Akiyama K, Isono K(1987). The physical map of the whole *E.coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. *Cell*, 50(3): 495-508
- Korenberg JR, Chen XN, Devon KL, Noya D, Oster-Granite ML, Birren BW(1999). Mouse molecular cytogenetic resource: 157 BACs link the chromosomal and genetic maps. *Genome Res*, 9: 514-523
- Kouprina N, Eldarov M, Moyzis R, Resnick M, Larionov V(1994). A model system to assess the integrity of mammalian YACs during transformation and propagation in yeast. *Genomics*, 21: 7-17
- Kulski JK, Shiina T, Anzai T, Kohara S, Inoko H(2002). Comparative genomic analysis of the MHC: the evolution of class I duplication blocks, diversity, and complexity from shark to man. *Immunol Rev*, 190: 95-122
- Kumánovics A, Takada, T and Lindahl KF(2003). Genomic organization of the mammalian MHC. *Annual Review of Immunology*, 21: 629-657
- Laird DJ, De Tomaso AW, Cooper MD, Weissman IL(2000). 50 Million years of chordate evolution: seeking the origins of adaptive immunity. *Proc Natl Acad Sci USA*, 97:6924-6926
- Lander ES, Botstein D(1986). Mapping complex genetic traits in humans: new methods using a complete RFLP linkage map. *Cold Spring Harb Symp Quant Biol*, 51 Pt 1: 49-62
- Langefors Å, Von Schantz T, Widegren B(1998). Allelic variation of Mhc class II in Atlantic salmon: a population genetic analysis. *Heredity*, 80: 568-575
- Langer PR, Waldrop AA, Ward DC(1981). Enzymatic synthesis of biotin-labeled polynucleotides: Novel nucleic acid affinity probes. *Proc Natl Acad Sci USA*, 78: 6633-6637

- Larin Z, Monaco AP, Lehrach H(1991). Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. *Proc Natl Acad Sci USA*, 88(10): 4123-4127
- Larionov V, Kouprina N, Nikolaishvili N, Resnick MA(1994). Recombination during transformation as a source of chimeric mammalian artificial chromosomes in yeast (YACs). *Nucleic Acids Res*, 22: 4154-4162
- Larsson M, Tissir F, Walter L, Günther E, Jacob H, Klinga-Levan K, Levan G(1996). Mapping of the rat ribosomal protein S18 gene (*Rps18*) to chromosome 20p12. *Mamm Genome*, 7: 90
- Laurenson K, Sillero-Zubiri C, Thompson H, Shiferaw E, Thirgood S, Malcom LJ (1998). Disease as a threat to endangered species: Ethiopian wolves, domestic dogs and canine pathogens. *Anim Cons*, 1: 273-280
- Leonardo ED & Sevidy JM(1990). A new vector for cloning large eukaryotic DNA segments in *Escherichia coli*. *Biotechnology*, 8: 841-844
- Le Paslier MC, Pierce RJ, Merlin F, Hirai H, Wu W, Williams DL, Johnston D, LoVerde PT, Le Paslier D(2000). Construction and characterization of a *Schistosoma mansoni* bacterial artificial chromosome library. *Genomics*, 65(2):87-94
- Li R, Mignot E, Faraco J, Kadotani H, Cantanese J, Zhao B, Lin X, Hinton L, Ostrander EA, Patterson DF, de Jong PJ(1999). Construction and characterization of an eightfold redundant dog genomic bacterial artificial chromosome library. *Genomics*, 58(1): 9-17
- Luo J, Lang M, Salzburger W, Siegel N, Stolting KN, Meyer A(2006). A BAC library for the goldfish *Carassius auratus auratus* (Cyprinidae, Cypriniformes). *J Exp Zool B Mol Dev Evol*, 306(6): 567-574
- Luo M, Kim H, Kudrna D, Sisneros NB, Lee SJ, Mueller C, Collura K, Zuccolo A, Buckingham EB, Grim SM, Yanagiya K, Inoko H, Shiina T, Flajnik MF, Wing RA, Ohta Y(2006). Construction of a nurse shark (*Ginglymostoma cirratum*) bacterial artificial chromosome (BAC) library and a preliminary genome survey. *BMC Genomics*, 7:106
- LÜ Z, Johnson WE, Menotti-raymond M, Yuhki N, Martenson JS, Mainka S, Huang SQ, Zheng ZH, Li GH, Pan WS, Mao XR, O'Brien SJ(2001). Patterns of genetic diversity in remaining giant panda populations. *Conservation Biology*, 15: 1596-1607

- Madden DR(1995). The three-dimensional structure of peptide-MHC complexes. *Annu Rev Immunol*, 13: 587-622
- Maier E, Hoheisel JD, McCarthy L, Mott R, Grigoriev AV, Monaco AP, Larin Z, Lehrach H(1992). Complete coverage of the *Schizosaccharomyces pombe* genome in yeast artificial chromosomes. *Nature Genetics*, 1:273-277
- Makinen A, Chowdhary B, Mahdy E, Andersson L, Gustavsson I(1989). Localization of the equine major histocompatibility complex (ELA) to chromosome 20 by in situ hybridization. *Hereditas*, 110(1): 93-96
- Malykh A, Malykh O, Polushin N, Kozyavkin S, Slesarev A(2004). Finishing "working draft" BAC projects by directed sequencing with ThermoFidelase and Fimers. *Methods Mol Biol*, 255: 295-308
- Martins-Wess F, Milan D, Drogemuller C, Vobeta-Nemitz R, Brenig B, Robic A, Yerle M, Leeb T(2003). A high resolution physical and RH map of pig chromosome 6q1.2 and comparative analysis with human chromosome 19q13.1. *BMC Genomics*, 4(1): 20
- Martinson JT, Sousa AB, Guethlein LA, Howard JC(1999). The gene conversion hypothesis of MHC evolution: a review. *Immunogenetics*, 50(3-4): 168-200
- Matsuo MY, Hayashi S, Nonaka(2003). Evolution of the vertebrate MHC revealed by physical analysis of medaka (*Oryzias Latipas*) MHC class I region. Symposium on comparative and functional genomics of medaka, 1418-1419
- McCormick MK, Antonarakis SE, Hieter P(1990). YAC cloning of DNA embedded in an agarose matrix. *Genet Anal Tech Appl*, 7(5): 114-118
- Menz MA, Klein RR, Mullet JE, Obert JA, Unruh NC, Klein PE(2002). A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP, RFLP and SSR markers. *Plant Mol Biol*, 48(5-6): 483-499
- Michaelson J(1983). Genetics of beta-2 microglobulin in the mouse. *Immunogenetics*, 17(3):219-60
- Milner CM, Campbell RD(2001). Genetic organization of the human MHC class III region. *Front Biosci*, 6: D914-D926
- Moxon ER(1995). Whole genome sequencing of pathogens: a new era in microbiology. *Trends Microbiol*, 3(9): 335-337
- Mumm S, Jermak C, Waeltz P, Schmatz M, Terrell J, McCauley B, Shoemaker M, Srivastava A, Kere J, Nagaraja R(1998). 22-Mb integrated physical and genetic map based on YAC/STS content spanning the interval DXS1125-DXS95 in human

- Xq12-q21.31. *Gene*, 208(2): 147-156
- Murray DL, Kapke CA, Evermann JF, Fuller TK(1999). Infectious disease and the conservation of free-ranging large carnivores. *Anim Cons*, 2: 241-254
- Nagaraja R, MacMillan S, Jones C, Masisi M, Pengue G, Porta G, Miao S, Casamassimi A, D'Urso M, Brownstein B, Schlessinger D(1998). Integrated YAC/STS physical and genetic map of 22.5 Mb of human Xq24-q26 at 56-kb inter-STS resolution. *Genomics*, 52(3): 247-266
- Nash WG, Wienberg J, Ferguson-Smith MA, Menninger JC, and O'Brien SJ(1998). Comparative genomics: tracking chromosome evolution in the family Ursidae using reciprocal chromosome painting. *Cytogenet Cell Genet*, 83: 182-192
- Neuman E, Sowers AE, Jordan CA(1989). *Electroporation and Electrofusion in Cell Biology*. Plenum, New York, NY
- Nei M, Gu X, Sitnikova T(1997). Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc Natl Acad Sci USA*, 94: 7799-7806
- Nishimura M, Obayashi H, Mizuta I, Hara H, Adachi T, Ohta M, Tegoshi H, Fukui M, Hasegawa G, Shigeta H(2003). TNF, TNF receptor type 1, and allograft inflammatory factor-1 gene polymorphisms in Japanese patients with type 1 diabetes. *Hum Immunol*, 64: 302-309
- Nobrega MA & Pennacchio LA(2004). Comparative genomic analysis as a tool for biological discovery. *J Physiol*, 554(Pt 1): 31-39
- O'Brien SJ, Nash WG, Wildt DE, Bush ME, Benveniste RE(1985). A molecular solution to the riddle of the giant panda's phylogeny. *Nature* 317:140-144
- O'Connor M, Peifer M, Bender W(1989). Construction of large DNA segments in *Escherichia coli*. *Science*, 244(4910):1307-12
- O'Farrell PH(1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*, 250(10): 4007-4021
- Ohta Y, Goetz W, Hossain MZ, Nonaka M, Flajnik MF(2006). Ancestral organization of the MHC revealed in the amphibian *Xenopus*. *J Immunol*, 176(6): 3674-3685
- Okamoto K, Makino S, Yoshikawa Y, Takaki A, Nagatsuka Y, Ota M, Tamiya G, Kimura A, Bahram S, Inoko H(2003). Identification of IκBL as the second major histocompatibility complex-linked susceptibility locus for rheumatoid arthritis. *Am J Hum Genet*, 72: 303-312
- Okubo K, Hori N, Matoba R, Niiyama T, Fukushima A, Kojima Y, Matsubara K

- (1992). Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nat Genet*, 2(3): 173-179
- Olson MV, Dutchik JE, Graham MY, Brodeur GM, Helms C, Frank M, MacCollin M, Scheinman R, Frank T(1986). Random-clone strategy for genomic restriction mapping in yeast. *Proc Natl Acad Sci USA*, 83(20): 7826-7830
- Olson MV, Hood L, Cantor C, Bostein D(1989). A common language for physical mapping of the human genome. *Science*, 245: 1434
- Orr HT, Lopez de Castro JA, Lancet D, Strominger JL(1979). Complete amino acid sequence of a papain-solubilized human histocompatibility antigen, HLA-B7. 2. Sequence determination and search for homologies. *Biochemistry*, 18(25): 5711-5720
- Osoegawa K & de Jong PJ(2004). BAC library construction. From: *Methods in Molecular Biology*, vol 255: Bacterial Artificial Chromosomes, Volume 1: Library Construction, Physical Mapping, and Sequencing. Edited by: S. Zhao and M. Stodolsky . Humana Press Inc., Totowa, NJ
- Osoegawa K, de Jong PJ, Frengen E, Ioannou PA(2001). Construction of bacterial artificial chromosome (BAC/PAC) libraries. In *Current protocols in molecular biology on CD-ROM*. Edited by Ausubel FM, Kingston RE et al, John Wiley & sons, Inc., unit 5.9
- Osoegawa K, Tateno M, Woon PY, Frengen E, Mammoser AG, Catanese JJ, Hayashizaki Y, de Jong PJ(2000). Bacterial artificial chromosome libraries for mouse sequencing and functional analysis. *Genome Res*, 10(1): 116-128
- Otting N, de Groot NG, Doxiadis GG, Bontrop RE(2002). Extensive Mhc-DQB variation in humans and non-human primate species. *Immunogenetics*, 54(4):230-239
- Ottova E, Simkova A, Martin JF, de Bellocq JG, Gelnar M, Allienne JF, Morand S(2005). Evolution and trans-species polymorphism of MHC class IIbeta genes in cyprinid fish. *Fish Shellfish Immunol*, 18(3):199-222
- Ouedraogo JT, Gowda BS, Jean M, Close TJ, Ehlers JD, Hall AE, Gillaspie AG, Roberts PA, Ismail AM, Bruening G, Gepts P, Timko MP, Belzile FJ(2002). An improved genetic linkage map for cowpea (*Vigna unguiculata* L.) combining AFLP, RFLP, RAPD, biochemical markers, and biological resistance traits. *Genome*, 45(1): 175-188
- Parham P & Ohta T(1996). Population biology of antigen presentation by MHC class

- I molecules. *Science*, 272: 67-74
- Parham P, Lomen CE, Lawlor DA, Ways JP, Holmes N, Coppin HL, Salter RD, Wan AM, Ennis PD(1988). Nature of polymorphism in HLA-A, -B, and -C molecules. *Proc Natl Acad Sci USA*, 85(11):4005-9
- Passarge E(2001). *Color atlas of genetics*. Thieme press, 234
- Penn DJ(2002). The scent of genetic compatibility: sexual selection and the major histocompatibility complex. *Ethology*, 108: 1-21
- Penn DJ, Damjanovich K, Potts WK(2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proc Natl Acad Sci USA*, 99: 11260-11264
- Ploegh H & Watts C(1998). Antigen recognition. *Curr Opin Immunol*, 10: 57-58
- Potts WK & Slev PR(1995). Pathogen-based models favoring MHC genetic diversity. *Immunol Rev*, 143:181-197
- Potts WK & Wakeland EK(1990). Evolution of diversity at the major histocompatibility complex. *Trends Ecol Evol*, 5: 181-186
- Qian Y, Jin L, Su B(2004). Construction and characterization of bacterial artificial chromosome library of black-handed spider monkey (*Ateles geoffroyi*). *Genome*, 47(2): 239-245
- Rabin M, Fries R, Singer D, Ruddle FH(1985). Assignment of the porcine major histocompatibility complex to chromosome 7 by in situ hybridization. *Cytogenet Cell Genet*, 39(3): 206-209
- Renard C, Hart E, Sehra H, Beasley H, Coghill P, Howe K, Harrow J, Gilbert J, Sims S, Rogers J, Ando A, Shigenari A, Shiina T, Inoko H, Chardon P, Beck S(2006). The genomic sequence and analysis of the swine major histocompatibility complex. *Genomics*, 88(1): 96-110
- RepeatMasker web server. (<http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker>)
- Riethman H, Birren B, Gnirke A(1997). Preparation, manipulation, and mapping of HMW DNA. In *Genome Analysis: A laboratory manual. Vol 1 Analysing DNA* (ed. by Birren et al.): 83-248. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
- Riley J, Ogilvie D, Anand R(1992). In Anand, R. (ed), *Techniques for the Analysis of Complex Genomes*. Academic Press, New York, 59-79
- Robic A, Jeon JT, Rey V, Amarger V, Chardon P, Looft C, Andersson L, Gellin J, Milan D(2001). Construction of a high-resolution RH map of the human 2q35

- region on TNG panel and comparison with a physical map of the porcine homologous region 15q25. *Mamm Genome*, 12(5): 380-386
- Rogel-Gaillard C, Piumi F, Billault A, Bourgeaux N, Save JC, Urien C, Salmon J, Chardon(2001). Construction of a rabbit bacterial artificial chromosome (BAC) library: application to the mapping of the major histocompatibility complex to position 12q1.1. *Mammalian Genome*, 12: 253-255
- Romanov MN, Koriabine M, Nefedov M, de Jong PJ, Ryder OA(2006). Construction of a California condor BAC library and first-generation chicken-condor comparative physical map as an endangered species conservation genomics resource. *Genomics*, 88(6):711-718
- Ross MT, Hoheisel JD, Monaco AP, Larin Z, Zehetner G, Lehrach H(1992). High density gridded YAC filters: their potential as genome mapping tools. In: Anand R (ed) *Techniques for the analysis of complex genomes*. Academic Press Ltd., London, 137-153
- Sambrook J, Russell DW(2001) *Molecular Cloning: A Laboratory Manual* 3rd edition, Cold Spring Harbor Laboratory Press
- Saito S, Nakamura Y(1993). A genetic linkage map with 16 RFLP loci on the distal long arm of human chromosome 6. *Genomics*, 15(1):103-106
- Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR, Fiddes CA, Hutchison CA, Slocombe PM, Smith M(1977). Nucleotide sequence of bacteriophage phi X174 DNA. *Nature*, 265(5596): 687-695
- Schena M, Shalon D, Davis RW, Brown PO(1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, 270: 467-470
- Schierman LW, Nordskog AW(1961). Relationship of blood type to histocompatibility in chickens. *Science*, 134: 1008-1009
- Seddon JM, Ellegren H(2002). MHC class II genes in European wolves: a comparison with dogs. *Immunogenetics*, 54(7):490-500
- Shan W & Hardham AR(2004). Construction of a bacterial artificial chromosome library, determination of genome size, and characterization of an Hsp70 gene family in *Phytophthora nicotianae*. *Fungal Genet Biol*, 41(3): 369-380
- Shawar SM, Vyas JM, Rodgers JR, Rich RR(1994). Antigen presentation by MHC class I-B molecules. *Annu Rev Immunol*, 12: 839-880
- Sheng Y, Mancina V, Birren B(1995). Transformation of *Escherichia coli* with large

- DNA molecules by electroporation. *Nucleic Acids Res*, 23: 1990-1996
- Shizuya H, Birren B, Kim UJ, Mancino V, Slepak T, Tachiiri Y, Simon, M(1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci USA*, 89: 8794-8797
- Snell GD(1981). Studies in histocompatibility. *Science*, 213: 172-178
- Snell GD & HigginsGF(1951). Alleles at the histocompatibility-2 locus in the mouse as determined by tumor transplantation. *Genetics*, 36: 306-10
- Sommer S(2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Front Zool*, 2: 16
- Srivastava AK, McMillan S, Jermak C, Shomaker M, Copeland-Yates SA, Sossey-Alaoui K, Mumm S, Schlessinger D, Nagaraja R(1999). Integrated STS/YAC physical, genetic, and transcript map of human Xq21.3 to q23/q24 (DXS1203-DXS1059). *Genomics*, 58(2): 188-201
- Strong SJ, Ohta Y, Litman GW, Amemiya CT(1997). Marked improvement of PAC and BAC cloning is achieved using electroelution of pulsed-field gel-separated partial digests of genomic DNA. *Nucl Acids Res*, 25: 3959-3961
- Suzuki K, Asakawa S, Iida M, Shimanuki S, Fujishima N, Hiraiwa H, Murakami Y, Shimizu N, Yasue H(2000). Construction and evaluation of a porcine bacterial artificial chromosome library. *Animal Genetics*, 31: 8-12
- Suzuki T, Ota T, Fujiyama A, Kasahara M(2004). Construction of a bacterial artificial chromosome library from the inshore hagfish, *Eptatretus burgeri*: A resource for the analysis of the agnathan genome. *Genes Genet Syst*, 79(4): 251-253
- Takahata N & Nei M(1990). Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics*, 124, 967-978
- Takahashi K, Rooney AP, Nei M(2000). Origins and divergence times of mammalian class II MHC gene clusters. *J Hered*, 91:198-204
- Teichmann SA, Chothia C and Gerstein M(1999). Advances in structural genomics. Current opinion in structural biology, 9: 390-399
- Tian Y, Nie WH, Wang JH, Yang YF, Yang FT(2002). Comparative chromosome painting shows the red panda (*Ailurus fulgens*) has a highly conserved karyotype. *Yi Chuan Xue Bao*, 29(2):124-127
- The MHC sequencing consortium(1999). Complete sequences and gene map of a

- human major histocompatibility complex. *Nature*, 401: 921-923
- Thorsen J, Zhu BL, Frengen E, Osoegawa K, de Jong PJ, Koop BF, Dvaidson WS, Høyheim B(2005). A highly redundant BAC library of atlantic salmon (*salmo salar*): an important tool for salmon projects. *BMC Genomics*
- Thursz MR, Thomas HC, Greenwood BM, Hill AVS(1997). Heterozygote advantage for HLA class-II type in hepatitis B virus infection. *Nat Genet*, 17: 11-12
- Trachtulec Z, Hamvas RM, Forejt J, Lehrach HR, Vincek V, Klein J(1997). Linkage of TATA-binding protein and proteasome subunit C5 genes in mice and humans reveals synteny conserved between mammals and invertebrates. *Genomics*, 44(1):1-7
- Vaiman D, Billault A, Tabet-Aoul K, Schibler L, Vilette D, Oustry-Vaiman A, Soravito C, Crihiu EP(1999). Construction and characterization of a sheep BAC library of three genome equivalents. *Mamm Genome*, 10(6): 585-587
- Velculescu VE, Zhang L, Vogelstein B, Kinzler KW(1995). Serial analysis of gene expression. *Science*, 270(5235): 484-487
- Vienne A, Shiina T, Abi-Rached L, Danchin E, Vitiello V, Cartault F, Inoko H, Pontarotti P(2003). Evolution of the proto-MHC ancestral region: more evidence for the plesiomorphic organisation of human chromosome 9q34 region. *Immunogenetics*, 55(7): 429-36
- Vignal A, Milan D, SanCristobal M, Eggen A(2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genet Sel Evol*, 34(3): 275-305
- Vincek V, Nizetic D, Golubic M, Figueroa F, Nevo E, Klein J(1987). Evolutionary expansion of the MHC class I loci in the mole rat, *Spalax ehrenbergi*. *Mol Biol Evol*, 4: 483-491
- Wan QH, Fang SG, Wu H, Fujihara T(2003). Genetic differentiation and subspecies development of the giant panda as revealed by DNA fingerprinting. *Electrophoresis*, 24: 1353-1359
- Wan QH, Zhu L, Wu H, Fang SG(2006). Major histocompatibility complex class II variation in the giant panda (*Ailuropoda melanoleuca*). *Molecular Ecology*, 15(9): 2441-2450
- Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL, Humphery-Smith I(1995). Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis*,

16(7): 1090-1094

- Watkins DI, Hodi FS, Letvin NL(1988). A primate species with limited major histocompatibility complex class I polymorphism. *Proc Natl Acad Sci USA*, 85: 7714-7718
- Weber JL, Myers EW(1997). Human whole-genome shotgun sequencing. *Genome Res*, 7(5): 401-409
- Whisson SC, Drenth A, Maclean DJ, Irwin JA(1995). Phytophthora sojae avirulence genes, RAPD, and RFLP markers used to construct a detailed genetic linkage map. *Mol Plant Microbe Interact*, 8(6): 988-995
- Wong GK, Yu J, Thayer EC, Olson MV(1997). Multiple-complete-digest restriction fragment mapping: generating sequence-ready maps for large-scale DNA sequencing. *Proc Natl Acad Sci USA*, 94(10): 5225-5230
- Woo SS, Jiang J, Gill BS, Paterson AH, Wing RA(1994). Construction and characterization of a bacterial artificial chromosome library of Sorghum bicolor. *Nucleic Acids Res*, 22: 4922-4931
- Woon PY, Osoegawa K, Kaisaki PJ, Zhao B, Catanese JJ, Gauguier D, Cox R, Levy ER, Lathrop GM, Monaco AP, de Jong PJ(1998). Construction and characterization of a 10-fold genome equivalent rat P1-derived artificial chromosome library. *Genomics*, 50(3): 306-316
- Woychik RP, Klebig ML, Justice MJ, Magnuson TR, Avner ED(1998). Functional genomics in the post-genome era. *Mutat Res*, 400(1-2): 3-14
- Xie T, Rowen L, Aguado B, Ahearn BE, Madan A, Qin SZ, Campbell RD, Hood L (2003). Analysis of the gene-dense major histocompatibility complex class III region and its comparison to mouse. *Genome research*, 2621-2636
- Xu HL, Qian YP, Nie WH, Chi JX, Yang FT, Su B(2004). Construction, characterization and chromosomal mapping of bacterial artificial chromosome (BAC) library of Yunnan snub-nosed monkey (*Rhinopithecus bieti*). *Chromosome Res*, 12(3):251-262
- Xu LL, Shanmugam N, Segawa T, Sesterhenn IA, McLeod DG, Moul JW, Srivastava S(2000). A novel androgen-regulated gene, PMEPA1, located on chromosome 20q13 exhibits high level expression in prostate. *Genomics*, 66(3): 257-263
- Yokobata K, Trenchak B, De Jong PJ(1991). Rescue of unstable cosmids by in vitro packaging. *Nucleic Acids Res*, 19: 403-404
- Yu J & Wong GK(1997). The mappers' torch song. *Genome Res*, 7(7): 666-668

- Yuan B, Vaske D, Weber JL, Beck J, Sheffield VC. (1997). Improved set of short-tandem-repeat polymorphisms for screening the human genome. *Am J Hum Genet*, 60(2): 459-460
- Yuan X, Zhang M, Ruan W, Song C, Ren L, Guo Y, Hu X, Li N(2006). Construction and characterization of a duck bacterial artificial chromosome library. *Anim Genet*, 37(6): 599-600
- Yuhki N, Beck T, Stephens RM, Nishigaki Y, Newmann K, O'Brien SJ(2003). Comparative genome organization of human, murine, and feline MHC class II region. *Genome Res*, 13(6A):1169-1179
- Yuhki N & O'Brien SJ(1990). DNA variation of the mammalian major histocompatibility complex reflects genomic diversity and population history. *Proc Natl Acad Sci USA*, 87(2): 836-840
- Zhang YP, Wang W, Su B(1995). Microsatellite DNAs and kinship indentification of giant panda. *Zoological Research*, 16: 301-306
- Zeng CJ, Yu JQ, Pan HJ, Wan QH, Fang SG(2005). Assignment of giant panda MHC class II gene cluster to chromosome 9q by fluorescence in situ hybridization. *Cytogenetic and Genome Research*, 109: 534H
- Zhu L, Ruan XD, Ge YF, Wan QH, Fang SG(2007). Low major histocompatibility complex class II DQA diversity in the Giant Panda (*Ailuropoda melanoleuca*). *BMC Genet*, 8:29
- Ziegler A, Kentenich H, Uchanska-Ziegler B(2005). Female choice and the MHC. *Trends Immunol*, 26(9):496-502
- Zimmer R & Verrinder Gibbins AM(1997). Construction and characterization of a large-fragment chicken bacterial artificial chromosome library. *Genomics*, 42(2): 217-226

## 附录 攻读博士学位论文期间发表论文

Zeng CJ, Yu JQ, Pan HJ, Wan QH, Fang SG. Assignment of giant panda MHC class II gene cluster to chromosome 9q by fluorescence in situ hybridization. *Cytogenetic and Genome Research*, 2005, 109: 534H

Zeng CJ, Pan HJ, Gong SB, Yu JQ, Wan QH, Fang SG. Giant panda BAC library construction and assembly of a 650-kb contig spanning major histocompatibility complex class II region. *BMC Genomics*, 2007(in press)

## 致 谢

本论文是在我的导师方盛国教授和万秋红副教授的悉心指导下完成的。从论文的选题、实验分析以及论文撰写的每个环节都凝聚着两位老师的心血。我依然记得和方盛国教授、万秋红副教授 8 年前相识于成都大熊猫繁育研究基地，两位老师高瞻远瞩的学术视野、富于创造性的科学思维、严谨的治学精神以及平易近人的性格给学生留下了深刻的印象。四年来，尽管在实验中遇到了种种意想不到的困难，也犯了许多许多的错误，但是两位老师依然鼓励支持我坚持到了最后。感谢两位老师在这四年里在我学习、生活以及工作上的关心和帮助，学生永生难忘！

能够和实验室的葛云法老师相识是学生的荣幸。葛老师认真的办事和学习精神、宽广的为人、幽默的谈吐给我留下了深刻的印象，他在我学习生活中也给予了很好的建议和鼓励，同时也成为生活中很好的朋友！非常感谢 Chris Wood 和席咏梅博士在我学习期间的鼓励和建议，他们积极认真的学术态度，海外的学术经历也给予了我很大的启迪。

在我的实验过程中，得到了中国科学院北京动物所陈大元研究员的热情的支持！陈先生和蔼的为人、对科学研究的热情和投入，堪称学术的典范。在实验材料上，感谢陈先生提供了大熊猫的部分细胞样品并提供了方便的学术科研条件。在我的实验过程中，也得到了中国科学院遗传与发育研究所的程祝宽研究员、张力博士的大力支持，为我的荧光原位杂交实验提供了便利的实验条件，程先生严谨的治学态度、创造性的科学思维以及对科学研究的执着，尤其是仍然坚持在科学研究的最前沿，与学生一同实验的治学精神深深感染了我，在此表示诚挚的谢意！

感谢实验室朝夕相处的兄弟姐妹们，他们是博士生阮向东、何陆平、潘慧娟、赵莎莎、陈艺燕、叶青、硕士生吴建会、龚少斌、类延乐、张培、王晓亮、荆伟；已经毕业的徐艳春博士、余建秋博士、章金涛博士、吴华博士、许强华博士、蒋萍萍博士、张蓓博士、贺培建博士、吴海龙博士、胡杰博士、朱亮博士、华东师范大学的曹立荣博士、王江硕士、陈笑硕士、林翔硕士、陈轩硕士、郭洋硕士以及即将进入实验室的博士生何珂、鹿博、硕士生连红以及本科生焦阳、蒋静玲，大家共同营造了一个良好的科学研究环境，良好的学术氛围，使我在实验

室度过了一段难忘的时光。

衷心感谢我在四川农业大学的硕士导师郑鸿培教授和朱庆教授以及成都大熊猫繁育研究基地的张志和主任、侯容研究员，是几位老师带领我走进濒危野生动物保护研究的学术殿堂。还有四川农业大学动物科技学院的苏晓萍老师，感谢您多年来给我学习工作中的关心和支持！

最后，我要深深感谢我的父母和岳父岳母，感谢您们多年来的理解、支持和无私的奉献！我还要深深感谢我的妻子罗颖女士，多年来默默关心、支持、鼓励我的生活和工作、历尽了生活的艰辛和委屈，本论文也凝聚着她的心血和汗水。感谢我的可爱、聪明活泼的儿子曾智，他带来了做父亲的责任以及工作的动力！有了您们无私的奉献和支持，才能让我全身心投入学习和工作，顺利完成学业！

曾长军

2007年8月于浙江大学

濒危野生动物保护遗传与繁殖教育部重点实验室

国家濒危野生动植物种质基因保护中心