

## 第四部分

## 第三章

## 分子标记——探索遗传多样性的工具

## 1 导言

DNA 标记广泛应用于基础研究（例

如系统发育分析和有用基因寻找）和应用研究（例如标记辅助育种、亲权鉴定和食品可追溯性）。本部分主要关注 DNA 标记在动物遗传资源多样性特性鉴定中的应用，以及寻找相关基因的功能

插文 70

DNA、RNA 和蛋白质

DNA 由成对的染色体组成，分别来自两个亲本。因此，个体的每个基因都由两个副本，成为等位基因，每一个在两条染色体中的其中一条。哺乳动物的基因分散在染色体内，间隔很大，主要是重复的 DNA 序列。基因由编码序列（外显子）构成，基因之间由内含子隔离。后者不含蛋白质编码信息，但有时候在基因表达的调节过程中发挥重要作用。由基因编码的指令分两步进行。第一步是将遗传信息转录（复制）为另一类型的核酸——核糖核酸（RNA）。外显子和内含子都被转录到一个初级信使 RNA（mRNA）的分子中去。接下来，这个分子将被编辑，去处内含子，将外显子连接起来，并且在每个 mRNA 的末端添加独一无二的特征，从而创造一个成熟的 mRNA 分子，并接着被转运到细胞质内的核糖体内。核糖体由核糖体 RNA 和蛋白质构成，是第二步的进行场所。第二步是将先前复制到 mRNA 中的遗传信息编译成一个多肽（一个

完整蛋白质，或一个蛋白质复合体的一条链）。mRNA 分子被读取或一次编译 3 个核苷酸（1 个密码子）。mRNA 的密码子和携带相应氨基酸给核糖体的 tRNA 分子的反密码子的互补确保了最新的成形多肽包含所需氨基酸的特定序列。

并不是所有的基因都被翻译成蛋白，有一些将它们的功能表达成 RNA 分子（例如翻译过程中的 rRNA 和 tRNA）。最近，人们又发现了 RNA 在 mRNA 编辑过程中和基因表达调节中的新作用（Storz 等，2005；Aravin 和 Tuschl，2005；Wienholds 和 Plasterk 2005）。当然，似乎没有编码 RNA 是各种调节过程的主要调节者（Bertone 等，2004 Clop 等，2006）。因此，三种类型的分子在调查细胞、组织和整个机体水平上的遗传特性的过程中是有用的，它们分别是：DNA，包含已编码的指令；RNA，将指令传输到细胞工厂；由指令编译成的蛋白质，生成功能细胞和机体。

性变异。重要的是要指出RNA和蛋白质也含有重要信息，因此，应该进行平行研究。下面探讨了它们在寻找功能性变异中发挥的作用。

生物多样性是DNA序列变异和环境影响的结果。遗传变异是绝对的，某个物种的每个个体，除了孪生子之外，都拥有一个独一无二的DNA序列。DNA变异是由单核苷酸的置换（单核苷酸多态性，SNPs），各种长度的DNA片断（从单个到几千个核苷酸）的插入或缺失，或DNA片断的重复或到位。DNA变异分为“中性”变异和“功能性”变异。如果变异没有引起代谢或表型试验的变化，因此不属于积极、消极或平衡选择，称之为“中性”变异；反之为“功能性”变异。编码序列的关键核苷酸的突变可以改变蛋白质的氨基酸组成，因而导致新的功能性变异。与原始的“野生型”相比，新陈代谢的效率可能会提高或降低，可能会彻底失去功能，甚至可能获得新的功能。调控区的突变可能会影响基因表达的水平 and 模式；例如，在不同的发育或生理时期，特定组织中蛋白表达的启动/中止，或表达的不足/过量。

虽然简单生物分子的分析已经充分证明了其在理解生物现象中非常有用，但是大规模的DNA，RNA和蛋白平行调查揭示了生物复杂性的解释和模拟中的新观点。以“一组学”为后缀的新兴学科正在诞生。在这些领域，DNA，RNA和蛋白质的制备、鉴定和测序以及大容量数据的

存储和分析方面的新进展正给我们带来新的认识革命。关于复杂生物学过程中的生物分子的一整套观点正在浮出水面。继续

#### 插文 71

##### 新的“基因组学”学科

基因组学绘制了个体和群体之中基因与遗传变异之间的图表。它使人们能够洞察遗传信息向代谢功能和表型特征的翻译，揭开了一系列生物学过程及其与环境因子的互作。基因组学包含一系列高通量技术（如蛋白质组学和代谢组学）与生物信息学技术的组合，其中生物信息学技术实现了大量数据的处理、分析和综合。

生物复杂性的调查是一个新的边界领域，它需要高通量的分子技术，计算机的高速处理和海量存储，各学科专家之间数据分析与综合（插文 70）的新方法。

#### 插文 72

##### 分子技术的最新进展

当前有关畜禽育种和遗传多样性保存的分子生物学研究的革命性进展包括：1. 大部分重要畜禽物种的全基因组序列的确定；2. 全基因组多态性测量技术的发展（例如，单核苷酸多态性的检测方法）；3. 测量大范围基因转录的微阵列技术（芯片技术）的建立。全基因组测序得到的信息（鸡的已经完成，猪和牛的已基本完成），加之与SNP技术的结合，将会加快基因的寻找速度。用于鉴定影响目标特征的染色体区域的数量性状基因定位（QTL）绘图，位于同一区域的候选基因的存在，及其在不同物种之间的表达方式（例如，通过微阵列和蛋白组分析）和功能的调查，将会共同鉴定关键基因，并且揭开目标特征生理学调节的复杂性。

## 第四部分

构基因组学、转录组学和蛋白组学之后,除了其他系统生物学,又发展起来了复杂程度更高代谢组学和互作组学(Hood等,2004; Box 69)。

## 2 分子技术在特性鉴定中的作用

遗传多样性信息是优化动物遗传资源的保存和利用策略的根本。作为保存的资源是有限的,因此经常需要区分优先序。新兴的分子技术能够对包含在大量特征(包括适应性特征)中的基因,以及导致功能性遗传变异的多样性(数量性状核苷酸,QTN)进行鉴定。然而,我们没有足够的知识来对基于功能性分子多样性的保存选择区分优先序,因此仍然需要替代措施。表型特征提供了给定的个体或群体所携带基因的功能变量的粗略平均估计。然而,大多数畜禽物种的表型都没有被记录。

作用一:在缺乏可靠的表型和QTN数据或打算补充现有数据时,最快、最划算的遗传多样性措施可以通过使用匿名的分子遗传标记进行多态性分析来获得。假设在中性标记上已经拥有一个特殊进化历史的唯一种群可能携带功能变异的唯一变量,匿名标记可能会为重要特征提供间接的功能基因信息。在畜禽的起源和驯化以及后来的迁徙调查中,分子技术也被证明非常有用,就像提供进化关系信息(系统发育树)和鉴定不同遗传

起源的混合种群的地理区域一样。3.1节概述了在种内和种间遗传多样性评估中应用的分子技术。

作用二:有效种群大小( $N_e$ )是估计一个种群中为下一代繁殖和贡献基因的有效动物数量的指标。 $N_e$ 与一个种群中的杂交和基因飘移水平密切相关,因此是评估种群风险程度的一个重要指示因子(第一章和第六章)。对育种种群的 $N_e$ 进行可靠估计的传统方法主要基于家谱数据或种群普查。关于繁殖成功率和时代间隔变异的必要数据,对于发展中国家的种群往往不可靠。因此,分子方法可能是一个理想的选择(详细介绍见3.2节)。

作用三:动物遗传资源管理的最高优先序是对具有独一无二性状的品种的保存。其中,在挑战性环境中的生存和繁殖以及抵抗侵染性病害的能力显得尤为重要,尤其是在发展中国家。复杂性状,例如适应性和疾病抗性,是不明显的,或不易测量的。它们可以通过特定环境适应性试验或相关寄主侵染试验来进行调查。然而,这些试验实施起来非常困难且十分昂贵,还会引起有关动物福利的关注。这正是为什么研究者对鉴定基因控制的复杂性状怀有极大兴趣的原因。这些基因可以通过多种不同的方法找到。3.3节描述了现代研究发展起来的工具。

### 3 分子技术概览

本章主要讲述了目前已发展起来的用于遗传多样性评估和功能变异定位的最重要的分子技术。插文73描述了怎样将遗传材料从生物材料中提取，并用于分析。插文74概括了常用的分子标记的属性，插文75讨论了取样(分子研究的一个非常重要的方面)。

#### 3.1 使用DNA标记评估遗传多样性的技术

##### 核酸DNA标记

很多标记都可以用于检测核算DNA的多态性。在遗传多样性研究中，最常用的标记便是微卫星。

##### 微卫星

目前，微卫星(插文74)是畜禽遗传特性鉴定研究中最常用的标记(Sunnucks, 2001)。它们的高突变率和共显性特征使种内和种间遗传多样性以及品种间的遗传混合的估计成为可能，即使它们紧密相关。

针对微卫星数据分析的突变模型选择已有一些争论(无限等位基因模型或渐进突变模型; Goldstein等, 1995)。然而，模拟试验的结果已经标明无限定位基因模型通常对物种内多样性评估有效(Takezaki和Nei, 1996)。

##### 插文73

##### DNA和RNA的提取和扩增

DNA, RNA和蛋白质分析的第一步是生物样本的提取和净化。目前有几个方案和试剂盒可用。所运用策略根据原材料和目标分子而定。例如，从整个红细胞或白细胞中提取DNA相对比较简单，然而对加工过的食品却相当困难。从胰腺中提取RNA非常困难，因为此器官中的死后降解非常快。DNA, RNA和蛋白质的纯度往往是获取可靠结果过程中最容易忽视的因素。从细胞中将DNA(或RNA)分离后，下一步就是获得数以千计或百万计的一个特殊基因或DNA片断的复制。DNA片断的复制可以通过微生物来完成，比较典型的是大肠杆菌，也可以通过聚合酶链式反应(PCR)在体外完成。卡里·穆利斯的PCR技术曾获得诺贝尔发明奖，几乎可以对任何一直序列的DNA片断进行指数扩增。PCR反应的关键部分是从一种能在高温下生存和繁殖的微生物——*Thermus aquaticus*中分离出来的DNA聚合酶。这种热稳定的Taq酶可以允许在PCR循环中进行链式复制，并在目标DNA的复制数量上产生几何增长。一个PCR循环包括三步：i) 在90~95℃下使DNA变性，将DNA分成两条模板；ii) 在45~65℃下退火，短单链寡核苷酸(引物)与靶标序列连接；iii) 新合成的DNA链在引物的引导和Taq酶的催化下延伸或延长，温度设定在72℃。次循环一般能被重复25~45次，来保证扩增用于检测的足够的扩增子。

蛋白质多态性是用于畜禽遗传研究的一种标记。然而，能够进行分析的多态性位点的数量和位在位点观测到的多态性水平常常很低，因此大大限制了其在遗传多样性研究中的应用。随着新技术的发展，DNA多态性已经成为标记的选择，来进行遗传变异的分子监测。

## 第四部分

插文 74  
常用 DNA 标记

限制性片段长度多态性(RFLP)是利用限制性内切酶能识别 DNA 分子的特异序列,并在特定序列处切开 DNA 分子(例如 EcoRI 在回文序列 GAATTC 定义的位点将 DNA 切开)。目前, RFLP 的最主要的用途是作为 PCR 的下游工具(PCR-RFLP),来检测在特定的限制性酶切位点具有不同序列的等位基因。一个基因片断首先通过 PCR 进行扩增,然后再暴露于一个只能切开其中一个等位基因特定的限制性内切酶。已消化的扩增子通常用电泳来溶解。

微卫星或简单重复序列(SSR)或简单串联重复(STR)由一类由 2~6 个碱基组成的基序串联重复而成的 DNA 序列(例如, CACACACACACACA)。它们通过真核细胞染色体来复制传播。微卫星相对较小,因此可以很容易地从各种资源(例如血、毛发、皮肤,甚至粪便)中提取的 DNA 中用 PCR 扩增出来。多态性可以在测序凝胶上显现出来,并且 DNA 自动测序的有效性还允许进行大量样本的高通量分析(Goldstein 和 Schlötterer, 1999; Jarne 和 Lagoda, 1996)。微卫星是超变量;它们通常可以展示在大量重复中互不相同的位点上的数十个等位基因。它们还是许多多样性研究(例如亲本分析和 QTL 作图)的标记选择,但是随着廉价的 SNP 分析方法的发展,

在不久的将来 SSR 将受到挑战。FAO 已经推荐了用于主要畜禽物种多样性研究的微卫星定位装置,该装置由国际动物遗传学会(ISAG)——FAO 动物遗传多样性顾问小组开发。(见 DAD-IS 文库)

小卫星具有与微卫星相同的特点,但是串联重复为 10 到几百个碱基。微卫星和小卫星都被认为是可变性串联重复序列(VNTRs)多态性。扩增片段长度多态性(AFLPs)是一种 DNA 指纹识别技术,可以通过 PCR 扩增检测 DNA 的限制性片断。

序列标签位点(STS)是染色体中仅出现一次的 DNA 片断,它的位置是已知的。它们不需要进行多态性分析,而被用于建立物理图。

单核苷酸多态性(SNPs)是单个核苷酸的变异,并不改变区域的 DNA 序列的总体长度。SNP 的发生遍布整个染色体。这种变异非常普遍,在人类染色体中,每 1000 个碱基中就有 1 个 SNP 出现(Sachinandam 等, 2001)。大多数 SNP 存在与非编码区,并且对个体的表型没有直接影响。然而,有些 SNP 在表达序列或影响基因表达的区域(启动子,增强子)中引发突变,并且可以诱导蛋白结构或调节的变化。这些 SNP 具有检测功能遗传变异的潜力。

每个种群的等位基因平均值(MNA),以及观察到和预期的杂合性(Ho和He)是评估种内多样性的两个最常用的参数。最简单的种间多样性评估参数是遗传分化因子和遗传固定因子。目前,已经有几个评估量被提议,例如 FST 和

GST,其中应用最广泛的是 FST(Weir 和 Basten, 1990),它通过计算种群之间的等位基因频率的标准方差来测量遗传分化程度。差异显著性可以通过种群之间的 FST 值来计算(Weir 和 Cockerham, 1984),以测量种群和遗传多样性划分之间的遗传分

## 插文 75

## 遗传材料的取样

样本收集是任何多样性研究的第一步也是最重要的一步。也就是说,样本在调查种群中应该独立且具有代表性。通常,如果分析了足量的独立标记,每个种群精选出30~50个个体进行取样就能够为品种特殊性和种内多样性提供所需的初步线索(例如20~30个微卫星; Nei and Roychoudhury, 1974; Nei, 1978)。然而,所需的实际数字可能视具体情况不同而不同,并且对于高度近亲繁殖的当地种群甚至低一些,对于广泛传播且分为不同生态型的种群可能会高一些。

在定义明确的品种中,独立样本的选择非常简单,因为可以通过畜禽血统书或家谱记录来选择。相反,由于没有可用的书面记录,对于半野生的种群,独立样本的选择将十分困难。在这种情况下,地理准则被大力推荐,也就是说,从大量在广阔地理区域传播的畜群中,在每个畜群中选出一个或少数几个独立的动物个体,并将它们集中在一起。取样地点、动物和畜群的相关地理记录和照片文档非常重要,可用来检查意外露宿情况下的杂交情况,或鉴定遗传多样性的兴趣地理模式。一套选择良好的样本是一中长期有价值的资源,它可以用来产生有价值的结果,即使利用比较落后的技术。相反,有偏差的样本将产生歪曲的结果,或即使利用最先进的分子技术都难以理解。

化缺乏的虚假设(例如, Mburu 等, 2003)。分子变异(AMOVA)(Excoffier 等, 1992)的递阶分析可以用来评估品种种群内和种群间的多样性分布。

通过对遗传距离的估计,微卫星数据也通常被用来评估种群和个体之间的遗传

关系(例如, Beja-Pereira 等, 2003; Ibeagha-Awemu 等, 2004; Joshi 等, 2004; Sodhi 等, 2005; Tapio 等, 2005)。最常用的测量遗传距离的方法是Nei氏标准遗传距离(DS)(Nei, 1972)。然而,对于密切相关的种群,基因飘移是遗传分化的主要影响因素,由于这种情况在家畜中经常发生,尤其是在发展中国家,因此推荐使用修正的Cavalli-Sforza遗传距离(DA)(Nei 等, 1983)。一般通过种族系统史的重建可以显现种间的遗传关系,最常用的就是邻近距离法(N-J)(Saitou 和 Nei, 1987)。然而,系统发育树重建的一个主要障碍就是家系的进化被假设为非网状的,也就是线性可以偏移,但线性之间不会交叉。这种假设在家畜中发生的几率很小,新的品种常常起源于两个或更多祖先品种的杂交育种。因此,必须慎重考虑系统发育树重建提供的品种进化的可视化。

多元分析和最近的贝叶斯分聚类分析方法已经被建议用于不同种群的微卫星数据的混合分析(Pritchard 等, 2000)。家畜最有意义的此种类型研究大概就是非洲牛的跨洲研究,它展现了起源,二次移动的遗传信号,以及非洲牛的畜牧主义的区别。

分子遗传数据与其他资源(如考古学证据和书面记录)的相互联合与补充,为畜禽品种的遗传多样性的起源、随后的移动和发展提供了有用信息。由于表型变异的数据有限,现有遗传多样性的作图允许对物种的功能性遗传变异可能发生的位置

## 第四部分

进行推测。

虽然人们非常希望对每个独立研究获得的微卫星数据进行组合分析,但实际上几乎不可能。因为大多数利用 DNA 标记的种群遗传研究局限于少量品种,并且通常来自一个国家 (Baumung 等, 2004)。在研究过程中,人们经常使用 FAO 推荐标记的不同种类,并且不同研究项目之间没有明确基因型的标准样本。不同的微卫星基因型系统的应用引起了相同位点上等位基因片断大小估计的差异。为了提高通用标记的使用率,FAO 正为主要畜禽提议一个最新的微卫星位点分类列表。FAO 推荐使用分级归类的标记,以达到独立调查研究中标记重叠的最大化。对于某些物种,可以利用标准动物的 DNA。例如,欧盟 Econogene 项目中使用的绵羊和山羊标准 DNA 试样已经被亚洲和非洲的许多其他项目所选用,并且可以通过 Econogene 项目的网站 (<http://www.econogene.eu>) 提出使用请求。

畜禽遗传多样性的大范围分析的案例只有几个。Hillel 等人 (2003) 和 SanCristobal 等人 (2006a) 分别调查了整个欧洲的鸡和猪的多样性; Hanotte 等人 (2002) 获得了几乎整个非洲大陆的牛的数据; Tapio 等人 (2005) 评估了北欧国家广大地区范围的绵羊多样性; Ca\_on 等人 (2006) 研究了欧洲、中东和远东地区的山羊多样性。然而,对于大多数畜禽来说,仍然缺乏全面的综述。正在进行的大范围项目之间的密切合作保证了在不久的

将来,对于一些品种,例如山羊和绵羊,遗传多样性全球估计的传递交流。其间,正在开发的新的数据分析方法将能实现对仅有少数品种,并且没有或仅有几个共同标记的数据集的综合分析 (Meta 分析) (Freeman 等, 2006)。畜禽多样性的这种全球设想将对重建畜禽种群,包括人类的起源和历史非常有价值。它也将加强以保存动物遗传多样性为目标的区域和局部热点。

#### 单核苷酸多态性

单核苷酸多态性也是遗传多样性研究除微卫星之外的可选技术。目前用于检测和定位 SNP 标记的方法已有几种 (参见 Syvänen, 2001)。作为等位基因标记, SNP 包含的信息内容非常少,因此要想达到 30 个微卫星位点的标准品所得到的信息水平就需要更大的样本量。然而,正在发展的分子生物学技术增加 SNP 定位的自动化水平,同时减少耗费。看来,在不久的将来,低成本的大量标记的并行分析方法将可能面世 (例如 Wong 等, 2004)。据此展望,大规模项目正在几种畜禽中进行,对数百万的 SNP 进行鉴定,以及对数千 SNP 进行验证,并且鉴定染色体中的单体型区断。与序列信息类似, SNP 可以对不同实验进行直接的对比分析和结合分析。

SNP 似乎是未来遗传多样性研究中非常吸引人的标记,因为它可以被方便地用于评估功能性或中性变异。然而, SNP

发明的最初阶段或 SNP 从数据库中的选出非常重要。SNP 可以通过多种实验方案产生(例如基因测序,单链构象多态性(SSCP),变性高效液相色谱(DHPLC)),或与公众基因和表达序列数据库中相同区域的多重序列进行电子标记和比较。当数据还没有被随机获得时,群体遗传参数的评估标准是不能够应用的。较为频繁的例子为:当 SNPS 开始在小样本(板,panel)中鉴定,然后又放入染色体这样的大样本中鉴定时,则先前小样本中出现的中间频率的 SNPS,与所期望的随机样本中等位基因变异的分布相比会出现偏差。SNPS 确实能够应用到未来群体遗传分析中去,但是那些明确可以发现 SNP 技术的统计方法一定要得到发展(Nielsen 和 Signorovitch, 2003; Clark 等, 2005)。

#### 扩增片段长度多态性

AFLPs 为双显性分子标记(Vos 等, 1995)。未知染色体区域的很多单核苷酸突变可以同时被检测出来,而且突变经常会出现在未知功能的基因中。由于 AFLP 具有显性遗传模式,因此其分析品种内多样性和近交系时会显示出它的缺点。然而,当分析品种间和相关物种间的多态性时,会体现出其信息含量高的特点。

#### 线粒体 DNA 标记

线粒体 DNA 多态性已经被广泛应用

于系统发生和遗传多态性分析。细胞质里线粒体携带的单倍体 mtDNA 具有母性遗传模式(即个体从它们的母本继承 mtDNA,而不是父本)和高突变率,并且不能重组。这些特性使进化生物学家可以通过评估 mtDNA 里的突变模式来重建种内和种间的进化关系。MtDNA 还可以提供一种检测畜禽品种或亚种之间杂交关系的快速方法(例如, Nijman 等, 2003)。

D 环的高变区或 MtDNA 控制区的序列多态性已经在家畜野生祖先的鉴定,遗传多样性地理模式的建立,以及家畜驯化的理解等方面做出了巨大贡献(Bruford 等, 2003)。现代欧洲牛的中东起源最近被 Troy 等人证实(2001)。该研究还鉴定了普通牛的四个母性来源,也证明了在人类新石器时代人类迁出新月沃土的过程中牛的遗传多样性的损失。同样,具有三个 mtDNA 来源的多重母性起源在山羊中非常显著(Luikart 等, 2001),并且亚洲和新月沃土是可能的起源中心。最近,一个 mtDNA 第三支系、第四支系和第五支系分别在中国绵羊(Guo 等, 2005)、中国山羊(Chen 等, 2005)和中国牛(Lai 等, 2006)中被发现。在亚洲鸡品种中,已发现 9 个不同的 mtDNA 支系(Liu 等, 2006),表明多重起源在南亚和东南亚。所有这些结果表明我们对家畜驯化和遗传多样性的了解还远远不够。家畜起源的深入讨论见第一部分第一章。

## 第四部分

## 3.2 利用标记估计有效种群大小

Hill 在 1981 年曾建议利用 DNA 多态性的配子相分布来估计有效种群大小 ( $N_e$ )。由于连锁的标记, 这种估计可以基于基因型连锁分子标记 (微卫星或 SNP)。连锁位点的等位基因频率的预期相关性是  $N_e$  和重组率的一个函数。因此,  $N_e$  可以通过观察到的不均衡性来估计。Haye 等人 (2003) 提出了一种基于染色体片段纯合性的类似方法。另外, 这种方法还具有估计早期世代  $N_e$  的潜力, 因此还可以判断现有的种群大小在过去是增加还是减少。通过示例数据集, 研究表明荷兰黑白花乳牛品种在过去经历了一个  $N_e$  的实质减少过程, 然而人类种群的有效种群大小在不断增加, 这与人口普查和家系研究一致。

## 3.3 针对功能变异的分子工具

基于地图位置的方法: 数量性状基因位点作图 (QTL 作图) 遗传标记表现为孟德尔法则性状; 换句话说, 它们遵循由孟德尔最先提出的分离和自由组合规律。位于同一个染色体上的两个基因是物理连锁的, 并具有一起遗传的趋势。在减数分裂过程中, 同源染色体之间的充足可能会打破这种连锁。位于同一染色体上的两个基因之间的重组频率取决于它们之间的距离。因此, 遗传标记之间的重组率是它们连锁程度的指示因子: 重组率越低, 标记之间间隔的就越近。遗传图谱的构建开发了这种特性, 并用来推断遗传作图中标记之间的近似距离和可能顺序。

作图实践通常根据选择性计划 (全

同胞或半同胞家族) 的结构化实验种群 (例如 F2 代或回交代) 或现有种群多态性标记的共分离来实现。具有几百至几千标记的中高密度遗传图谱对大多数畜禽都适用。

为鉴定一个特定性状的 QTL, 可以用一套均匀分布于基因组的已测定位置的分子标记来确定特例家族的基因型 (插文 76)。已有很多统计方法可以用来推断特定标记间隔内是否存在重要的 QTL, 但是所有方法都基于畜禽家族拥有一个高水平的连锁不平衡: 即染色体的很大片断在从亲代向子代的传递中没有进行重组。

QTL 作图相通常采用精妙的 QTL 作图位置 (QTL 精细定位)。为完成这项任务, 需要对附加标记及以上目标区域发生的所有其他重组事件进行分析。最近, 人们又设计了一个巧妙方法, 并应用于染色体上 BTA14 区域的精细作图, 该区域携带了一个与牛奶脂肪含量和其他性状有关的重要 QTL (Farnir 等, 2002)。这种方法利用了过去世代的历史性重新组合将作图位置限制到相对较小的 3.8 cM 区域, 一个允许基因定位克隆的大小 (DGAT1) (Grisart 等, 2002)。

通过精细定位, 人们就可以从鉴定区域的基因中找到决定生产性状的基因。候选基因也可以在相同的畜禽 (例如, 当充足的 EST 图谱可以利用时, 或基因组被完全测序时) 或模式动物的直系同源区域中找到, 从模式生物中我们可以得到全部的基因组信息。

## 插文 76

## QTL 作图

如果一个目标性状存在一个 QTL, 未知的相关基因 (Q 和 q) 的正、负变异的等位基因将与 M1 标记附近的等位基因共分离, 其中 M1 (M1 和 m1) 能够在实验室内测定基因型。例如, 让我们假设 M1 和 m1 分别与 Q 和 q 进行共分离, 因此 M1 和 Q 在同一条染色体上紧密相邻, m1 和 q 在同源染色体上紧密相邻 (MIQ and m1q)。

让我们也假设测定了一个起源于 F1 代杂合个体杂交的 F2 代种群的基因型。根据基因型, F2 代的后代基于它们标记基因型 (M1M1 和 m1m1; M2M2 和 m2m2; ... MnMn 和 mnmn) 进行分组, 接下来对群体的平均表现型进行比较。如果没有 QTL 与给定的标记 (例如 M2) 连锁, 那么 M2M2 和 m2m2 后代的表型值之间对目标特性来说就检测不到重大差异。相反, 如果后代按照它们在标记 M1 上的基因型来分组, 那么 M1M1 组在 QTL 初大部分将是 QQ, m1m1 组大部分将是 qq。在这种情况下, 将能观察到后代平均值之间的重大差异, 因此

能检测到 QTL 的存在。对于某些品种, 品系和品种一般通过商业杂种繁育, 例如家禽和猪, 这种措施可以在实验种群 (F2, BC) 中完成, 然而对于反刍动物, 通过要用到两代 (父女设计, DD) 或三代 (孙女设计, GDD) 家系。在父女设计中, 杂合在父本 (第一代) 中的标记被分离并传递给女儿 (第二代), 并且通过第二代来收集表型数据。在孙女设计中, 杂合在祖父本 (第一代) 中的标记被分离并传递给它的半同胞儿子 (第二代), 半同胞儿子的表型可根据孙女 (第三代) 的表型推断得到。

QTL 作图试验的结果鉴定了染色体区域, 一般可以生成半条染色体, 并利用其检测目标性状的重要影响因素。现代研究正积极利用作图来鉴定适应 QTL 影响的性状。这些性状在鸡中的例子包括对沙门氏菌的侵染和排泄物不断增加的抗性 (Tilquin 等, 2005), 肺部高血压综合症发生的敏感性 (Rabie 等, 2005), 以及牛的锥虫病 (Hanotte 等, 2002)。

有时候, 一个意外的资源可以带来基因功能的关键信息。肌肉生成抑制素基因就是这种情况, 它的功能首先在小鼠中被发现, 然后才在牛的双肌基因事先被标记的染色体区域内被发现 (McPherron 和 Lee, 1997)。

很明显, 鉴定一个复杂性状的相关基因 (数量性状基因, QTG) 和功能突变 (QTN) 仍然是一项重要任务, 还需要开发几种方法来减少位置候选基因的数量。在这个层面上, 基因的功能信息是根本。然而, 我们对基因组测序和 cDNA (互补

DNA) 测序得到的大多数基因的可能功能仍然知之甚少。这就是为什么调查基因表达模式能够与前面所讲的定位方法结合起来提供有用信息的原因, 并为复杂性状鉴定候选基因。这种组合方法被称为“遗传基因组学” (Haley 和 de Koning, 2006)。基因表达模式调查的最新进展在下一节中介绍。

目前, 人们正在研究利用遗传标记检测环境适应基因的替代方法 (插文 77)。它们现在正处于试验阶段, 并且还需要深入研究来评估这些方法的效能。

## 第四部分

## 插文 77

## 种群基因组学方法

携带相关基因的基因组区域的一个替代性鉴定方法最近已被提议。它由通过“种群基因组学”方法而进行的“选择签字”组成(Black等, 2001; Luikart等, 2003): 利用种群基因组学进行QTL作图的三个主要原则分别是: (1) 跨基因组的中立(不确定性)位点将同样受到基因飘移、种群统计和种群进化史的影响; (2) 选择位点往往会有不同表现, 因此显示出变异的“异常值”模式, 多样性的损失(如果位点处于一个平衡选择将会增加多样性), 连锁不平衡以及Gst/Fst指示因子的增加/减少; (3) 通过搭便车效应选择也会影响连锁标记, 承认“选择签字”的检测(异常值效应), 通常可以通过测定大量染色体标记的基因型以及鉴定“异常值”的类别来检测。这种方法利用的是品种(或品种内的亚种群)水平的表型数据, 而不是个体水平的, 因而精细地补充了传统的家系内QTL作图方法。

种群基因组学方法也可以鉴定强选择压下的基因以及最终固定在品种内的基因, 尤其是含有适应极端环境、抗病性等性状的基因。这

些性状对动物可持续育种非常重要, 其中大部分难于或不可能通过传统的QTL作图或相关方法进行调查。最近, 研究者分别利用理论观点(Beaumont和Balding, 2004; Bamshad和Wooding, 2003)和自然种群中不同类型标记的具体试验(AFLPs分析: Campbell and Bernatchez, 2004; 微卫星: Kayser等, 2003; 单核苷酸多态性(SNPs): Akey等, 2002)对种群基因组学的潜力进行了调查。经过初步分析, 绵羊的MYH1(阻凝蛋白1), MEG3(callypige), 和CTSB(蛋白酶B)基因表现出重大的异常行为(Pariset等, 2006)。

同一个项目, 基于空间分析方法(SAM)的一种新方法已被设计用来检测家畜和野生动物的签名自然选择(Joost, 2006)。利用此方法得到的初步结果与运用种群基因组学的理论模型(例如由Beaumont和Balding在2004年开发的模型)所得到的结果一致。既然它被设计用来鉴定与选择标记有关的环境参数, 因此与传统的方法相比, SAM又前进了一步。

QTL作图的根本目标是鉴定QTG, 并最终鉴定QTN。虽然迄今为止在家畜中只有少数几例, 但这些却是能够对标记辅助育种和保存决策产生直接影响的突变类型。在不久的将来, QTG和QTN的数量将不断增加, 因此我们需要对基于功能性状和突变的保存模型进行充分考虑。

## 基因表达的调查模式

过去, 只能在表型水平上对特定性状(如适应性和抗性)的表达进行测量。现

在, 转录组(细胞或组织中所有转录产物的集合体)和蛋白质组可以直接通过高通量技术, 例如差异显示(DD)(Liang和Pardee, 1992), cDNA-AFLP(Bachem等, 1996), 基因表达的系列分析(SAGE)(Velculescu等, 1995; 2000), 质谱, 以及蛋白质和DNA微阵列分析。这些技术代表了RNA和蛋白分析的突破, 从而实现了特定时间范围内组织中所有实际表达的基因的平行分析。因此, 这些技术将有助于可能支撑许多复杂形状的网络的编码。

组学技术往往被比喻成在一幅米开朗基罗壁画前打开灯光,而不是使用一个仅能看到局部的火把。全貌图使其表现出的含义得以理解,使其美丽得以欣赏。事实上,这些技术目前正面临技术运用和数据分析的困难和所需的巨大花费。同类细胞样本的分离相当困难,并且是许多基因表达谱研究的一个重要先决条件。大量平行测定的单个测验花费较低,但是整个试验花费却相当高。实验设备非常昂贵,并且所有的实验阶段都需要很高的实验技能。除了常规困难之外, RNA 分析与 DNA 分析相比,还面临新的困难。因为 RNA 非常容易降解,必须从新陈代谢非常活跃的组织中非常小心地提取。实际上,样品保存和处理是成功进行 RNA 分析试验的一个关键。然而,纳米技术在生物分析中的应用为解决这些问题提供了广阔的应用前景 (Sauer 等, 2005)。

数据处理是一个更深层次的问题。分子数据集,例如基因表达谱,可能会在相当短的时间内产生。然而,不同实验室之间的数据需要进行标准化来满足不同生物数据集一致性分析的要求。标准化的一致性,也就是互联数据库的建立,是对分子网络进行有效分析的基础。

### 转录谱

本部分主要介绍了 SAGE 和微阵列技术。最近的一些综述描述了其他技术 (例如, Donson 等, 2002)。SAGE 产生组织或细胞系的全部表达谱。它包括全部

mRNA 库的构建, mRNA 库能够在细胞激活的特定步骤对整个表达或失活的转录进行数量分析。它基于三条原则: (i) 从每个 mRNA 转录中给定区间中获得的一个短序列标签 (9~14bp), 包含唯一鉴定一个特定转录的足够信息; (ii) 序列标签可以连锁在一起形成 DNA 分子 (串联体), 串联体可以被克隆和测序——串联体克隆的测序可以进行大量单个标签的快速鉴定; (iii) 转录的表达水平可以通过一个特定标签被观察到的次数而量化。

微阵列在独立试验中可被用于比较两个生物系统 (例如通常环境和挑战性环境中的动物) 的数千基因的 mRNA 表达水平。微阵列技术还可以提供对广泛适应生物暴露因素的基因表达的时空模式的理解。

将极少量的 DNA 溶液印迹在一个无孔玻片材料上, 斑点直径为 100~150  $\mu\text{m}$ 。现在, 大约 50 000 个 DNA (cDNAs) 被点在显微镜载玻片上。DNA 微阵列分析包括数百个已知基因和几千个未知基因。微阵列可以用 cDNA 片断或预制的寡核苷酸作标记, 其中后者具有较高的灵敏度和再现性, 但仅限于已知序列的设计。微矩阵的利用基于“杂交”原则, 也就是两个单链 DNA 或一个 DNA 和 RNA 序列相互暴露, 然后测量形成的双链分子的数量。mRNA 的表达可以被定性和定量。它显示了一个组织的基因活性, 并且往往与这种 mRNA 诱导的蛋白表达直接相关。

## 第四部分

基因表达谱有利于生物学机理的理解,因此推动了候选基因的鉴定。例如,牛锥虫病抗性表达中的基因库已分别通过 SAGE (Berthier 等, 2003) 和 cDNA 微阵列分析 (Hill 等, 2005) 得到了鉴定。很多基因表达的平行调查可以鉴定控制表型且通过微分表达分析仍未检测到的主效基因。例如,这些主效基因可以拥有在同一水平表达的不同等位基因,可以不同程度上提高下游基因的表达。在这种情况下,主效基因既可以通过挖掘代谢途径的现有知识来找到,也可以通过表达 QTL 方法找到 (Lan 等, 2006)。根据这种方法,下游基因的表达水平需要在隔离种群中测量。每个基因的转录量可以利用上一节中的方法获得。

## 蛋白谱

蛋白结构(转译为蛋白谱),蛋白—蛋白、蛋白—核酸以及蛋白—小分子之间的互动,以及蛋白在真核细胞中的时空表达的系统研究对复杂生物现象的理解是至关重要的。蛋白是活体细胞及其功能的基本结构。

蛋白的结构可以通过 X-射线衍射或核磁共振来显现。X-射线衍射需要大量的结晶蛋白,而结晶蛋白往往较难获得。为了解蛋白质的功能以及蛋白与蛋白在分子水平上的互动,测定细胞或组织中所有蛋白的结构将非常有用。然而,目前此项工作尚未完成。有趣的是,出现在蛋

白合成过程中不同蛋白变异(插入和/或转译后修饰)的数量要比染色体中的基因数量大的多。

质谱(一种测定分子量的分析技术)与色谱或电泳分离技术的联机是目前鉴定细胞中的内生蛋白,描绘转译后修饰的特性以及测定蛋白丰度的可选方法(Zhu 等, 2003)。二维凝胶电泳是能够在一次单独实验中分离和显现大量蛋白质(>10 000)的唯一方法。将蛋白质斑点从凝胶上切下,进行蛋白酶消化,然后将蛋白质用质谱进行鉴定(Aebersold 和 Mann, 2003)。然而,二维凝胶电泳的标准化和自动化非常困难,使用已得到的蛋白模式作为蛋白质组基准图仅仅适用于少数情况。

液相色谱作为一种辅助技术,容易实现自动化,并且可以直接与质谱进行联机。基于微阵列的亲合蛋白质组方法可以作为蛋白谱的替代方法(Lueking 等, 2003),并且可以用于检测蛋白与蛋白之间的互动。这些信息是生物途径算法模型的基础。然而,结合特异性仍然是蛋白质微阵列分析方法应用中的一个问题,因为交叉反应不能被精确预测。对于蛋白与蛋白之间的互动检测,还有替代方法,例如酵母双杂交系统(Fields 和 Song, 1989)。然而,目前使用的方法仍然不能对结合蛋白进行定量检测,并且对于观察到的互动对生理性的蛋白—蛋白互作的可能表现程度仍然不清楚。

基于阵列的方法还可以用于检测活体

和离体的 DNA—蛋白互作 (Sauer 等, 2005), 鉴定结合于基因表达调控序列的未知蛋白。DNA 微阵列的使用可以有效监测键合 DNA 复合体的核提取物, 然而, 蛋白微阵列主要用于鉴定蛋白质组范围水平上的未知 DNA 键合蛋白质。将来, 这两种方法将展现对转录调节网络更细致的洞察力。

预测蛋白功能的许多方法都基于预测蛋白于其他蛋白的同源性及其在细胞内的位置。蛋白功能的预测非常复杂, 并且需要检测蛋白—蛋白互作以及蛋白与其他分子互作的技术, 因为蛋白质在这些结合过程中实现其功能。

#### 4 生物信息学的作用

如果没有能力对呈指数上涨的生物学数据进行分析, 那么开发高通量的技术将没有任何用处。那些需要以电子形式存储的数据库要与设计好的特定软件结合起来 (插文 78), 以便数据的更新、调出和修复。这些信息必须是容易理解的, 并且可以灵活地调出以保证信息可以得到修复。这些信息可以用来分析新陈代谢途径以及蛋白质和核酸的作用。

生物信息学是把现存的不同新、旧数据整合的非常重要的工具。它能够模拟分子系统的结构、功能和动力学机制, 因此对于提出假设和推动试验进程都非常有帮助。

#### 插文 78

#### 分子生物学数据库

##### 一些现有的收集分子生物学信息的数据库

##### DNA 序列数据库

- 欧洲分子生物学实验室 (EMBL): <http://www.ebi.ac.uk/embl/index.html>
- 基因库: <http://www.ncbi.nlm.nih.gov/>
- 日本 DNA 数据库 (DDBJ): <http://www.ddbj.nig.ac.jp>

##### 蛋白质数据库:

- 国际蛋白质数据库 SWISS-PROT: <http://www.expasy.ch/sprot/sprot-top.html>
- 蛋白质信息资源 (PIR): <http://pir.georgetown.edu/pirwww/>
- 蛋白质数据库 (PDB): <http://www.rcsb.org/pdb/>

##### 基因鉴定利用站: Bio-Portal

- 基因组网站: <http://www.hgmp.mrc.ac.uk/GenomeWeb/nuc-geneid.html>
- BCM 搜索引擎: <http://searchlauncher.bcm.tmc.edu/>
- MOLBIOL: <http://www.molbiol.net/>
- 佩德罗生物分子研究工具: [http://www.biophys.uni-duesseldorf.de/BioNet/Pedro/research\\_tools.html](http://www.biophys.uni-duesseldorf.de/BioNet/Pedro/research_tools.html)
- ExPASy 分子生物学服务器: <http://www.expasy.ch/>

##### 畜禽特殊兴趣数据库:

- <http://locus.jouy.inra.fr/cgi-bin/bovmap/intro.pl>
- <http://www.cgd.csiro.au/cgd.html>
- <http://www.ni.bbsrc.ac.uk/cgi-bin/arkdb/browsers/>
- <http://www.marc.usda.gov/genome/genome.html>
- <http://www.ncbi.nlm.nih.gov/genome/guide/pig/>
- <http://www.ensembl.org/index.html>
- <http://www.tigr.org/>
- <http://omia.angis.org.au/>
- <http://www.livestockgenomics.csiro.au/ibiss/>
- <http://www.thearkdb.org/>
- <http://www.hgsc.bcm.tmc.edu/projects/bovine/>

## 第四部分

## 5 结论

分子特征在揭开历史、鉴定多样性、特异性和动物遗传资源结构方面都起到了非常重要的作用。它能够对小种群动物提供一个辅助的作用以防止种内过渡的近亲繁殖。大量的调查结果表明,种群内和种群间的多样性,有一些是在相当大的范围内。但是,这些研究很琐碎,很难进行比较和综合。而且,相应品种在世界范围内广泛的调查还没有进行。同样,开发将现存的、部分重叠的数据整合的方法具有非常重要的战略意义,这样能够使将来在世界范围用作参考数据的样本和标记得到保证。一个可以方便收集本土种质资源的网络可以使世界范围内的调查变得方便,通过适当管理下可以为科学团体所利用。

分子标记技术正日新月异地发展,微卫星标记很可能成为 SNPs 技术的补充。由于这些分子标记在基因组中大量存在,并且又适应于自动生产和评分,所以具有光明的前景。但是,利用 SNP 的方法进行动物品种多样性分析的可信度还需要进行进一步探究。为了避免实验结果偏离真值,这个课题需要利用充分的批判性剖析方法进行处理。

数据分析方法也在不断进步。新的分析方法可以直接进行多态性研究,而不必按照先前的通过调查做出的种群结构假设,鉴定适应基因的多样性探索(例如:

利用种群遗传学,见插文 77) 以及不同来源信息的综合,包括社会经济学参数和环境参数,来设定保存优先序(见第六章)。正确的取样策略的采用和表型信息和环境信息的系统收集仍然是开发具有潜力的新技术和新方法的必要要求。

除中性变异之外,目前开展的研究正积极寻找影响主要性状的基因。与疾病抗性、繁殖率和产品品质均是具有最高优先权的性状。大量的策略和高通量组学技术最终都应用于此。QTN 的鉴定为动物遗传资源管理提供了新的机遇和挑战。适应性多样性信息可以在表型和中性遗传多样性等方面对 QTN 进行补充,可以整合到动物遗传资源管理和保存的决定工具中去。特定种群中适应性性状的独特等位基因的鉴定或重组可能会加强它们保存工具和目的性状利用的可信度。基因辅助选择还具有减少现存的选择效率在产业化生产系统的大种群与不能有效应用的遗传评估系统和育种规划的本地小种群之间的差距的潜力。然而,标记和基因辅助选择也不总是最好的方法。考虑到其对近亲繁殖的种群结构和比例上长期和短期的影响,还要考虑到环境、社会经济条件的花费和收益,尤其是对人类生活的影响,这些选择需要基于具体情况进行评估和优化。

对于其他的先进技术,我们希望它们能够使世界范围内的分子特征领域的科学进展受益。因此,为了人类的现在和将来着想,应提高利用和保存世界动物遗传资源的认识。

## 插图 79

## 术语：分子标记

本部分用到的定义如下：

**候选基因：**任何可能引起动物可见特征改变的基因（例如：疾病抗性，牛奶蛋白的生产或生长）。由于这种基因位于染色体的特殊区域，可能与性状的控制有关，或者是它的蛋白质产物与控制性状相关（例如：牛奶蛋白生产基因）。

**DNA：**基因组中由脱氧核糖核酸编码的遗传信息，储存在细胞核中。DNA 是由脱氧核糖、磷酸、四种碱基（腺嘌呤、鸟嘌呤、胞嘧啶、胸腺嘧啶）组成，具有两条双螺旋结构的链。一条链上的 A 与另一条链上的 T 通过两个氢键相连，C 与 G 配对，通过三条氢键相连，因此两条链是彼此互补的。

**互补 DNA (cDNA)：**是由 mRNA 反转录产生的 DNA 序列。此类 DNA 含有外显子和 5',3' 区域的非编码区，但是不含有内含子 DNA。

**遗传标记：**可以很容易地通过分子或者表型分析得到的 DNA 多态性。这种标记可以存在于基因中，或者存在于未知功能的 DNA 序列中。由于 DNA 片段与染色体上与其相邻的 DNA 片段一起遗传，因此标记可以作为间接的方法，用来跟踪那些没有被定义，但是大概位置已知的基因。

**单倍型：“单倍体基因型”**的缩写，是个体染色体的遗传构成。对于二倍体生物，单倍型包含每个位点的一对等位基因。它也可能指与单个染色体相关联的一组标记（例如：单核苷酸多态性 - SNPs）。有了这个知识，就可以认为鉴定一个单倍型区域上的几个等位基因就可以明确地鉴定这个区域位点的其他多态性。这种信息对于调查复杂性状的遗传性

非常有价值。

**连锁：**在同一个染色体上相互关联的基因或者标记。连锁的基因或标记具有共同遗传的倾向。

**连锁不平衡 (LD)：**用于研究两个或多个位点（不一定要在相同的染色体上）的非随机相关联的等位基因的种群遗传学术语。与连锁不同，连锁不平衡所描述的染色体上的两个或多个位点有重组方面的限制。连锁不平衡描述的是这样一种情况，某种群中等位基因或者遗传标记所发生的重组事件的频率多于或者少于理想上随机发生的等位基因的重组率。连锁不平衡是由于基因之间适当的交换或者是一些非适应性的过程引起的。例如种群的结构，近亲繁殖和一些随机的影响。在种群遗传学中，连锁不平衡是用来说明单倍型在两个或多个位点上的分布特征的。

**基因芯片技术：**是指可以同时研究大量相互作用的基因和细胞中调控网络中各条通道上的基因的一种新方法。这种方法可以利用机器人准确地将包含有功能 DNA 的样品点到玻璃介质上。研究人员将荧光标记粘在他们想要研究细胞中的 mRNA 或 cDNA 上。这些被标记的探针可以与介质上的 cDNA 链相结合。然后将介质进行显微扫描，可以衡量每个荧光点的亮度，亮度的强弱可以说明与 mRNA 特意结合的程度。

**引物：**用于进行聚合酶链式反应的一段单链寡核苷酸序列。

**RNA：**核糖核苷酸。含有组成 DNA 的四种碱基中三种碱基的单核苷酸链（A、C、G）。RNA 中，T 被 U 代替。

## 第四部分

## 参考文献

- Aebersold, R. & Mann, M. 2003. Mass spectrometry-based proteomics. *Nature*, 422 (6928): 198–207. Review.
- Ajmone-Marsan, P., Negrini, R., Milanese, E., Bozzi, R., Nijman, I.J., Buntjer, J.B., Valentini, A. & Lenstra, J.A. 2002. Genetic distances within and across cattle breeds as indicated by biallelic AFLP markers. *Animal Genetics*, 33: 280–286.
- Akey, J.M., Zhang, G., Zhang, K., Jin, L. & Shriver, M.D. 2002. Interrogating a high-density SNP map for signatures of natural selection. *Genome Research*, 12(12): 1805–14.
- Aravin, A. & Tuschl, T. 2005. Identification and characterization of small RNAs involved in RNA silencing. *Febs Letters*, 579(26): 5830–40.
- Bachem, C.W.B., Van der Hoeven, R.S., De Bruijn, S.M., Vreugdenhil, D., Zabeau, M. & Visser, R.G.F. 1996. Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analyses of gene expression during potato tuber development. *The Plant Journal*, 9: 745–753.
- Bamshad, M. & Wooding, S.P. 2003. Signatures of natural selection in the human genome. *Nature Reviews Genetics*, 4(2): 99–111. Review.
- Baumung, R., Simianer, H. & Hoffmann, I. 2004. Genetic diversity studies in farm animals – a survey. *Journal of Animal Breeding and Genetics*, 121: 361–373.
- Beaumont, M.A. & Balding, D.J. 2004. Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, 13(4): 969–80.
- Beja-Pereira, A., Alexandrino, P., Bessa, I., Carretero, Y., Dunner, S., Ferrand, N., Jordana, J., Laloe, D., Moazami-Goudarzi, K., Sanchez, A. & Cañon, J. 2003. Genetic characterization of southwestern European bovine breeds: a historical and biogeographical reassessment with a set of 16 microsatellites. *Journal of Heredity*, 94: 243–50.
- Berthier, D., Quere, R., Thevenon, S., Belemsaga, D., Piquemal, D., Marti, J. & Maillard, J.C. 2003. Serial analysis of gene expression (SAGE) in bovine trypanotolerance: preliminary results. *Genetics Selection Evolution*, 35 (Suppl. 1): S35–47.
- Bertone, P., Stolc, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X., Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S., Gerstein, M. & Snyder, M. 2004. Global identification of human transcribed sequences with genome tiling arrays. *Science*, 306: 2242–2246.
- Black, W.C., Baer, C.F., Antolin, M.F. & DuTeau, N.M. 2001. Population genomics: genome-wide sampling of insect populations. *Annual Review of Entomology*, 46: 441–469.
- Bruford, M.W., Bradley, D.G. & Luikart, G. 2003. DNA markers reveal the complexity of livestock domestication. *Nature Reviews Genetics*, 4: 900–910.
- Buntjer, J.B., Otsen, M., Nijman, I.J., Kuiper, M.T. & Lenstra, J.A. 2002. Phylogeny of bovine species based on AFLP fingerprinting. *Heredity*, 88: 46–51.
- Campbell, D. & Bernatchez, L. 2004. Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Molecular Biology and Evolution*, 21(5): 945–56.
- Cañon, J., Garcia, D., Garcia-Atance, M.A., Obexer-Ruff, G., Lenstra, J.A., Ajmone-Marsan, P., Dunner, S. & The ECONOGENE Consortium. 2006. Geographical partitioning of goat diversity in Europe and the Middle East. *Animal Genetics*, 37: 327–334.
- Chen, S.Y., Su, Y.H., Wu, S.F., Sha, T. & Zhang, Y.P. 2005. Mitochondrial diversity and phylogeographic structure of Chinese domestic goats. *Molecular Phylogenetics and Evolution*, 37: 804–814.
- Clark, A.G., Hubisz, M.J., Bustamante, C.D., Williamson, S.H. & Nielsen, R. 2005. Ascertainment bias in studies of human genome-wide polymorphism. *Genome Research*, 15: 1496–1502.

- Clop, A., Marcq, F., Takeda, H., Pirottin, D., Tordoir, X., Bibe, B., Bouix, J., Caiment, F., Elsen, J.M., Eychenne, F., Larzul, C., Laville, E., Meish, F., Milenkovic, D., Tobin, J., Charlier, C. & Georges, M. 2006. A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nature Genetics*, 38: 813–818.
- De Marchi, M., Dalvit, C., Targhetta, C. & Cassandro, M. 2006. Assessing genetic diversity in indigenous Veneto chicken breeds using AFLP markers. *Animal Genetics*, 37: 101–105.
- Donson, J., Fang, Y., Espiritu-Santo, G., Xing, W., Salazar, A., Miyamoto, S., Armendarez, V. & Volkmuth, W. 2002. Comprehensive gene expression analysis by transcript profiling. *Plant Molecular Biology*, 48: 75–97.
- Excoffier, L., Smouse, P.E. & Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131: 479–491.
- Farnir, F., Grisart, B., Coppieters, W., Riquet, J., Berzi, P., Cambisano, N., Karim, L., Mni, M., Moiso, S., Simon, P., Wagenaar, D., Vilkki, J. & Georges, M. 2002. Simultaneous mining of linkage and linkage disequilibrium to fine map quantitative trait loci in outbred half-sib pedigrees: revisiting the location of a quantitative trait locus with major effect on milk production on bovine chromosome 14. *Genetics*, 161: 275–287.
- Fields, S. & Song, O. 1989. A novel genetic system to detect protein–protein interactions. *Nature*, 340: 245–246.
- Freeman, A.R., Bradley, D.G., Nagda, S., Gibson, J.P. & Hanotte, O. 2006. Combination of multiple microsatellite data sets to investigate genetic diversity and admixture of domestic cattle. *Animal Genetics*, 37: 1–9.
- Goldstein, D.B., Linares, A.R., Cavalli-Sforza, L.L. & Feldman, M.W. 1995. An evaluation of genetic distances for use with microsatellite loci. *Genetics*, 139: 463–471.
- Goldstein, D.B. & Schlötterer, C. 1999. *Microsatellites: evolution and applications*. New York: Oxford University Press.
- Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., Spelman, R., Georges, M. & Snell, R. 2002. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Research*, 12: 222–231.
- Guo, J., Du, L.X., Ma, Y.H., Guan, W.J., Li, H.B., Zhao, Q.J., Li, X. & Rao, S.Q. 2005. A novel maternal lineage revealed in sheep (*Ovis aries*). *Animal Genetics*, 36: 331–336.
- Haley, C. & de Koning, D.J. 2006. Genetical genomics in livestock: potentials and pitfalls. *Animal Genetics*, 37(Suppl 1): 10–12.
- Hanotte, O., Bradley, D.G., Ochieng, J.W., Verjee, Y. & Hill, E.W. 2002. African pastoralism: genetic imprints of origins and migrations. *Science*, 296: 336–339.
- Hayes, B.J., Visscher, P.M., McPartlan, H.C. & Goddard, M.E. 2003. A novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Research*, 13: 635–643.
- Hill, E.W., O’Gorman, G.M., Agaba, M., Gibson, J.P., Hanotte, O., Kemp, S.J., Naessens, J., Coussens, P.M. & MacHugh, D.E. 2005. Understanding bovine trypanosomiasis and trypanotolerance: the promise of functional genomics. *Veterinary Immunology and Immunopathology*, 105: 247–258.
- Hill, W.G. 1981. Estimation of effective population size from data on linkage disequilibrium. *Genetics Research*, 38: 209–216.
- Hillel, J., Groenen, M.A., Tixier-Boichard, M., Korol, A.B., David, L., Kirzhner, V.M., Burke, T., Barre-Dirie, A., Crooijmans, R.P., Elo, K., Feldman, M.W., Freidlin, P.J., Maki-Tanila, A., Oortwijn, M., Thomson, P., Vignal, A., Wimmers, K. & Weigend, S. 2003. Biodiversity of 52 chicken populations assessed by microsatellite typing of DNA pools. *Genetics Selection Evolution*, 35: 533–557.

## 第四部分

- Hood, L., Heath, J.R., Phelps, M.E. & Lin, B. 2004. Systems biology and new technologies enable predictive and preventative medicine. *Science*, 306: 640–643.
- Ibeagha-Awemu, E.M., Jann, O.C., Weimann, C. & Erhardt, G. 2004. Genetic diversity, introgression and relationships among West/Central African cattle breeds. *Genetics Selection Evolution*, 36: 673–690.
- Jarne, P. & Lagoda, P.J.L. 1996. Microsatellites, from molecules to populations and back. *Tree*, 11: 424–429.
- Joshi, M.B., Rout, P.K., Mandal, A.K., Tyler-Smith, C., Singh, L. & Thangaraj, K. 2004. Phylogeography and origin of Indian domestic goats. *Molecular Biology and Evolution*, 21: 454–462.
- Joost, S. 2006. *The geographical dimension of genetic diversity: A GIScience contribution for the conservation of animal genetic resources*. École Polytechnique Fédérale de Lausanne, Switzerland. (PhD thesis)
- Kayser, M., Brauer, S. & Stoneking, M. 2003. A genome scan to detect candidate regions influenced by local natural selection in human populations. *Molecular Biology and Evolution*, 20: 893–900.
- Lai, S.J., Liu, Y.P., Liu, Y.X., Li, X.W. & Yao, Y.G. 2006. Genetic diversity and origin of Chinese cattle revealed by mtDNA D-loop sequence variation. *Molecular Phylogenetics and Evolution*, 38: 146–54.
- Lan, L., Chen, M., Flowers, J.B., Yandell, B.S., Stapleton, D.S., Mata, C.M., Ton-Keen Mui, E., Flowers, M.T., Schueler, K.L., Manly, K.F., Williams, R.W., Kendzierski, C. & Attie, A.D. 2006. Combined expression trait correlations and expression quantitative trait locus mapping. *PLoS Genetics*, 2: 51–61.
- Liang, P. & Pardee, A.B. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*, 257: 967–997.
- Liu, Y.P., Wu, G.S., Yao, Y.G., Miao, Y.W., Luikart, G., Baig, M., Beja-Pereira, A., Ding, Z.L., Palanichamy, M.G. & Zhang, Y.P. 2006. Multiple maternal origins of chickens: out of the Asian jungles. *Molecular Phylogenetics and Evolution*, 38: 12–19.
- Lueking, A., Possling, A., Huber, O., Beveridge, A., Horn, M., Eickhoff, H., Schuchardt, J., Lehrach, H. & Cahill, D.J. 2003. A nonredundant human protein chip for antibody screening and serum profiling. *Molecular and Cellular Proteomics*, 2: 1342–1349.
- Luikart, G., England, P.R., Tallmon, D., Jordan, S. & Taberlet, P. 2003. The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, 4: 981–994.
- Luikart, G., Gielly, L., Excoffier, L., Vigne, J.D., Bouvet, J. & Taberlet, P. 2001. Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proceedings of the National Academy of Science USA*, 98: 5927–5932.
- Mburu, D.N., Ochieng, J.W., Kuria, S.G., Jianlin, H. & Kaufmann, B. 2003. Genetic diversity and relationships of indigenous Kenyan camel (*Camelus dromedarius*) populations: implications for their classification. *Animal Genetics*, 34(1): 26–32.
- McPherron, A.C. & Lee, S.J. 1997. Double muscling in cattle due to mutations in the myostatin gene. *Proceedings of the National Academy of Science USA*, 94: 12457–12461.
- Negrini, R., Milanese, E., Bozzi, R., Pellecchia, M. & Ajmone-Marsan, P. 2006. Tuscany autochthonous cattle breeds: an original genetic resource investigated by AFLP markers. *Journal of Animal Breeding and Genetics*, 123: 10–16.
- Nei, M. 1972. Genetic distance between populations. *The American Naturalist*, 106: 283–292.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89: 583–590.
- Nei, M. & Roychoudhury, A.K. 1974. Sampling variances of heterozygosity and genetic distance. *Genetics*, 76: 379–390.

- Nei, M., Tajima, F. & Tateno, Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *Journal of Molecular Evolution*, 19: 153–170.
- Nielsen, R. & Signorovitch, J. 2003. Correcting for ascertainment biases when analyzing SNP data: applications to the estimation of linkage disequilibrium. *Theoretical Population Biology*, 63: 245–55.
- Nijman, I.J., Otsen, M., Verkaar, E.L., de Ruijter, C. & Hanekamp, E. 2003. Hybridization of banteng (*Bos javanicus*) and zebu (*Bos indicus*) revealed by mitochondrial DNA, satellite DNA, AFLP and microsatellites. *Heredity*, 90: 10–16.
- Pariset, L., Cappuccio, I., Joost, S., D'Andrea, M.S., Marletta, D., Ajmone Marsan, P., Valentini A. & ECONOGENE Consortium 2006. Characterization of single nucleotide polymorphisms in sheep and their variation as an evidence of selection. *Animal Genetics*, 37: 290–292.
- Pritchard, J.K., Stephens, M. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155: 945–959.
- Rabie, T.S., Crooijmans, R.P., Bovenhuis, H., Vereijken, A.L., Veenendaal, T., van der Poel, J.J., Van Arendonk, J.A., Pakdel, A. & Groenen, M.A. 2005. Genetic mapping of quantitative trait loci affecting susceptibility in chicken to develop pulmonary hypertension syndrome. *Animal Genetics*, 36: 468–476.
- Saitou, N. & Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4: 406–425.
- Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L., Hunt, S.E., Cole, C.G., Coggill, P.C., Rice, C.M., Ning, Z., Rogers, J., Bentley, D.R., Kwok, P.Y., Mardis, E.R., Yeh, R.T., Schultz, B., Cook, L., Davenport, R., Dante, M., Fulton, L., Hillier, L., Waterston, R.H., McPherson, J.D., Gilman, B., Schaffner, S., Van Etten, W.J., Reich, D., Higgins, J., Daly, M.J., Blumenstiel, B., Baldwin, J., Stange-Thomann, N., Zody, M.C., Linton, L., Lander, E.S. & Altshuler, D.; International SNP Map Working Group. 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, 409: 928–933.
- SanCristobal, M., Chevalet, C., Haley, C.S., Joosten, R., Rattink, A.P., Harlizius, B., Groenen, M.A., Amigues, Y., Boscher, M.Y., Russell, G., Law, A., Davoli, R., Russo, V., Desautels, C., Alderson, L., Fimland, E., Bagga, M., Delgado, J.V., Vega-Pla, J.L., Martinez, A.M., Ramos, M., Glodek, P., Meyer, J.N., Gandini, G.C., Matassino, D., Plastow, G.S., Siggins, K.W., Laval, G., Archibald, A.L., Milan, D., Hammond, K. & Cardellino, R. 2006a. Genetic diversity within and between European pig breeds using microsatellite markers. *Animal Genetics*, 37: 189–198.
- SanCristobal, M., Chevalet, C., Peleman, J., Heuven, H., Brugmans, B., van Schriek, M., Joosten, R., Rattink, A.P., Harlizius, B., Groenen, M.A., Amigues, Y., Boscher, M.Y., Russell, G., Law, A., Davoli, R., Russo, V., Desautels, C., Alderson, L., Fimland, E., Bagga, M., Delgado, J.V., Vega-Pla, J.L., Martinez, A.M., Ramos, M., Glodek, P., Meyer, J.N., Gandini, G., Matassino, D., Siggins, K., Laval, G., Archibald, A., Milan, D., Hammond, K., Cardellino, R., Haley, C. & Plastow, G. 2006b. Genetic diversity in European pigs utilizing amplified fragment length polymorphism markers. *Animal Genetics*, 37: 232–238.
- Sauer, S., Lange, B.M.H., Gobom, J., Nyarsik, L., Seitz, H. & Lehrach, H. 2005. Miniaturization in functional genomics and proteomics. *Nature Reviews Genetics*, 6: 465–476.
- Sodhi, M., Mukesh, M., Mishra, B.P., Mitkari, K.R., Prakash, B. & Ahlawat, S.P. 2005. Evaluation of genetic differentiation in *Bos indicus* cattle breeds from Marathwada region of India using microsatellite polymorphism. *Animal Biotechnology*, 16: 127–137.
- Storz, G., Altuvia, S. & Wassarman, K.M. 2005. An abundance of RNA regulators. *Annual Review of Biochemistry*, 74: 199–217.
- Sunnucks, P. 2001. Efficient genetic markers for population biology. *Tree*, 15: 199–203.

## 第四部分

- Syvänen, A.C. 2001. Accessing genetic variation genotyping single nucleotide polymorphisms. *Nature Reviews Genetics*, 2: 930–941.
- Takezaki, N. & Nei, M. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, 144: 389–399.
- Tapio, M., Tapio, I., Grisli, Z., Holm, L.E., Jeppsson, S., Kantanen, J., Miceikiene, I., Olsaker, I., Viinalass, H. & Eythorsdottir, E. 2005. Native breeds demonstrate high contributions to the molecular variation in northern European sheep. *Molecular Ecology*, 14: 3951–3963.
- Tilquin, P., Barrow, P.A., Marly, J., Pitel, F., Plisson-Petit, F., Velge, P., Vignal, A., Baret, P.V., Bumstead, N. & Beaumont, C. 2005. A genome scan for quantitative trait loci affecting the *Salmonella* carrier-state in the chicken. *Genetics Selection Evolution*, 37: 539–61.
- Troy, C.S., MacHugh, D., Bailey, J.F., Magee, D.A., Loftus, R.T., Cunningham, P., Chamberlain, A.T., Sykes, B.C. & Bradley D.G. 2001. Genetic evidence for Near-Eastern origins of European cattle. *Nature*, 410: 1088–1091.
- Velculescu, V.E., Vogelstein, B. & Kinzler, K.W. 2000. Analyzing uncharted transcriptomes with SAGE. *Trends in Genetics*, 16: 423–425.
- Velculescu, V.E., Zhang, L., Vogelstein, B. & Kinzler, K.W. 1995. Serial analysis of gene expression. *Science*, 270: 484–487.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J. & Kuiper, M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23: 4407–1444.
- Weir, B.S. & Basten, C.J. 1990. Sampling strategies for distances between DNA sequences. *Biometrics*, 46: 551–582.
- Weir, B.S. & Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, 38: 1358–1370.
- Wienholds, E. & Plasterk, R.H. 2005. MicroRNA function in animal development. *FEBS Letters*, 579: 5911–5922.
- Wong, G.K., Liu, B., Wang, J., Zhang, Y., Yang, X., Zhang, Z., Meng, Q., Zhou, J., Li, D., Zhang, J., Ni, P., Li, S., Ran, L., Li, H., Zhang, J., Li, R., Li, S., Zheng, H., Lin, W., Li, G., Wang, X., Zhao, W., Li, J., Ye, C., Dai, M., Ruan, J., Zhou, Y., Li, Y., He, X., Zhang, Y., Wang, J., Huang, X., Tong, W., Chen, J., Ye, J., Chen, C., Wei, N., Li, G., Dong, L., Lan, F., Sun, Y., Zhang, Z., Yang, Z., Yu, Y., Huang, Y., He, D., Xi, Y., Wei, D., Qi, Q., Li, W., Shi, J., Wang, M., Xie, F., Wang, J., Zhang, X., Wang, P., Zhao, Y., Li, N., Yang, N., Dong, W., Hu, S., Zeng, C., Zheng, W., Hao, B., Hillier, L.W., Yang, S.P., Warren, W.C., Wilson, R.K., Brandstrom, M., Ellegren, H., Crooijmans, R.P., van der Poel, J.J., Bovenhuis, H., Groenen, M.A., Ovcharenko, I., Gordon, L., Stubbs, L., Lucas, S., Glavina, T., Aerts, A., Kaiser, P., Rothwell, L., Young, J.R., Rogers, S., Walker, B.A., van Hateren, A., Kaufman, J., Bumstead, N., Lamont, S.J., Zhou, H., Hocking, P.M., Morrice, D., de Koning, D.J., Law, A., Bartley, N., Burt, D.W., Hunt, H., Cheng, H.H., Gunnarsson, U., Wahlberg, P., Andersson, L., Kindlund, E., Tammi, M.T., Andersson, B., Webber, C., Ponting, C.P., Overton, I.M., Boardman, P.E., Tang, H., Hubbard, S.J., Wilson, S.A., Yu, J., Wang, J., Yang, H.; International Chicken Polymorphism Map Consortium. 2004. A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature*, 432: 717–722.
- Zhu, H., Bilgin, M. & Snyder, M. 2003. Proteomics. *Annual Review of Biochemistry*, 72: 783–812.