



蓝藻分子遗传学又十年

孔任秋 徐旭东 王业勤

(中国科学院水生生物研究所, 武汉 430072)

MOLECULAR GENETICS OF CYANOBACTERIA: THE DECADE AFTER LAST REVIEW

KONG Ren-qiu, XU Xu-dong and WANG Ye-qin

(*Institute of Hydrobiology, The Chinese Academy of Sciences, Wuhan, 430072*)

关键词: 蓝藻; 分子遗传学

Key words: Cyanobacteria; Molecular genetics

中图分类号: Q949.22 文献标识码: A 文章编号: 1000-3207(2001)06-0620-11

蓝藻是一类进行产氧光合作用的原核生物, 又称蓝细菌, 广泛分布于地球上各种类型的水环境和一些陆生环境, 是海洋和内陆水体中最重要的原初生产者。蓝藻分子遗传学在上个世纪八十年代奠定基础, 在许多方面借鉴其他类群微生物分子遗传学已发展的技术和方法, 取得了诸多成绩^[1], 但总体上处于开拓阶段。这一时期的研究方法一般是建立基因转移系统, 筛选突变株, 单个或少数几个基因的克隆和表达分析等。进入九十年代后随着被鉴别基因数量的增加和更有效的研究方法的引入, 蓝藻生物学已经能够着眼于整个基因调控途径或调控网络的解析。尤其在九十年代后五年里, 若干蓝藻全基因组序列分析的完成 (<http://www.science.richmond.edu/~biology/cyanonews>), 显示蓝藻生物学已经跟上微生物学乃至整个生物学的步伐, 跨入了基因组时代。本文将着重从分子遗传学角度阐述过去十年里这一领域的重要发展, 而不是试图包罗蓝藻分子生物学已经涉及的各个方面。有关基因调控途径或调控网络, 作者也只举出生物钟和异形胞分化这两个代表这一时期研究水平的例证。

1 基因组计划

蓝藻模式生物全基因组测序计划的完成是过去十年该领域最为重要的成就。其中,

收稿日期: 2001-07-10; 修订日期: 2001-07-23

基金项目: 中国科学院留学回国人员择优基金; 中国科学院“百人计划”项目资助

作者简介: 孔任秋, (1964—), 女, 山东济宁人; 副研究员; 从事微生物分子遗传学研究

通讯作者: 徐旭东, Email: xux@ihb.ac.cn

集胞藻(*Synechocystis* sp.)PCC6803 是单细胞、可异养生长的种类,具有天然的 DNA 转化系统,被作为光合系统分子生物学研究的重要模式^[2]。该种全基因组序列于 1996 年公布,是最早完成序列分析的几种微生物之一,而在光合生物中则是第一个。在 3.6 Mb 的基因组中,有 3168 个可能的蛋白编码基因,占基因组总长的 87.0%;还有两个 rRNA 基因丛和 42 个 tRNA 基因。基因组中有 99 个类似插入因子(IS)的结构,但只有 26 个可能编码完整的转座酶^[3]。另一种被测序的单细胞蓝藻是聚胞藻(*Synechococcus* sp.)WH8102,基因组大小为 2.72Mb,有 2426 个蛋白编码基因。该种与其他聚胞藻种类广泛分布于全世界海洋,是地球上主要的原初生产者之一;不具有任何明显的胞外结构,却能够推动自身以 25mm/sec 的速度在海水中运动;可进行遗传操作^[4]。已完成基因组测序的丝状固氮蓝藻有鱼腥藻(*Anabaena* sp.)PCC7120 和点形念珠藻(*Nostoc punctiforme*) ATCC29133。鱼腥藻 7120 基因组大小为 6.4Mb,目前正进行基因组数据的分析。该种在八十年代至九十年代初建立了较完善的遗传分析系统,是一维生物体细胞分化研究的模式^[5]。点形念珠藻除异形胞分化外,还可进行异养生长,产生厚壁孢子和藻殖段,侵染并与植物共生^[6]。在鱼腥藻 7120 发展起来的基因转移系统和转座子诱变方法均适用于该种,但转座诱变效率较低^[7]。其基因组大小为 9.2Mb,有 7432 个蛋白编码基因。在以上几例中,集胞藻 6803 基因组数据公布较早,已对蓝藻分子遗传学研究起到重要推动作用。这些种类之间以及它们和原绿藻之间的比较基因组学研究尚无报道。

2 分子剪接

2.1 DNA 重排

早在八十年代,Golden 等人即发现了鱼腥藻 7120 固氮酶基因丛在异形胞发育晚期有两个 DNA 重排发生,即在 *nifD* 基因内有一个 11kb 的因子,在 *fdxN* 基因内有一个 55kb 的因子,在各自编码的剪接酶的作用下以环状形式从基因组中脱落下来并使 *nifD* 和 *fdxN* 重新成为完整基因^[1]。九十年代以后,Golden 实验室又发现了位于吸氢酶的大亚基基因 *hupL* 内的第三个重排因子,大小为 10.5kb^[8]。与前两个重排反应相似,*hupL* 因子两侧各有一个 16bp 的序列被该因子自身携带的剪接酶基因 *xisC* 产物识别,剪接释放出 10.5kb 的环状 DNA。*hupL* 的重排也发生于异型胞分化晚期。Andrey V. Matveyev 等人以脉冲场凝胶电泳的方法观察到了所有三个重排反应。

2.2 内含子剪接

除 DNA 水平的剪接反应以外,在九十年代还发现了 RNA 和蛋白质水平的剪接因子,即内含子(*Intron*)和蛋白内含子(*Intein*)。原核生物的内含子有 I 组、II 组和古细菌型三种类型的剪接方式。真细菌具有 I 组、II 组内含子的 RNA 有核酶活性,自身催化剪接反应;而古细菌型内含子的剪接需要有内切核酸酶的催化。I 组内含子在剪接后成为一个环状结构,而 II 组内含子成为一个套马索结构,十分类似于真核生物核基因内含子的剪接反应方式。蛋白内含子是某些成熟蛋白质前体序列中的一段间插肽段,可通过自身催化切除;与此同时,两侧蛋白质外显子发生连接。目前已在原核生物和真核生物细胞器中发现有 100 多种蛋白内含子(<http://www.neb.com/neb/inteins.html>)。内含子中往往有编码核酸内切酶的阅读框,而蛋白内含子则可能具有核酸内切酶结构域。这些核酸内切酶对于

这些基因的寄生序列在演化过程中向其他基因的转移起关键作用。

蓝藻中已发现的 I 组内含子均位于 $tRNA_{UAA}^{leu}$ 、 $tRNA^{Met}$ 的反密码子环中。其中 $tRNA_{UAA}^{leu}$ 内含子仅发现于蓝藻和叶绿体,被认为是蓝藻内共生演化的证据之一^[9]。蓝藻 $tRNA_{UAA}^{leu}$ 内含子属 IC3 亚组,即在 P_8 区域有 CCC - - - GGG 碱基配对和 AA 突出^[10]。 P_8 区域的这一 IC3 特征结构与 L2 区 GAAA 环相作用,有助于维持 RNA 的高级结构^[11]。I 组内含子中有内部引导序列(IGS)与 5' 和 3' 外显子相配对,分别形成 P1、P10 区域,对于两端剪接位点相并列进行反应起重要作用。但在蓝藻 $tRNA_{UAA}^{leu}$ 内含子 P1 区域较短(仅 3 个碱基对),而 P10 区几乎没有。实验证明, $tRNA$ 反密码子臂的配对补偿了这一结构上的缺陷,从而使得内含子的剪接得以有效进行^[12]。

蓝藻 $tRNA_{UAA}^{leu}$ 内含子大小在 216—308 碱基范围内,均无内切核酸酶阅读框(ORF)。而在已发现的 7 种 $tRNA^{Met}$ 内含子中,集胞藻 6803 的内含子具有 ORF,为 $tRNA$ 内含子在演化过程中的横向转移提供了可能。将蓝藻不同种的 $tRNA^{Met}$ 内含子的核心序列的演化关系与依据 16S rRNA 序列拟定的系统进化树相比较,发现二者歧异颇多,表明在演化过程中可能存在横向转移^[13]。 $tRNA_{UAA}^{leu}$ 内含子可能不是单一起源的。根据其保守区域 P、V、R、S,或根据 346 个核苷同源性的比较,可将蓝藻 $tRNA_{UAA}^{leu}$ 内含子分为聚类 I 和聚类 II^[10]。有些种具有两个 $tRNA_{UAA}^{leu}$ 基因,分别插有聚类 I 或聚类 II 内含子。聚类 II 中三个微囊藻株的序列与 α -紫细菌的 $tRNA_{CCU}^{arg}$ 内含子同源性高于与许多其他 $tRNA_{UAA}^{leu}$ 内含子的同源性,可能具有独立的起源^[14];而许多聚类 I 内含子的演化关系与 16S rRNA 进化树也不相一致^[10]。以上证据表明蓝藻内含子的演化是一个较为复杂的过程。

在 *Calothrix* 中还发现了 II 组内含子,其中携有一个 1752 碱基的阅读框产物,类似于逆转录酶的聚合酶结构域,并有锌指结构。与 I 组内含子不同,II 组内含子插在一个可能编码单链 DNA 结合蛋白的 ORF 内,这是目前惟一报道的插于蛋白编码基因的蓝藻内含子^[15]。

2.3 蛋白内含子

蓝藻蛋白内含子的发现直接得益于基因组学成果。随着集胞藻 6803 全基因组序列的公布,已先后在 4 个基因识别出了蛋白内含子,它们是解旋酶(Helicase)基因 *dnaB*、螺旋酶(Gyrase)基因 *gyrB*、DNA 聚合酶 III γ/τ 亚基基因 *dnaX* 和 α 亚基基因 *dnaE*。蛋白内含子往往有 A、B、C、D、E、H、F、G 依次排列的保守序列区。其中 A、B 位于其 N-端,F、G 位于 C-端,对于自我剪接反应是必须的。C、D、E、H 区对于蛋白内含子在演化过程中向其他基因转移起作用,C、E 区是 Dod-型核酸内切酶区。在以上 4 个基因中仅有 *dnaB* 的蛋白内含子具有所有 8 个保守区域,而其余三个缺少 C、D、E、H 区,其中 DnaX 蛋白内含子具有 Dod 结构域,但与 DnaX 本身处于不同的阅读框。GyrB 中的蛋白内含子具有一个类似于 EC1H - HC3H 家族的核酸内切酶区域,在所有蛋白内含子中是十分独特的^[16]。

DnaB 蛋白内含子在集胞藻 6803 和海洋红色嗜热盐菌插在 *dnaB* 相同位置,并且二者的序列同源性达 54%,而 DnaB 序列本身同源性仅有 37%,显示有水平转移关系^[17]。对 DnaB 蛋白内含子中心区域的缺失研究表明,N-端 106 个氨基酸和 C-端 48 个氨基酸,即相当于 A、B 和 F、G 区域的部分,对于 DnaB 前体蛋白的剪接成熟是必须的。实际上紫球藻

叶绿体 DnaB 蛋白内含子就只有 A、B 和 F、G 区域。A 区 N-端的一个半胱氨酸和 G 区 C-端的天冬酰胺-丝氨酸对于该蛋白内含子的自我剪接起关键作用^[18]。

除 DnaB 的顺式剪接外,集胞藻 6803 *dnaE* 基因分为两段(*dnaE-n* 和 *dnaE-c*),在基因组上相距达 745kb,两段基因独立表达,翻译成两个蛋白质前体片断,再剪接形成 DnaE 成熟蛋白,这种方式称反式剪接。在 DNA 聚合酶中 *dnaE* 编码的 α 亚基有 β 亚基和 τ 亚基结合区域,这一区域在集胞藻 DnaE 中被蛋白内含子插入。将 *dnaE-n* 和 *dnaE-c* 在大肠杆菌表达,也可完成自我剪接反应,形成完整的 DnaE,而蛋白内含子部分被切下成为两个肽段^[19]。

3 σ (Sigma)因子和调控蛋白

3.1 σ 因子

蓝藻主要在转录水平进行基因表达的调控。负责转录的 RNA 聚合酶由核心酶和 σ 因子两部分组成全酶。在蓝藻中组成核心酶的亚单位除了 α 、 β 、 β' 蛋白外,还有一个 γ 蛋白。蓝藻的 β' 和 γ 亚基分别相当于大肠杆菌 β' 亚基的 C-端和 N-端,此点与叶绿体 RNA 聚合酶(相应亚基称为 β' 和 β'')相同。各类 σ 因子与核心酶组成的全酶可识别不同类型的启动子。在真细菌中 σ 因子在进化关系上分为 σ^{70} 和 σ^{54} 家族。研究较多的 σ^{70} 家族又可分为几组。1 组为主要 σ 因子,参与指数生长期大多数基因的转录,对细胞生存是必须的,又称为管家 σ 因子。2 组包括所有类似主要 σ 因子但又非必需的 σ 因子。3 组在结构上有别于 1、2 组,参与特殊操纵子的转录。

Anabaena sp. PCC7120 的 *SigA* 是在八十年代第一个发现的蓝藻 σ 因子,九十年代后又发现 *SigB* 和 *SigC*。*SigA* 属于 1 组 σ 因子;*SigB* 和 *SigC* 属于 2 组。其中,*SigB* 和 *SigC* 基因在去除氮源后瞬时诱导表达,但二基因突变并不导致对异形胞分化或固氮能力的任何影响^[20]。*Synechococcus* sp. PCC7002 的 Sigma 因子有 *SigA*、*SigB*、*SigC*、*SigD*、*SigE*。*SigB* 和 *SigC* 基因在氮饥饿或碳饥饿条件下转录水平显著提高,但 *SigC* 基因转录又迅速下降;*SigE* 可能对于生长稳恒期特异的基因表达起调控作用^[21]。*Synechococcus* sp. PCC7942 至少有一个 1 组 σ 因子 *rpoD1* 和三个 2 组 σ 因子 *rpoD2*、*rpoD3*、*rpoD4*。其中,*rpoD2* 基因影响 *Synechococcus* sp. PCC7942 某些基因表达的昼夜节律^[22]。在 *Synechocystis* sp. PCC6803 全基因组序列中搜寻到 8 个 σ 因子基因,其中 *SigH*、*SigG* 和 *SigF* 属 3 组 σ 因子。通过对这三个基因的敲除研究,发现 *SigF* 突变株在高盐压力下应答蛋白的诱导严重缺陷,在含 855mM NaCl 培养基上 7 天内死亡,而野生型可以耐受。*SigF* 突变株对于短时光强处理也十分敏感。另外,*SigF* 还与细胞表面纤毛的形成有关。突变株细胞表面没有纤毛,并失去野生型的向光运动能力^[23]。*SigH* 基因受外界压力的激活而表达,特别是在细胞受到热击时转录水平显著提高。点形念珠藻可与苔类植物 *Anthoceros punctatus* 共生。该藻 2 组 σ 因子 *SigH* 可能参与调控这一过程。*A. punctatus* 释放的藻殖段诱导因子(HIF)可诱导 *SigH* 基因转录;*SigH* 基因突变株对 *A. punctatus* 感染效率约为野生型的六倍^[24]。

将 *Synechococcus* 7942 的 *RpoD1*、*RpoD3* 和 *RpoD4* 分别重组表达并纯化,与该种蓝藻 RNA 聚合酶核心酶重组,以离体系统测试对大肠杆菌或蓝藻多种启动子的转录活性。结果表明,不同组合的全酶识别启动子的特性是相似的,可能都从含有 Pribnow-10 区保守序

列的真细菌启动子起始转录,但对于不同启动子又显然存在偏好^[25]。如果不同的 σ 因子都能识别 Pribnow 序列,那么 2 组 σ 因子如何在功能上有别于主要 σ 因子? 其中一种可能性即是通过与其他调控蛋白相作用。

3.2 NtcA 和 NtcB

除不同的 σ 因子以外,NtcA 和 NtcB 等调控因子对于 RNA 聚合酶与某些基因启动子的识别、结合也是必需的。NtcA 类似 CAP(Catabolite activator protein)家族成员,其 C 端有一个螺旋-转角-螺旋结构,通常与启动子-35 区 GTAN₆TAC 序列结合;受调控基因在启动子-10 区有 TAN₃T 特征序列,与 *E. coli* σ^{70} 家族启动子一致。NtcB 类似于 LysR 家族激活蛋白,在其 N 末端附近也有一螺旋-转角-螺旋结构。大多数 LysR 家族成员结合位点在受控基因转录起始位点上游约 65 碱基处,具有 TN₁₁A 保守序列特征。

NtcA 可控制多种氮调控基因的表达,如 *Synechococcus* 7942 的谷氨酰氨合成酶基因 *glnA*、亚硝酸盐还原酶基因 *nir* 等,对于利用硝酸盐生长是必需的。鱼腥藻 *ntcA* 基因突变株也不能在含硝酸盐的培养基上生长,同时还失去了异形胞分化和固氮生长能力。在固氮酶操纵子 *nifHDK*,谷氨酰氨合成酶基因 *glnA*,异形胞分化基因 *hetC* 启动子区域都证实有 NtcA 结合位点^[26]。另一个调控蛋白 NtcB 对于鱼腥藻 *nir* 启动子的激活也是必需的^[27],而在 *Synechococcus* 7942 中仅参与亚硝酸盐对 *nir* 操纵子的正调控。

4 信号传导途径

蓝藻感受外界环境的变化,并相应地调整有关基因的表达、重新整合生理状态,这一过程需要有信号传导途径的介导,其主体是感应环境变化的组氨酸激酶和相应接受磷酸基团的调控蛋白,称二组分信号传导系统(Two-component signal transduction system)。这一调控机制和不同 σ 因子以及 NtcA、NtcB 等调控蛋白之间可能存在交叉或耦联的关系。在蓝藻中也发现有真核生物类型的 Ser/Thr 激酶以及 Ser/Thr 和 Tyr 磷酸酶^[28],但除了集胞藻 6803 的 *spkA* 证实为爬行运动所必需以外,有关基因突变性状都不显著,其生理作用和调节方式有待深入研究。而二组分信号传导系统在互补色适应、生物钟节律调制、细胞分化、向光运动等均起关键作用。其中,对于 *Calothrix* sp. PCC7601(*Fremyella diplosiphon*)藻胆体组分调节过程中信号传导的研究最为系统。

Calothrix 7601 在红光中关闭藻红蛋白基因 *cpeBA*、开启藻蓝蛋白基因丛 *cpcB2A2*;而在绿光中则相反,故藻体外观上呈现与光质互补的颜色。曾有人设想存在一个植物光敏色素类似成分参加的感光 and 基因调控机制,这一假设在九十年代得到了证实。*Calothrix* 7601 互补色适应的信号传导途径至少有 RcaE、RcaF 和 RcaC 三个组分。其中,RcaE N-端序列与光敏色素类光感受器中结合发色团的区域同源,C-端有作为感受器的组氨酸激酶中 H, N, G1, F 和 G2 保守序列^[29]。RcaF 是一个响应调节蛋白,但无 DNA 结合区域。二者编码基因 *rcaE* 和 *rcaF* 仅相距 12bp,同向排列。RcaC 是一个特殊的响应调节蛋白,其 N-端和 C-端各有一个磷酸基团受体区域,但定点突变证明仅 N-端区域对互补色适应必需;中部又有一个与某些组氨酸激酶相似的 H2 序列,和一个 DNA 结合序列^[30]。根据大量的遗传学证据,Kehoe 和 Grossman 提出互补色适应调控模型。在感受红光时,RcaE 发挥激酶功能,首先在激酶结构域的一个保守的组氨酸发生自我磷酸化,再把磷酸基团传递给 RcaF

的一个保守的天冬氨酸残基。该磷酸基团相继传递到 RcaC 中部激酶区域的一个组氨酸和 N-端的天冬氨酸受体。被磷酸化的 RcaC 发挥调节功能,致使 *cpeBA* 关闭、*cpcB2A2* 开启。而在接受绿光时,RcaE 发挥磷酸酶的功能,RcaC 不被 RcaF 磷酸化,导致 *cpeBA* 开启、*cpcB2A2* 关闭。在这一模型中共有两个组氨酸和两个天冬氨酸残基参与了磷酸基团的传递,可见信号传导过程有时如同一个级联式的“继电器”系统,包含多个步骤。

与 RcaE 具有类似结构的还有 CikA,PlpA,Cph1 等,其中,CikA 在调制蓝藻生物节律中发挥作用;PlpA 缺失时则不能适应在蓝光中生长。Cph1 是第一个被识别的蓝藻光敏色素蛋白,通过其组氨酸激酶区域介导响应调控蛋白 RcpI 的红光-远红光可逆磷酸化,但其生理功能不明^[31]。虽然如此,感光区域(或其他理化因子感受区域)并非一定与激酶区域处于同一蛋白分子。另一个蓝藻光敏色素蛋白 Cph2 具有两个发色团结合区域,其中 N-端位点可与藻蓝胆素或光敏色素自我催化结合,但并无组氨酸激酶区域。

集胞藻 PCC6803 的向光性是一个比互补色适应更复杂的过程。这种爬行运动依赖于 Type IV 纤毛。从细胞感受光源到纤毛趋动的向光运动依靠一个信号传导途径的介导。业已证明,*sl10038*,*sl10039*,*sl10041*,*sl10042* 和 *sl10043* 在被插入失活时均导致负向光性,即藻细胞背离光源运动。其中,*sl10041* 具有一个与 RcaE 结合发色团保守序列同源的区域,而 *sl10038*,*sl10039* 与大肠杆菌有关趋化运动基因 *cheY* 相似,*sl10043* 与 *cheA* 相似^[32]。CheA 是个组氨酸激酶,在大肠杆菌中,将来自氨基酸受体 Tsr 或 Tar 的信号传递给响应调控蛋白 CheY。推测 *sl10041* 产物是向光运动的光感受器,通过 CheA 和 CheY 类似蛋白激活纤毛运动机制,使细胞向光爬行。这一系统缺失导致负向光性,表明还存在其他控制细胞运动的光感受体和信号传导途径。

在更为复杂的生命活动中,二组分信号系统往往作为庞大的基因调控网络中的部分环节,有时起到核心作用,但有时可能仅在网络各部分或者阶段间起协调作用。这一现象分别在蓝藻生物节律和异形胞分化的分子机制中得到体现。

5 蓝藻生物钟

适应地球昼夜周期,蓝藻的基因表达具有生物节律现象,称为蓝藻生物钟。研究证明,当生物钟节律与外界光照节律更为接近时,藻细胞具有较强的竞争优势,具有适应意义。事实上蓝藻与真核生物的生物钟无任何联系,各自具有独立的起源,这也从另一个方面体现了生物钟对蓝藻生命活动的重要性。

利用整合于聚球藻 7942 基因组的 P_{psbAI} -*luxAB*,Kondo T. 等人发现荧光信号的变化满足生物钟的三个判断标准:1) 经 12h 光照/12h 黑暗(L/D)处理后,在持续光照条件下能在较长时间内保持周期约 1d 的节律性震荡;2) L/D 信号的时相或短时黑暗能够调制生物钟时相;3) 在生理耐受限度内的温度变化对生物钟周期没有显著影响,即存在温度补偿反应。通过对固体培养基上聚球藻 7942:: P_{psbAI} -*luxAB* 藻落的同步定位观察和记录,证实其荧光素酶基因的表达同样呈现出节律性,为突变株的筛选和生物钟基因的克隆奠定了基础^[33]。利用 EMS 诱变获得的 19 个点突变,无论表型为周期长度变化、图式变化、或无节律,均定位于 *kaiA-kaiB-kaiC* 基因内。其中 A 基因单独转录,B 和 C 共同转录;与 *psbAI* 基因相似,*kai* 基因的表达也呈现出节律性^[34]。

对 *kaiA-kaiB-kaiC* 逐个敲除,发现任何一个基因缺失都将使生物钟节律消失。*kaiA* 与 P_{inc} 融合,在 IPTG 诱导超量表达时会使 *kaiBC* 表达水平过高而不能显示节律。相反,*kaiC* 过量表达会抑制 *kaiA* 和 *kaiBC* 基因表达。更为重要的是,对 *kaiC* 短暂的诱导会改变生物钟的时相。这种改变的程度与诱导所处的时相有关:先于 *kaiC* 表达峰值时可使后续周期时相前移,后于 *kaiC* 表达峰值则时相后滞;人为诱导与 *kaiC* 预定的峰值之间的时差决定了相位迁移的大小。因此,*kaiA* 对 *kaiBC* 起正调控作用,而 *kaiC* 起负调控作用,它们共同构成了生物钟的核心,并形成周期性振荡。其中,KaiC 蛋白由两个重复的结构域组成,即 1-260 位氨基酸的 CI 区和 261-519 位的 CII 区。CI 和 CII 十分相似,各自可与 KaiA、KaiB 或 KaiC 自身相作用^[35]。KaiB 分子之间也可自相作用。

SasA 是一个感应组氨酸激酶,它与 KaiC 的复合体在细胞中持续存在,而 KaiB 与 KaiC 的复合体则呈节律性变化。*SasA* 突变株表现为:在持续光照下生长不受影响,但在 L/D 培养条件下,生长缓慢;*KaiBC* 表达量减少到 5%—10%,*kaiA* 表达量约减少一半,二者节律性均消失。*SasA* 过量表达时抑制 *kaiBC* 的表达,同时也使之失去节律性。当 *SasA* 被短暂诱导之后,它对生物钟时相的改变方向(延迟或提前)及改变程度与短时诱导 KaiC 对时相的改变情况正好相反。*SasA* 极可能与一个响应调控因子(RR)形成二组份信号传导系统,对 *kaiABC* 组成的生物钟核心起着放大器的作用^[36]。

CikA 的 N 端氨基酸序列显示其是一个细菌光敏色素,可能起到光接收器的作用,但缺少与 bilin 共价结合的 Cys 或 His 残基。在 CikA 中部有组氨酸激酶特征性的 H, N, D/F 和 G 保守序列;在 C 端还有响应调控因子的接收区结构域。*CikA* 突变株生物钟周期较野生种缩短约 2h,但更重要的是,在不同时间以一个 5h 的暗处理基本不能对生物钟时相发生调制;而在野生种则可提前或后移 8—10h。因此,CikA 在生物钟的信号输入途径中发挥重要作用^[37]。此外,聚球藻的 *pex* 基因具有延长生物钟周期的作用,可能对生物钟的作用方式有修饰作用。

Liu 等发现聚球藻基因组中表达水平呈现节律性的基因可大致分为 5 类^[38]。其中 1, 2, 3 类波形是基本对称的 Sine 曲线,它们之间的区别在于时相不同,如 1 类和 2 类波峰、波谷所处时相恰好相反。4 类波形不对称,呈锯齿状;其余不规则波形归入第 5 类。同一细胞内存在不同的周期和时相,反映了从生物钟核心发出的节律信号在不同输出途径中分别进行了调制。因此,蓝藻生物钟的组织包括信号输入、核心部分和不同的信号输出途径。

6 异形胞的分化

鱼腥藻的丝体由一系列细胞组成,在缺氮诱导的条件下,其中一些间隔分布的细胞分化成异形胞进行固氮作用。在从营养细胞到异形胞的分化过程中细胞停止分裂;细胞壁外形成双层包被结构,外面是多糖层,里面是糖脂层,共同形成异形胞内的固氮酶防氧屏障;细胞内类囊体膜结构和相关酶系也经历了深刻改组,总体上表现为失去光合放氧活性而呼吸速率大大提高。这一生理过程的背后可能隐藏着蓝藻生物学中最为庞大的基因调控网络^[5]。综合基因调控和细胞学证据,作者认为异形胞发育过程应分三个主要阶段,即启动分化、细胞分裂终止和形态建成。

在异形胞发育的起始阶段,*hetR* 基因起关键作用。该基因发生突变时失去形成异形

胞的能力,而克隆于高拷贝质粒或由铜离子诱导的启动子使之过量表达则导致形成成串异形胞,而且不被硝酸盐或氨盐抑制^[39]。*hetR* 基因具有自我调控特征,即 *hetR* 是其自身缺氮诱导表达的条件;在缺氮诱导后 0—6h 表达持续增加,并且表达 *hetR* 的细胞呈间隔分布。以生物化学和点突变手段证明,*HetR* 蛋白很可能是一个丝氨酸型蛋白酶^[40],在发育起始阶段诱导表达,降解某种抑制蛋白或激活某种调控蛋白从而启动异形胞的发育。

可能与 *hetR* 处于同一调控途径的有 *patA* 基因。*patA* 突变株几乎只能在丝体的末端形成异形胞,从而完全改变了鱼腥藻异形胞的分布图式。该基因序列显示其产物类似二组分信号传导系统中的调控蛋白,但无 DNA 结合位点。在 *patA* 突变株过量表达 *hetR* 基因,仍只能在丝体末端产单个异形胞,推测 *patA* 对于 *hetR* 充分发挥其生理功能是必须的^[39,41]。

hetC 基因编码一个消耗 ATP 转运蛋白质或多肽的运载体,在缺氮诱导后约 3.5h,几乎与 *hetR* 基因同步,已经能够检测到 *hetC* 基因的显著表达,并且也被定位到原异形胞^[42]。*hetC* 基因对于异形胞发育早期细胞分裂的终止起关键作用。*hetC* 突变株在缺氮诱导时可启动 *hetR* 基因在间隔分布细胞的表达,但这些细胞不能终止分裂,从而不能启动异形胞的形态建成。已经证明,在成熟异形胞中细胞分裂基因 *ftsZ* 不能表达。*hetC* 如何转运某种蛋白或多肽控制细胞分裂基因的表达将成为涉及细胞分裂和分化两个重要领域的饶有兴趣的问题。

与 *hetR*、*hetC* 等调控分化的基因不同,*patS* 对异形胞分化起抑制作用。这种抑制作用对于鱼腥藻异形胞分布图式的形成具有重要作用。计算机模拟研究表明异形胞可能产生某种沿丝体扩散的物质抑制其他细胞的分化,从而形成鱼腥藻的异形胞间隔分布图式。*patS* 基因产物或衍生物可能就是这样一个抑制物^[43]。该基因仅编码 17 个氨基酸,过量表达时完全抑制异形胞的形成;而缺失则得到许多成串异形胞,并且间隔异形胞的营养细胞数变少。*patS* 基因的 17 个氨基酸中仅 C-端的 5 个氨基酸 RGSGR 对于抑制异形胞的发育是必须的。*patS* 基因在原异形胞和异形胞特异表达,可能释放出 C-端 5 肽,沿丝体扩散,抑制相邻营养细胞启动分化,从而参与决定异形胞发育的图式。

在形态发生上包被的多糖外层比里面的糖脂层形成较早。大约在缺氮诱导后 6—7h,*hepA* 基因开始在间隔分布的细胞特异表达^[44]。该基因突变时异形胞不能形成多糖层;糖脂层能够形成,但由于失去外层的保护易于松脱。*hepB*、*hepC* 基因突变也导致多糖层不能形成,这两个基因产物可能催化多糖的合成。*hepA* 的表达受一个组氨酸激酶基因 *hepK* 的调控^[45]。

参与糖脂层合成或沉积的基因,如 *hetM* (*hglB*) *hglC*、*hglD*、*hetN* 和 *hglK* 等,可能并非都受缺氮诱导,但至少 *hetM* 是在缺氮之后的 9h 开始表达^[44]。糖脂在细胞内合成之后必须转运到细胞壁外,沉积形成糖脂层。*devBAC* 基因产物可能形成一个利用 ATP 的细胞膜上的转运体。*dev* 基因突变时,细胞内仍合成糖脂,但不能形成包被中的糖脂层^[46]。糖脂层的沉积还受到细胞表面结构的影响^[47]。

在以上决定异形胞分化的不同阶段基因之间存在着调控关系:在 *hetR* 突变株中 *hepA*、*devA*、*hetM* 等基因不能表达^[44]; *hetC* 基因对 *hetR* 的表达几乎没有什么影响,但对 *hepA* 基因的表达是必须的^[42]; *hetR* 和 *hetC* 基因都受 *nucA* 基因的调控。*NucA* 调节异形胞发育

不同阶段的基因表达,包括异形胞成熟后固氮酶基因丛中 DNA 重排的发生。*hetR* 和 *hetC* 等基因的表达也被 *patS* 基因编码的 5 肽 RCSCR 抑制;但去除启动子的 *patS* 基因置于 *he-pA* 启动子控制之下同样能形成与野生型相同的异形胞发育图式^[43]。

7 展望

如上所述,蓝藻分子遗传学在上个世纪九十年代取得了诸多重要进展,其中基因组数据的积累奠定了未来十年本学科发展的基础。一些复杂调控网络的解析将依赖于全基因组规模的对于基因功能的系统研究。目前蓝藻基因组数据中约有半数基因功能不明甚至毫无线索,需要高通量的方法弄清它们在生命活动中的作用。在其他微生物模式种业已建立的功能基因组学方法主要包括以下层次:(1) 基因的系统“敲除”;(2) 转录组学;(3) 蛋白组学;(4) 代谢组学和(5) 生物信息学。其中,转录组学、蛋白组学方法已在集胞藻 6803 获得应用;生物信息学方法在所有基因组数据的分析中都是必需的,但有关蛋白相互作用和基因组的比较等方面仍有很大拓展空间。基因的系统插入失活将提供有关基因功能的最直接的证据,是包括本实验室在内的若干研究小组目前正在努力的目标。与以上各层次相结合,代谢组学方法将进一步描绘出更为精细的蓝藻生命活动图景。

蓝藻可能在一些具有广泛意义的生物学问题上显示特殊的研究价值,如生物膜的发生、广宿主内共生等重要生命活动或生命现象的遗传基础。类囊体膜是蓝藻进行光合作用的场所,也有呼吸电子传递链,但对于具有异养生活能力的蓝藻,如集胞藻 6803,失去类囊体膜的突变株仍然可利用糖源正常生长^[48]。对集胞藻 6803 进行系统的基因“敲除”或对随机插入突变株进行定向筛选均可能发现一系列与类囊体发生相关的基因。光合膜系的发生可能成为生物膜发生研究的一个重要模式,而其他生物膜系几乎不可能进行类似的遗传学研究。蓝藻中还存在一个特殊的种——*Gloeobacter violaceus*,代表着演化的一个侧支。该种具有蓝藻的光合系统、藻胆体和光合色素,但均位于细胞质膜,而细胞内没有类囊体膜结构。日本 Kazusa DNA 研究所已宣布对该种进行基因组序列分析,预测将对类囊体膜发生的研究起到推动作用。念珠藻的广宿主共生现象是微生物与植物相互作用的一个十分有趣的模式,其宿主涉及到苔类、蕨类、裸子植物和被子植物,而根瘤菌的共生范围局限于豆科植物,弗兰克氏菌局限于 8 科木本植物。对点形念珠藻的功能基因组学研究将可能在这一问题上取得突破。

蓝藻生物学的其他方面也将在分子遗传学层次得到深入研究。如蓝藻细胞壁的形成,蓝藻的运动,细胞分裂等。在应用方面,水华蓝藻已引起较多关注,如微囊藻毒素合成基因几乎全部获得解析^[49],微囊藻全基因组测序正在进行。总之,蓝藻分子遗传学在下一个十年将全面进入基因组时代,其研究深度和规模将是空前的。但是,对于调控途径的解析仍然不可能脱离对具体基因或产物相互作用关系的细致分析。在这一领域,我国已出现若干有一定水准的研究队伍,预计在未来十年内将取得更大的成绩。

参考文献:

- [1] 王业勤,徐旭东,黎尚豪. 蓝藻分子遗传学十年研究进展[J]. 水生生物学报,1991,15:356—367
- [2] Bryant DA. Molecular biology of cyanobacteria [C]. Dordrecht/Boston/London: Kluwer Academic Publishers. 1994

- [3] Kaneko T, Sato S, Kotani H, et al. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions [J]. *DNA Res*, 1996, **3**:109—136
- [4] Brahmsha B. A genetic manipulation system for oceanic cyanobacteria of the genus *Synechococcus* [J]. *Appl Environ Microbiol*, 1996, **62**:1747—1751
- [5] Wolk C P. Heterocyst formation in *Anabaena* [C]. Brun Y V and Shimkets L J. Prokaryotic Development. Washington, DC: American Society for Microbiology, 2000. 83—104
- [6] Meeks J C. Symbiosis between nitrogen-fixing cyanobacteria and plants [J]. *BioScience*, 1998, **48**:266—276
- [7] Cohen M F, Wallis J G, Campbell E L, et al. Transposon mutagenesis of *Nostoc* sp. strain ATCC 29133, a cyanobacterium with multiple cellular differentiation alternatives. *Microbiology*, 1994, **140**(Pt12):7324—7329
- [8] Carrasco C, Buettner JA, Golden JW. Programed DNA rearrangement of a cyanobacterial *hupL* gene in heterocysts [J]. *Proc Natl Acad Sci USA*, 1995, **92**:791—5
- [9] Xu MQ, Kathe SD, Goodrich-Blair H, et al. Bacterial origin of a chloroplast intron: conserved self-splicing group I introns in cyanobacteria [J]. *Science*, 1990, **250**:1566—70
- [10] Rudi K, Jakobsen KS. Complex evolutionary patterns of tRNA_{CAU}^{Leu} group I introns in cyanobacterial radiation [J]. *J Bacteriol*, 1999, **181**:3445—51
- [11] Ikawa Y, Naito D, Aono N, et al. A conserved motif in group IC3 introns is a new class of GNRA receptor [J]. *Nucleic Acids Res*, 1999, **27**:1859—65
- [12] Zaug AJ, MaEvoy MM, Cech TR. Self-splicing of the group I intron from *Anabaena* pre-tRNA: requirement for base-pair of the exons in the anticodon stem [J]. *Biochemistry*, 1993, **32**:7946—53
- [13] Paquin B, Kathe SD, Nierzwicki-Bauer SA and Shub DA. Origin and evolution of group I introns in cyanobacterial tRNA genes [J]. *J Bacteriol*, 1997, **179**:6798—6806
- [14] Paquin B, Heinfling A, Shub DA. Sporadic distribution of tRNA_{CAU}^{Leu} introns among α -purple bacteria: evidence for horizontal transmission and transposition of a group I intron [J]. *J Bacteriol*, 1999, **181**:1049—53
- [15] Ferat J.-L., Michel F. Group II self-splicing introns in bacteria [J]. *Nature*, 1993, **364**:358—61
- [16] Gorbalenya AE. Non-canonical inteins [J]. *Nucleic Acids Res*, 1998, **26**:1741—8
- [17] Liu XQ, Hu Z. A DnaB intein in *Rhodothermus marinus*: indication of recent intein homing across remotely related organisms [J]. *Proc Natl Acad Sci USA*, 1997, **94**:7851—6
- [18] Mathys S, Evans Jr TC, Chute IC, et al. Characterization of a self-splicing miniintein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation [J]. *Gene*, 1999, **231**:1—13
- [19] Wu H, Hu Z, Liu XQ. Protein trans-splicing by a split intein encoded in a split DnaE gene of *Synechocystis* sp. PCC6803 [J]. *Proc Natl Acad Sci USA*, 1998, **95**:9226—31
- [20] Brahmsha B, Haselkorn R. Identification of multiple RNA polymerase sigma factor homologs in the cyanobacterium *Anabaena* sp. Strain 7120: Cloning, expression, and inactivation of the *sigB* and *sigC* genes [J]. *J Bacteriol*, 1991, **174**:7273—7282
- [21] Gruber T M, Bryant D A. Characterization of the alternative σ -factors SigD and SigE in *Synechococcus* sp. Strain PCC 7002. SigE is implicated in transcription of post-exponential-phase-specific genes [J]. *Arch Microbiol*, 1998, **169**:211—219
- [22] Tsinoremas N F, Ishiura M, Kondo T, et al. A sigma factor that modifies the circadian expression of a subset of genes in cyanobacteria [J]. *EMBO J*, 1996, **15**:2488—2495
- [23] Bhaya D, Watanabe N, Ogawa T, et al. The role of an alternative sigma factor in motility and pilus formation in the cyanobacterium *Synechocystis* sp. strain PCC6803 [J]. *Proc Natl Acad Sci USA*, 1999, **96**:3188—3193
- [24] Campbell E L, Brahmsha B, Meeks J C. Mutation of an alternative sigma factor in the cyanobacterium *Nostoc punctiforme* results in increased infection of its symbiotic plant partner, *Anthoceros punctatus* [J]. *J Bacteriol*, 1998, **180**:4938—4941
- [25] Goto-Seki A, Shirokane M, Masuda S, et al. Specificity crosstalk among group 1 and group 2 sigma factors in the cyanobacterium *Synechococcus* sp. PCC7942: in vitro specificity and a phylogenetic analysis [J]. *Mol Microbiol*, 1999, **34**:473—484
- [26] Muro-Pastor A M, Valladares A, Flores E, et al. The *hetC* gene is a direct target of the NtcA transcriptional regulator in cya-

- nobacterial heterocyst development [J]. *J Bacteriol*, 1999, **181**:6664—6669
- [27] Frias J E, Flores E, Herrero A. Activation of the *Anabaena nir* operon promoter requires both NtcA (CAP family) and NtcB (LysR family) transcription factors [J]. *Mol Microbiol*, 2000, **38**:613—625
- [28] Zhang C C. Bacterial signaling involving eukaryotic-type protein kinases [J]. *Mol Microbiol*, 1996, **20**:9—15
- [29] Kehoe D M, Grossman A R. Similarity of a chromatic adaptation sensor to phytochrome and ethylene receptors [J]. *Science*, 1996, **273**:1409—1412
- [30] Kehoe D M, Grossman A R. New classes of mutants in complementary chromatic adaptation provide evidence for a novel four-step phosphorelay system [J]. *J Bacteriol*, 1997, **179**:3914—3921
- [31] Yeh K C, Wu S H, Murphy J T, et al. A cyanobacterial phytochrome two-component light sensory system [J]. *Science*, 1997, **277**:1505—1508
- [32] Yoshihara S, Suzuki F, Fujita H, et al. Novel putative photoreceptor and regulatory genes required for the positive phototactic movement of the unicellular motile cyanobacterium *Synechocystis* sp. PCC6803 [J]. *Plant Cell Physiol*, 2000, **41**:1299—1304
- [33] Kondo T and Ishiura M. Circadian rhythms of cyanobacteria: monitoring the biological clocks of individual colonies by bioluminescence [J]. *J Bacteriol*, 1994, **176**:1881—5
- [34] Ishiura M, Kutsuna S, Aoki S, et al. Expression of a gene cluster *kaiABC* as a circadian feedback process in cyanobacteria [J]. *Science*, 1998, **281**:1519—23
- [35] Iwasaki H, Taniguchi Y, Ishiura M, et al. Physical interactions among circadian clock proteins, KaiA, KaiB, KaiC, in cyanobacteria [J]. *EMBO J*, 1999, **18**:1137—45
- [36] Iwasaki H, Williams SB, Kitayama Y, et al. A KaiC-interacting sensory histidine kinase, SasA, necessary to sustain robust circadian oscillation in cyanobacteria [J]. *Cell*, 2000, **101**:223—233
- [37] Schmitz O, Katayama M, Williams SB et al. CikA, a bacteriophytochrome that resets the cyanobacterial circadian clock [J]. *Science*, 2000, **289**:765—768
- [38] Liu Y, Tsinoremas NF, Johnson CH, et al. Circadian orchestration of gene expression in cyanobacteria [J]. *Genes Dev*, 1995, **9**:1469—78
- [39] Buikema WJ, Haselkorn R. Expression of the *Anabaena hetR* gene from a copper-regulated promoter leads to heterocyst differentiation under repressing conditions [J]. *Proc Natl Acad Sci USA*, 2001, **98**:2729—34
- [40] Zhou R, Wei X, Jiang N, Li H, Dong Y, Hsi KL and Zhao J. Evidence that HetR protein is an unusual serine-type protease [J]. *Proc Natl Acad Sci USA*, 1998, **95**:4959—63
- [41] Liang J, Scappino L, Haselkorn R. The *patA* gene product, which contains a region similar to CheY of *Escherichia coli*, controls heterocyst pattern formation in the cyanobacterium *Anabaena* 7120 [J]. *Proc Natl Acad Sci USA*, 1992, **89**:5655—5659
- [42] Xu X, Wolk CP. Role for *hetC* in the transition to a nondividing state during heterocyst differentiation in *Anabaena* sp. [J]. *J Bacteriol*, 2001, **183**:393—396
- [43] Yoon HS and Golden JW. Heterocyst pattern formation controlled by a diffusible peptide [J]. *Science*, 1998, **282**:935—938
- [44] Cai Y, Wolk CP. *Anabaena* sp. strain PCC7120 responds to nitrogen deprivation with a cascade-like sequence of transcriptional activations [J]. *J Bacteriol*, 1997, **179**:267—271
- [45] Zhu J, Kong R, Wolk CP. Regulation of *hepA* of *Anabaena* sp. strain by elements 5' from the gene and by *hepK* [J]. *J Bacteriol*, 1998, **180**:4233—4242
- [46] Fiedler G, Arnold M, Hannus S, et al. The DevBCA exporter is essential for envelope formation in heterocysts of the cyanobacterium *Anabaena* sp. strain PCC7120 [J]. *Mol Microbiol*, 1998, **27**:1193—1202
- [47] Xu X, Khudyakov I, Wolk CP. Lipopolysaccharide dependence of cyanophage sensitivity and aerobic nitrogen fixation in *Anabaena* sp. strain PCC7120 [J]. *J Bacteriol*, 1997, **179**:2884—91
- [48] Westphal S, Heins L, Jurgen S, et al. *Vipp1* deletion mutant of *Synechocystis*: A connection between bacterial phage shock and thylakoid biogenesis [J]? *Proc Natl Acad Sci USA*, 2001, **98**:4243—4248
- [49] Tillett D, Dittmann E, Erhard M, et al. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide-polyketide synthetase system [J]. *Chem Biol*, 2000, **7**:753—764