

鸡基因组研究进展*

顾志良²⁾ 张勇¹⁾ 朱大海^{1)**} 李辉²⁾

(¹⁾ 哈尔滨工业大学生命科学与工程系, 哈尔滨 150001; ²⁾ 东北农业大学动物科技学院, 哈尔滨 150030)

摘要 随着人类基因组计划实施而开展的动物基因组计划受到了科学界和各国政府的支持. 无论是作为一种实验用模式生物, 还是作为一种农业经济动物, 鸡都有着独特的生物学特性和经济学价值, 因此, 开展鸡基因组研究是十分有意义的. 综述了近年来鸡基因组研究 (包括鸡基因组的有关参数、遗传连锁图、物理图谱、比较定位、表达序列标签和数量性状座位定位等方面) 所取得的成就并对其前景进行了讨论.

关键词 鸡, 基因组, 比较基因组学

学科分类号 Q343

随着人类基因组计划的实施而开展的动物基因组计划受到了科学界和各国政府的支持, 欧盟和美国等都启动了动物基因组计划. 鸡不仅是全世界广泛饲养并具有重要经济价值的禽类, 而且还是生命科学极有价值的模式动物, 越来越引起科学家的重视. 欧美等发达国家推出了国际鸡基因组作图计划. 1992 年英国动物健康研究院 Bamstead 等^[1] 率先完成了利用限制性酶切片长度多态性 (RFLP) 技术构建鸡基因组的初始连锁图, 图谱包含了 100 个 DNA 标记, 分布于 18 个连锁群, 覆盖了 585 cM, 这是鸡的第一张分子标记连锁图. 随后 1994 年 Levin 等^[2] 发表的鸡常染色体遗传连锁图包含了 98 个 DNA 标记, 覆盖了 590 cM. 近几年来由于科学家们的努力, 世界各国实验室共同的合作, 加之生物技术手段的发展, 使鸡基因组的研究取得了重大的进展. 本文将从鸡的基因连锁图谱、物理图谱、比较定位和数量性状座位 (QTL) 定位等方面对这一领域中所取得的成果进行综述.

1 鸡基因组的组成和结构

鸡共有 39 对染色体, 根据大小可分为大、中和微型染色体. 大型染色体利用分带技术很容易区分, 而一些微小染色体不能用细胞学方法来分辨识别. 大型染色体标准化分带核型已于 1993 年由鸟类核型标准化国际委员会确认. 现在, 寄希望于 cDNA, 粘粒 (Cosmid) 和酵母人工染色体 (YAC) 克隆作探针进行荧光原位杂交来分析微型染色体. Rodionov 等 (1992 年) 报道鸡的基因组 DNA 总长度为 1.2×10^9 bp, 遗传长度为 2 950~ 3 200 cM, 其中最大 5 对染色体的遗传长度约占 40%, Z 染色体遗传长度

为 145 cM, 而 W 染色体的遗传长度只有 50 cM, 鸡的每 100 万个碱基对的平均距离为 2.4~ 2.6 cM, 大约相当于人类雌性基因组平均遗传距离的 3 倍^[3]. 根据 East Lansing (E) 和 Compton (C) 连锁图谱估计鸡基因组总的遗传长度为 3 800 cM^[4], Bloom (1993 年) 的研究结果为 8 对大型染色体和 Z 及 W 性染色体占总 DNA 的 72%, 而 Smith^[4] 的结果是这些染色体占 82%. E 遗传连锁图的 1~ 8 条大染色体 1 cM 相当于 440 kb, 而 C 图谱中 1 cM = 350 kb, 平均 1 cM = 396 kb^[4].

2 遗传图谱

鸡染色体图谱的研究可以追溯到 1908 年, Spillman 揭示了斑纹 (barring) 基因是性连锁的, 标志着鸡染色体图研究的开始. 1936 年, Hutt 发表了鸡的第一个基因连锁图, 这也是畜禽的第一张基因连锁图, 这张图谱仅包含了 5 个连锁群的 18 个座位^[5]. 利用经典的基因图谱构建方法, 对常染色体和性染色体上的许多基因作了定位, 一些与生长和发育有关的突变表型如矮小型、快慢羽已经定位在 Z 染色体上, 而且集中分布在 19 cM 范围内, 在 5 个连锁群中大约有 50 个基因, 其中 3 个连锁群在同一条染色体上. 利用基因操作技术研究鸡的基因图谱在 20 世纪 80 年代初期开始有零星的

* 国家杰出青年基金 (300250) 和国家自然科学基金 (30070552) 资助项目.

** 通讯联系人. 哈尔滨工业大学生命科学与工程系分子细胞发育生物学实验室.

Tel: 0451-6412865, E-mail: dahaizhu@hotmail.com

收稿日期: 2001-11-20, 接受日期: 2002-01-31

报道, 但其真正在世界范围内引起人们注意的是在 80 年代末期.

在鸡基因图谱的构建中, 一个重要的里程碑是在英国和美国建立了两个特殊的回交群体. 1992 年, 英国动物健康研究院 Bumstead 等^[1]采用抗病性不同的两个白来航近交系建立了 Compton 群体. 美国密歇根州立大学 Crittenden 等^[6]用一只部分近交的红色原鸡公鸡与一只高度近交的白来航母鸡杂交后代回交后建立了 East Lansing 群体. 这两个回交群在第一届国际遗传学会家禽基因定位研讨会上作为国际参照候选群体得到公认. 还有一个 Wageningen (WAU) 资源群体是基于 3 代全同胞半同胞群体, F1 是两个极端的白洛克肉鸡品系的杂交后代, F1 和 F2 代测定基因型, 用来建立 WAU 连锁图, F3 代测定表现型 (生产性能), 这些数据用于 QTL 的分析^[7]. 目前围绕三个参照群已建立了三张连锁图 (C 连锁图、E 连锁图和 WAU 连锁图).

近年来遗传图谱的进展主要表现在分子遗传学标记, 特别是微卫星标记和扩增片段长度多态性 (AFLP) 标记不断发现, 遗传图谱上定位的遗传标记不断增加, 标记密度不断提高. Cheng 等^[8] (1995 年) 对遗传图谱进行了补充, 使得 E 参照系上的标记数达到 273 个 (增加了 86 个微卫星标记), 其中 243 个分布于 32 个连锁群, 连锁群覆盖了 1 420 cM, 平均间距为 6.7 cM. Groenen 领导的研究小组在开发和定位鸡的微卫星标记做了大量工作, 开发和定位了大量的微卫星标记^[7]. 1998 年 Groenen^[9]报道的综合微卫星连锁图上的座位数为 1 364 个 (其中微卫星标记 430 个), 分布在 27 条常染色体、连锁群和 Z 染色体上的遗传总长度为 3 062 cM, 450 个标记是可识别的基因或表达序列标签 (EST). Knorr 等^[10] (1999 年) 用 AFLP 技术在 E 参照群的图谱中来增加标记密度, 204 个 AFLP 标记被定位在 E 图谱上, 使得 E 图谱上的标记增加 25%; Herbergs 等^[11] (1999 年) 在 WAU 家系中发现了 475 个多态的 AFLP 标记, 其中 344 个被定位在 WAU 连锁图上, 这样 AFLP、微卫星连锁图在 33 个连锁群中有 805 个座位. 最近报道 E 群体上的 42 个连锁群共分布了 1 100 个标记 (其中 479 个为微卫星座位), 202 个 I 型标记的定位对开发比较定位带来了便利, 这个图谱还包含了 200 个 AFLP 标记, 连锁群共覆盖了 3 700 cM; WAU 连锁图有 573 个微卫星标记. 基于 3 个群体

的第二代综合图谱 (consensus map) 包含了 1 889 个标记, 分布于 50 个连锁群, 覆盖了鸡基因组的 3 800 cM^[12].

3 物理图谱

建立物理图谱的常用方法为染色体原位杂交 (CISH) 和体细胞杂交, 最理想的是用荧光染色体原位杂交 (FISH) 技术. 有一些连锁群通过原位杂交已经被定位在染色体上, 大染色体核型标准化推动了物理图谱的建立. 微小染色体的基因定位出现的一些问题主要是采用过去的分带技术产生非特异性的结果, 采用 cDNA、Cosmid、细菌人工染色体 (BAC)、噬菌体人工染色体 (PAC) 和 YAC 克隆作探针荧光原位杂交可以解决这一问题, 另一问题是常发生单一的染色体交叉, 而且 50% 的这种交叉出现在视野区域, 这样微小染色体不能出现所期望的重组, 只能检测出单点连锁群^[3]. 鸡物理图谱的构建要做的工作包括大片段插入克隆的文库 (YAC 和 BAC)、体细胞杂交系和放射性杂交系的建立. 目前鸡 YAC 文库和 BAC、PAC 文库均已建成.

Morrisson^[13] (1998 年) 报道了将鸡的 BAC 和 PAC 文库克隆, 用原位杂交的方法定位在染色体上, 同时也定位在 E 和 C 连锁图上, 用 18 个标记将物理图和连锁图整合 (integrated map) 起来, 7 个标记在大染色体上, 11 个标记定位在微小染色体上. Smith (2000 年) 报道了鸡的大染色体上的物理和连锁图整合结果, 根据 E、C、WAU 参照群中获得的大量遗传图谱信息, 将 14 个标记的物理和遗传图谱整合在一起, 发表了 8 条大染色体和 Z 染色体的综合图, 给出了 1~3 号染色体和 Z 染色体的物理和遗传距离的关系. 中国农业大学农业生物技术国家重点实验室也已经构建了鸡的 BAC 文库, 相信不久的将来我国学者也会在这一方面获得大的进展.

4 比较定位

早期的研究认为, 比较鸟类与哺乳动物基因的核苷酸序列会发现其同源性相差甚远, 而且基因组的长度及结构也有很大的差异, 鸟类基因组 DNA 长度一般只有哺乳动物的三分之一, 通常看来比较定位是不可取的. 但是进行鸡和哺乳动物连锁图的比较, 可发现鸟类和哺乳动物同一染色体上某些位点有保守性. 鸡 E2、C1 连锁群与人第 6 号染色体

上的 ESR 和 FYN 基因有同线性, 而人 MYB 位点与 ESR、FYN 位点连锁, 用荧光原位杂交法已知 MYB 定位在鸡的 3 号染色体上, 如果同一染色体上的 ESR、MYB、FYN 是保守的, 那么鸡的 C1 和 E2 连锁群也分布在 3 号染色体上. 而比较定位技术的成功应用是依赖于保守基因更详细图谱和基因组数据库.

利用序列的保守性, 根据人和鼠的序列保守区设计基因的引物进行比较定位, Smith (1997 年) 用这种方法将 21 个基因定位于连锁群或染色体, 5 个同线性连锁群被发现. 随着比较图谱的继续积累, 发现人和鸡的基因组基因在相当程度上是保守的, 甚至超过了人和鼠的基因组的同源性. 然而最近的研究结果证实早期的假设, 禽类性染色体进化是独立于哺乳类的性染色体^[14, 15], 而且微小染色体上的基因密度高于大染色体. 人与鸡的基因组中许多差别可通过恰当的结构比较定位来实现. Groenen 等 (1999 年) 将 15 个新的基因定位于连锁群上, 而其中 13 个基因的人同源位点是已知的, 另外 2 个基因通过原位杂交定位, 这些标记扩展了鸡和人的比较图谱. 他们的结果证实了鸡的 2、7 号染色体和 E29C09W09 连锁群上与人存在保守区, 而且在 7 号染色体上也检测到另外一个保守区.

5 表达序列标签

表达序列标签 (expressed sequence tags, ESTs) 是揭示基因组容量的一种快速而有效的方法, 在新基因的发现、基因作图和基因组序列中编码区域的确定等方面起重要作用. ESTs 序列的测定为我们提供了一个较好的建立特异组织基因表达图谱的途径. 随着功能基因组学的开展, 作为基因组学和功能基因组学之间桥梁的 ESTs 将会起到更加显著的作用. 以美国 Delaware 大学为主进行的鸡 EST 计划 (Delaware Institute Chick EST Project) 共收集了来自淋巴组织、脂肪、肝、肌肉和垂体的 50 000 多个鸡 ESTs. 中国农业大学构建了鸡下丘脑 cDNA 文库, 测定了鸡下丘脑 ESTs, 哈尔滨工业大学分子与细胞发育生物学实验室正在从事不同家鸡骨骼肌的 ESTs 研究.

6 QTL 定位及 QTL 连锁的标记

美国农业部禽病与肿瘤研究所和密歇根州立大学利用已建立的由马立克氏病易感和抗性极端差异的两个品系杂交群组成的资源群体, 检测到对马立

克氏病 (MD) 的易感性有显著影响的 5 个 QTL, 分别位于 Ch2, Ch4, Ch7, Ch8 和 E16 连锁群上^[16]; 进一步研究又发现了多个新的 QTL 与马立克氏病有关, 分别位于 Ch1, Ch2, Ch4, Ch7, Ch8 和 E41 连锁群上. 以色列的研究小组发现 2 号染色体上的 ADL0146 与 SRBC 和 NDV 的抗体量有关, 31 连锁群上的 ADL0290 影响抗 NDV 的抗体量, 34 连锁群的 ADL0298 与抗大肠杆菌的抗体和生活力有关. 通过区间定位, 进一步证实了这种显著效应和 QTL 的存在, 以及免疫抗性和疾病抗性的相关^[17].

Groenen 领导的研究小组^[18]正在大规模开展鸡的 QTL 定位研究, 通过全基因组扫描法对一些早期生长性状、饲料转化率和屠宰性状的 QTL 定位分析结果如下: 检测位于 24 个常染色体连锁群上 368 个标记, 发现影响体重最可能的 QTL 在 1 号染色体的 24 cM 处. 用全同胞区间定位的方法对 27 个常染色体上的 420 个标记检测基因型, 发现 4 个 QTL, 分别是位于 1 号染色体 235 cM 处的 23~48 日龄采食量的 QTL (进一步研究发现该 QTL 影响 23~48 日龄的生长速度及 48 日龄体重), 位于 WAU26 的 16 cM 处潜在的与 23~48 日龄采食量的 QTL, 位于 4 号染色体 147 cM 处的影响两种采食量的 QTL, 位于 2 号染色体 41 cM 处随体重调整的采食量的 QTL, 另外还检测到影响屠宰率的 QTL 在 1 号染色体的 466 cM 处, 影响肉色的 QTL 在 2 号染色体的 315~361 cM 处^[19].

中国农业大学在建立资源家系的基础上用 352 对微卫星标记分成了 44 组进行多重 PCR 的基因组扫描, F0 代、F1 代个体的标记基因分型工作已经完成, F2 个体也在进行全基因组扫描; 采用候选基因的方法发现细胞外脂肪酸结合蛋白基因为控制鸡 12 周龄腹脂重、腹脂率的主基因. 哈尔滨工业大学分子与细胞发育生物学实验室对鸡的 Myostatin 基因 5' 和 3' 侧翼序列的 SNP 分析发现了两种基因型间胸肌和胸肌率具有显著的差异. 东北农业大学通过对鸡 PPAR、UCP、OBR 基因多态性的研究, 发现 3 个突变位点对腹脂重和腹脂率有显著的影响.

7 展 望

作为一种有用的模式动物, 鸡的价值并没有被充分认识. 首先, 鸡和人的保守同线性水平非常高, 从鸡和人比较定位发现进化过程中相当部分的

染色体区段是非常保守的, 物理定位的基因中也发现许多基因是保守的, 已经定位在连锁图上的 235 个已知基因中有 204 个在人上被定位; 其次, 鸡基因组只有人的三分之一大小, 其基因组中的重复序列少, 内含子长度小, 其基因组结构有其独特性 (微小染色体中的基因密度高于大染色体). 这些特点无疑决定了鸡是除河豚、小鼠和大鼠外研究人类基因组和生物进化规律的又一个很好的模式生物. 作为鸟类的代表, 鸡的红细胞有细胞核, 方便提取 DNA; 鸡胚为体外发育, 易于获取不同发育时期的胚胎样本. 此外, 在比较基因组学的研究中, 尤其是在涉及到骨骼肌发生 (myogenesis) 和繁殖的研究中, 鸡有着其他模式动物无可比拟、不可取代的优势. 例如, 肉鸡 (在生长性状上拥有强势)、蛋鸡 (在繁殖性状上拥有强势), 在基因组水平上具有大体相同的遗传背景和与其特定性状相关的微小差异. 这为我们进行比较基因组学研究提供了良好的素材. 因此, 研究鸡基因组对我们揭示人类基因组的结构和功能、人类疾病分子机理和生物进化规律具有重要作用.

作为一种经济动物, 鸡的基因组研究最终是以重要经济性状 QTL 定位和重要基因的位置克隆为目标. 目前鸡基因图谱还有许多工作要做: 针对 QTL 定位, 虽然从基因组上的平均遗传距离来看, 标记的密度已经可满足全基因组扫描定位 QTL 的需要, 但是存在着在整个基因组中标记密度不均匀性, 有些区域过疏, 针对位置克隆仍然需要增加连锁图上的标记; E、C、WAU 三个连锁图谱中还有一些不一致的结果, 需要进一步验证, 而且 3 个图谱的进一步综合也是至关重要的; 遗传图谱和物理图谱的整合虽然进行了一些工作, 但物理图谱上定位的位点数远低于连锁图上的标记数; 鸡和人等其他生物的基因比较图谱也是需要进一步做的工作, 这是物理定位一些功能基因的很好途径; QTL 定位与猪等其他家畜相比进展较慢, 检测定位的 QTL 太少, 将要花费更大的人力和物力来进行研究.

无论是作为一种实验用模式动物, 还是作为一种经济动物, 鸡都有着自身的优势. 遗传学已进入基因组学时代, 我国有丰富的鸡品种资源, 是国际上一座极其宝贵的鸡的基因库, 进行鸡的基因组计划无疑具有其深远的意义.

参 考 文 献

- 1 Bumstead N, Palyga J. A preliminary linkage map of the chicken genome. *Genomics*, 1992, **13** (3): 690~ 697
- 2 Levin I, Santangelo L, Cheng H, *et al.* An autosomal genetic linkage map of the chicken. *J Hered*, 1994, **85** (2): 79~ 85
- 3 Burt D W, Bumstead N, Bitgood J J, *et al.* Chicken genome mapping: a new era in avian genetics. *TIG*, 1995, **11** (5): 190 ~ 194
- 4 Smith J, Burt D W. Parameters of the chicken genome (*Gallus gallus*). *Anim Genet*, 1998, **29** (4): 290~ 294
- 5 Bitgood J J. *Poultry Breeding and Genetics*. Amsterdam: Elsevier Science Publishers Ltd, 1990. 469~ 495
- 6 Crittenden L B, Provencher L, Santangelo L, *et al.* Characterization of a Red Jungle Fowl by White Leghorn backcross reference population for molecular mapping of the chicken genome. *Poult Sci*, 1993, **72** (2): 334~ 348
- 7 Crooijmans R P, van Oers P A, Strijk J A, *et al.* Preliminary linkage map of the chicken (*Gallus domesticus*) genome based on microsatellite markers: 77 new markers mapped. *Poult Sci*, 1996, **75** (6): 746~ 754
- 8 Cheng H H, Levin I, Vallejo R L, *et al.* Development of a genetic map of the chicken with markers of high utility. *Poult Sci*, 1995, **74** (11): 1855~ 1874
- 9 Groenen M A, Crooijmans R P, Veenendaal A, *et al.* A comprehensive microsatellite linkage map of the chicken genome. *Genomics*, 1998, **49** (2): 265~ 274
- 10 Knorr C, Cheng H H, Dodgson J B. Application of AFLP markers to genome mapping in poultry. *Anim Genet*, 1999, **30** (1): 28~ 35
- 11 Herbergs J, Siwek M, Crooijmans R P, *et al.* Multicolour fluorescent detection and mapping of AFLP markers in chicken (*Gallus domesticus*). *Anim Genet*, 1999, **30** (4): 274~ 285
- 12 Groenen M A, Cheng H H, *et al.* A consensus linkage map of the chicken genome. *Genome Research*, 2000, **10** (1): 137~ 147
- 13 Morisson M, Pitel F, Fillon V, *et al.* Integration of chicken cytogenetic and genetic maps: 18 new polymorphic markers isolated from BAC and PAC clones. *Anim Genet*, 1998, **29** (5): 348~ 355
- 14 Fridolfsson A K, Cheng H, Copeland N G, *et al.* Evolution of the avian sex chromosomes from an ancestral pair of autosomes. *Proc Natl Acad Sci USA*, 1998, **95** (14): 8147~ 8152
- 15 Nanda I, Shan Z, Scharl M, *et al.* 300 million years of conserved synteny between chicken Z and human chromosome 9. *Nat Genet*, 1999, **21** (3): 258~ 259
- 16 Vallejo R L, Bacon L D, Liu H C, *et al.* Genetic mapping of quantitative trait loci affecting susceptibility to Marek's disease virus induced tumors in F2 intercross chickens. *Genetics*, 1998, **148** (1): 349~ 360
- 17 Yonash N, Cheng H H, *et al.* DNA microsatellites linked to quantitative trait loci affecting antibody response and survival rate in meat-type chickens. *Poult Sci*, 2001, **80** (1): 22~ 28
- 18 Van Kaam J B, Groenen M A, Bovenhuis H, *et al.* Whole genome scan in chickens for quantitative trait loci affecting growth and feed efficiency. *Poult Sci*, 1999, **78** (1): 15~ 23
- 19 van Kaam J B, Groenen M A, Bovenhuis H, *et al.* Whole genome scan in chickens for quantitative trait loci affecting carcass traits. *Poult Sci*, 1999, **78** (8): 1091~ 1099

Progress on The Chicken Genome Project^{*}

GU ZhiLiang²⁾, ZHANG Yong¹⁾, ZHU DaHai¹⁾^{**}, LI Hui²⁾

⁽¹⁾ *Department of Life Science and Biotechnology, Harbin Institute of Technology, Harbin 150001, China;*

²⁾ *College of Animal Science & Technology, Northeast Agricultural University, Harbin 150031, China)*

Abstract With the development of the human genome project, the chicken genome research has made great progress and has significant impact on both animal breeding and basic biological research. The current development of the chicken genome research was reviewed in including the parameter of chicken genome, genetic linkage map, physical map, comparative genome map, expression sequence tag (EST) and identification of quantitative trait loci (QTL) in chicken and the future prospect of the chicken genome research was also discussed.

Key words chicken, genome, comparative genomics

^{*} This work was supported by grants from The National Foundation of Science for Outstanding Young Scientists (300250) and the National Natural Science Foundation of China (30070552).

^{**} Corresponding author. Molecular and Cellular Developmental Biology Laboratory, Department of Life Science and Biotechnology, Harbin Institute of Technology, Harbin 150001.

Tel: 86-451-6412865, E-mail: dahaizhu@hotmail.com

Received: November 20, 2001 Accepted: January 31, 2002