

蓝细菌 16S rRNA 双甲基化酶基因 *ksgA* 的进化分析、蛋白结构预测及功能验证

王婧蓝 胡 胜 王 莉 陈雯莉

华中农业大学农业微生物学国家重点实验室, 武汉 430070

摘要 为研究核糖体小亚基 rRNA 双甲基化酶基因 *ksgA* 的进化关系、蛋白结构和功能, 利用 MEGA 5.2 软件构建了 37 种蓝细菌 *ksgA* 基因和 16S rDNA 的系统进化树, 通过 SWISS-MODEL 分析和预测 KsgA 的蛋白质结构模型, 并通过克隆集胞蓝细菌 *Synechocystis* sp. strain PCC 6803、鱼腥蓝细菌 *Anabaena* sp. strain PCC 7120 和聚球蓝细菌 *Synechococcus elongates* sp. strain PCC 7942 的 *ksgA* 基因对大肠杆菌 *ksgA* 突变体进行了异源互补。结果表明, *ksgA* 基因在蓝细菌中与 16S rDNA 进化分支高度相似, 且体现出蓝细菌进化的聚类特征; 蓝细菌 KsgA 甲基化酶与该蛋白家族的其他蛋白质拥有共同的保守结构、催化位点和配体结合位点; 克隆的 3 种蓝细菌 *ksgA* 基因均能在一定程度上互补大肠杆菌 *ksgA* 基因的功能。

关键词 *ksgA*; 蓝细菌; 双甲基化修饰; 序列分析; 蛋白结构预测; SWISS-MODEL

中图分类号 Q 931 **文献标识码** A **文章编号** 1000-2421(2016)06-0001-10

蛋白质是生命活动的主要体现者, 蛋白质的功能和结构在进化中的联系一直都是生命科学相关领域的研究热点^[1]。近年来, 1 个在动物、植物、微生物中都极端保守的蛋白质家族——KsgA/Dim1 蛋白质家族, 因其功能和结构上进化的特殊性引起了研究者的关注。KsgA/Dim1 家族的蛋白质是典型的 S-腺苷甲硫氨酸(SAM)依赖的甲基化酶, 能够以 SAM 为甲基供体, 催化甲基基团的转移^[2]。这个家族蛋白质最核心的功能是参与核糖体小亚基 rRNA 的修饰, 使其 3' 末端颈环上 2 个相邻的腺嘌呤发生双甲基化^[3]。

在核糖体成熟的过程中, 核糖体的修饰是一个重要且必需的步骤。相对于 rRNA 自身的保守性, rRNA 的修饰在不同物种中表现出惊人的多样性。而 KsgA/Dim1 对 16S rRNA 的双甲基化修饰作用, 是目前已知仅有的 3 个保守的 16S rRNA 修饰之一^[4]。

在原核生物中, 作为将 16S rRNA 末端 A1518 和 A1519(*E. coli* rRNA 碱基序列) 双甲基化的修饰酶^[5], *ksgA* 的突变体在蛋白质翻译过程中可能会出现高频错误而使细胞承担不可预知的风险, 但是

目前的相关研究显示 *ksgA* 对于大多数原核生物的生存似乎不是必需的^[6]; 同时, *ksgA* 的缺失使这些微生物出现多样化的表型: 大肠杆菌^[3]、淋球菌^[7]等细菌的 *ksgA* 突变体能够表现出春雷霉素(kasugamycin)抗性; 金黄色葡萄球菌^[6]等少量种类的细菌则表现出对温度的敏感。在真核生物中, KsgA/Dim1 家族的同源蛋白的功能差异甚大。在绿色植物和部分真核藻类中, 细胞核染色体上存在并表达多个不同的该家族的基因, KsgA/Dim1 家族同源蛋白 PFC(Paleface 1)是修饰叶绿体核糖体的双甲基化酶, Dim1B 修饰 18S rRNA。PFC 的缺失突变体对温度敏感, 在低温下其叶绿体无法正常发育^[8]。在酿酒酵母(*Saccharomyces cerevisiae*)中, KsgA/Dim1 家族的同源蛋白 sc-mtTFB 已没有甲基化酶的活性, 而成为线粒体转录因子 B(sc-mtTFB), 是线粒体转录装置的组分之一^[9]。人类则拥有 2 种形式的 mtTFB, 不同于 sc-mtTFB, 人类的 h-mtTFB1 和 h-mtTFB2 都足够保守, 作为线粒体转录因子的同时仍拥有甲基化酶的活性^[10]。

蓝细菌是地球上已知的最早生物之一, 遍布于世界上几乎所有的生境, 并演化出多种多样的适应

收稿日期: 2016-01-06

基金项目: 国家自然科学基金项目(31570048); 中央高校基本科研业务费专项(2014PY003)

王婧蓝, 硕士研究生。研究方向: 蓝细菌分子生物学。E-mail: wangjinglan1989@gmail.com

通信作者: 陈雯莉, 博士, 教授。研究方向: 环境微生物学。E-mail: wlchen@mail.hzau.edu.cn

性形态、构造以及生存机制,同时对地球生态环境的改变和生命系统的演化产生了巨大的影响^[11]。作为在进化上较早独立分支的类群,蓝细菌保留了大量的古老遗传信息,成为分子进化领域中不可多得的材料和不容忽视的研究对象^[12]。基于此,本研究选择蓝细菌的 KsgA 蛋白质家族进行系统进化分析和结构预测,并初步验证其生物学功能,以期为蛋白质结构与功能共同进化理论提供新的实验证据。

1 材料与方法

1.1 材 料

本研究使用的菌株和质粒见表 1。

表 1 本研究使用的菌株及质粒
Table 1 Strains or plasmids used in this study

菌株及质粒 Strain or plasmid	特点 Description	来源 Source
菌株 Strains		
<i>Anabaena</i> PCC 7120	野生型 Wild type	法国巴斯德菌种保藏中心 Pasteur Culture Collection
<i>Synechocystis</i> PCC 6803	野生型 Wild type	法国巴斯德菌种保藏中心 Pasteur Culture Collection
<i>Synechococcus elongatus</i> PCC 7942	野生型 Wild type	法国巴斯德菌种保藏中心 Pasteur Culture Collection
<i>E. coli</i> MG1655	野生型 K-12 菌株 Wild-type K-12 strain	笔者所在实验室保藏 The author's laboratory
<i>E. coli</i> PR7	<i>thr</i> -1, <i>leuB</i> 6 (<i>Am</i>), <i>lacY</i> 1, <i>rna</i> -19, <i>pnp</i> -7, <i>rpsL</i> 132 (<i>strR</i>), <i>malT</i> 1(λ R), <i>xyl</i> -7, <i>mtlA</i> 2, <i>thiE</i> 1	笔者所在实验室保藏 The author's laboratory
<i>E. coli</i> Q13	<i>rna</i> -19, <i>gdhA</i> 2, <i>his</i> -95, <i>tyrA</i> 6, <i>relA</i> 1, <i>pnp</i> -13, <i>spoT</i> 1, <i>metB</i> 1	笔者所在实验室保藏 The author's laboratory
<i>E. coli</i> JC12	<i>fhuA</i> 2, <i>lacY</i> 1 or <i>lacZ</i> 4, <i>tsx</i> -1, <i>glnV</i> 44 (<i>AS</i>), <i>gal</i> -6, <i>rfbC</i> 1, <i>purF</i> 1, <i>xyl</i> -7, <i>mtlA</i> 2, <i>metB</i> 1	笔者所在实验室保藏 The author's laboratory
<i>E. coli</i> JC411	<i>leuB</i> 6 (<i>Am</i>), <i>fhuA</i> 2, <i>lacY</i> 1, <i>glnV</i> 44 (<i>AS</i>), <i>gal</i> -6, <i>hisG</i> 1 (<i>Fs</i>), <i>rfbC</i> 1, <i>galP</i> 63, <i>argG</i> 6, <i>rpsL</i> 104, <i>malT</i> 1 (λ R), <i>xyl</i> -7, <i>mtlA</i> 2, <i>metB</i> 1	笔者所在实验室保藏 The author's laboratory
<i>E. coli</i> BW25113	Δ (<i>araD</i> - <i>araB</i>) 567, Δ <i>lacZ</i> 4787 (: : <i>rrnB</i> -3), <i>rph</i> -1, Δ (<i>rhaD</i> - <i>rhaB</i>)568, <i>hsdR</i> 514	笔者所在实验室保藏 The author's laboratory
<i>E. coli</i> EMGksgA2	MG1655 <i>ksgA</i> 2	笔者所在实验室保藏 The author's laboratory
<i>E. coli</i> EMGksgA4	MG1655 <i>ksgA</i> 4	笔者所在实验室保藏 The author's laboratory
<i>E. coli</i> TPR201	PR7 <i>ksgA</i> 40	笔者所在实验室保藏 The author's laboratory
<i>E. coli</i> FS242	Q13 <i>ksgA</i> 23	笔者所在实验室保藏 The author's laboratory
<i>E. coli</i> FS131	JC12 <i>ksgA</i> 19	笔者所在实验室保藏 The author's laboratory
<i>E. coli</i> FS232	JC411 <i>ksgA</i> 19	笔者所在实验室保藏 The author's laboratory
<i>E. coli</i> FS233	JC411 <i>ksgA</i> 23	笔者所在实验室保藏 The author's laboratory
<i>E. coli</i> JW0050-3	BW25113 Δ <i>ksgA</i> 733::kan	笔者所在实验室保藏 The author's laboratory
质粒 Plasmids		
pBluescript II KS	Cloning vector	安捷伦科技有限公司 Agilent Technologies
<i>psk-alr</i> 3229		本研究构建 This study
<i>psk-sll</i> 0708		本研究构建 This study
<i>psk-syn</i> 0311		本研究构建 This study

春雷霉素(kasugamycin)购自上海晶纯生化科技股份有限公司,引物合成和基因测序由上海生工生物工程有限公司完成。总 DNA 抽提试剂盒购自麦克莱博生物科技有限公司。

1.2 序列比对和进化分析

用于分析的 37 种蓝细菌及其他细菌的 16S rDNA、核酸序列来源于 NCBI GenBank 数据库和 Cyanobase 数据库。采用 Clustal W (version2.0.1)^[13] 软件对序列进行多序列比对,通过 ML(maximum likelihood)方法由 MEGA 5.2 软件构建系统进化树,进化树以 Bootstrip10000 自检拟合,得出各分支置信值。16S rDNA 使用模型为 Kimura2-parameter, *ksgA* 基因使用模型为 generalized time reversible (GTR)^[14]。

1.3 蛋白质三维结构同源建模

通过 SWISS MODEL(<http://swissmodel.expasy.org//SWISS-MODEL.html>)分析和预测得到 *ksgA* 的蛋白质模型^[15]。由 QMEAN server(<http://swissmodel.expasy.org/qmean/cgi/index.cgi>)

对蓝细菌 KsgA 蛋白质模型的三维结构合理性进行评估^[16]。

1.4 染色体 DNA 提取

收集细胞,破碎后采用总 DNA 抽提试剂盒抽提蓝细菌总 DNA。

1.5 PCR 与克隆的构建

本研究所用 PCR 引物序列见表 2。提取 *Synechocystis* sp. strain PCC 6803、*Synechococcus elongatus* sp. strain PCC 7942 和 *Anabaena* sp. strain PCC 7120 3 种蓝细菌的总 DNA,以 3 种总 DNA 为模板分别进行 PCR 扩增,得到 3 种蓝细菌 *ksgA* 基因的 PCR 产物,长度分别为 875、860、836 bp。反应条件:95 ℃,预变性 5 min;94 ℃,1 min;63 ℃,30 s;68 ℃,2 min;30 个循环;68 ℃延伸 10 min。限制性内切酶 *Pst* I、*Xho* I 酶切 PCR 产物,将其插入 pBluescript II SK(-)酶切位点 *Pst* I、*Xho* I 之间,常规化学转化并挑取抗性转化子,培养后抽提质粒,

表 2 本研究使用的引物
Table 2 Primers used in this study

引物 Primer	序列 Sequence	特点 Description
<i>alr3229-psk-S</i>	CCTTCTCGAGATGGTACGACCGC-CAAGCTCTTTG	<i>Anabaena</i> sp. strain PCC 7120 <i>ksgA</i> 基因 <i>alr3229</i> PCR 上游引物 Upstream primer of <i>alr3229</i>
<i>alr3229-psk-A</i>	CCTTCTGCAGTTACTCTACTCCAA-CAAGTTTGCT	<i>Anabaena</i> sp. strain PCC 7120 <i>ksgA</i> 基因 <i>alr3229</i> PCR 下游引物 Downstream primer of <i>alr3229</i>
<i>sll0708-psk-S</i>	CCTTCTCGAGATGGCTTTTCGAC-CCCGCAAACGCT	<i>Synechocystis</i> sp. strain PCC 6803 <i>ksgA</i> 基因 <i>sll0708</i> PCR 上游引物 Upstream primer of <i>sll0708</i>
<i>sll0708-psk-A</i>	CCTTCTGCAGTTAGGTTGGGGGTA-AAAAAGTCGGC	<i>Synechocystis</i> sp. strain PCC 6803 <i>ksgA</i> 基因 <i>sll0708</i> PCR 下游引物 Downstream primer of <i>sll0708</i>
<i>syn0311-psk-S</i>	CCTTCTCGAGATGCCTGCTCG-CAAACGCTTTGGTC	<i>Synechococcus elongatus</i> sp. strain PCC 7942 <i>ksgA</i> 基因 <i>synpcc7942-0311</i> PCR 上游引物 Upstream primer of <i>synpcc7942-0311</i>
<i>syn0311-psk-A</i>	CCTTCTGCAGTCATGCAGTCGGGTTA-ATGGCGCTG	<i>Synechococcus elongatus</i> sp. strain PCC 7942 <i>ksgA</i> 基因 <i>synpcc7942-0311</i> PCR 下游引物 Downstream primer of <i>synpcc7942-0311</i>

酶切验证,得到重组质粒 *psk-sll0708*、*psk-syn0311*、*psk-alr3229*。

1.6 大肠杆菌突变体春雷霉素抗性互补检测

将重组质粒 *psk-sll0708*、*psk-syn0311*、*psk-alr3229* 经常规化学转化,转入大肠杆菌 *ksgA* 突变体中得到互补菌株。将 pBluescript II KS 经常规化学转化,转入大肠杆菌 *ksgA* 突变体中得到空载体对照菌株。将互补菌株、正对照菌株、负对照菌株和空载体对照菌株在 LB 液体培养基中 37 ℃培养至对数生长期($D_{600}=0.4\sim0.6$),调整不同样品的菌液浓度至 D_{600} 一致,梯度稀释并取菌液 5 μ L 点至含春雷霉素 400 mg/L 的 LB 固体平板上,37 ℃培养过

夜,观察细菌生长状态。

2 结果与分析

2.1 蓝细菌 *ksgA/dim1* 同源基因的进化分析

通过 Blast 软件分析及相关基因注释,确定了 37 种全基因组测序蓝细菌的 *ksgA/dim1* 的同源基因(以下称蓝细菌 *ksgA*),各基因的具体描述见表 3。*ksgA* 基因在 37 种蓝细菌的染色体组中均有且仅有 1 份拷贝。

对 37 种蓝细菌 16S rDNA 和 *ksgA* 同源基因的核酸序列进行了系统发生分析,选取大肠杆菌、枯草芽胞杆菌、绿硫菌、新月柄杆菌、拟诺卡氏菌及红

表 3 37 种蓝细菌假定的 *ksgA* 基因
Table 3 List of putative *ksgA* from thirty seven cyanobacterial genomes

物种 Species	<i>ksgA</i> 基因 <i>ksgA</i> gene	长度/bp Length
<i>P. marinus</i> MED4	PMM0933	1 084
<i>P. marinus</i> MIT9312	PMT9312-0866	1 081
<i>P. marinus</i> MIT9313	PMT0621	843
<i>P. marinus</i> SS120	Pro0763	1 092
<i>P. marinus</i> str. AS9601	A9601-09271	1 073
<i>P. marinus</i> str. MIT9211	P9211-07111	1 103
<i>P. marinus</i> str. MIT9215	P9215-09571	1 073
<i>P. marinus</i> str. MIT9301	P9301-09521	1 073
<i>P. marinus</i> str. MIT9303	P9303-16171	1 095
<i>P. marinus</i> str. MIT9515	P9515-10161	1 081
<i>P. marinus</i> str. NATL1A	NATL1-09471	1 138
<i>P. marinus</i> str. NATL2A	PMN2A-0278	1 138
<i>A. platensis</i> NIES-39	NIES39-D07300	1 073
<i>S. elongatus</i> PCC 6301	<i>syc1202-d</i>	1 092
<i>S. elongatus</i> PCC 7942	<i>synpcc7942-0311</i>	840
<i>S.</i> CC9311	<i>sync-1594</i>	1 084
<i>S.</i> CC9605	<i>syncc9605-1187</i>	1 073
<i>S.</i> CC9902	<i>syncc9902-1283</i>	1 073
<i>S.</i> JA-2-3B'a(2-13)	CYB-0986	1 103
<i>S.</i> JA-3-3Ab	CYA-0917	1 092
<i>S.</i> PCC 7002	<i>synPCC 7002-A0897</i>	822
<i>S.</i> RCC307	<i>synRCC307-1315</i>	1 073
<i>S.</i> WH7803	<i>synWH7803-1366</i>	1 095
<i>S.</i> WH8102	<i>synW1052</i>	909
<i>S.</i> PCC 6803	<i>sl10708</i>	855
<i>T. elongates-1</i>	<i>tlr0659</i>	798
<i>T.</i> IMS101	<i>Tery-0708</i>	1 100
<i>G. violaceus</i> PCC 7421	<i>gll1991</i>	1 065
<i>M. aeruginosa</i> NIES-843	MAE-04530	1 049
<i>N. punctiforme</i> ATCC29133	<i>Npun-R4912</i>	1 119
<i>C.</i> PCC 8801	PCC 8801-4416	1 065
<i>C.</i> PCC 7424	PCC 7424-4477	1 068
<i>C.</i> PCC 7425	<i>Cyan7425-3485</i>	1 111
<i>C.</i> ATCC51142	<i>cce-2498</i>	1 057
<i>A. marina</i> MBIC11017	AM1-1753	1 052
<i>A.</i> PCC 7120	<i>atr3229</i>	816
<i>A. variabilis</i> ATCC29413	<i>Ava-4886</i>	816

假单胞菌 6 种模式生物作为对照。由于用最大似然法和邻接法构建出来的进化树的高级拓扑结构一致,此处仅展示最大似然法的结果。

16S rDNA 序列进化树(图 1)显示蓝细菌可聚集为 3 个大的分支,其中第 1 个分支为海洋聚球蓝细菌和海洋原绿蓝细菌,分离于进化早期;第 2 个分支的蓝细菌形态和生境多样,包含所有已知固氮蓝细菌,其中固氮单细胞蓝细菌与固氮丝状蓝细菌(产异形胞)的位于不同的分支。第 3 分支为嗜热蓝细菌。

蓝细菌 *ksgA* 基因进化树(图 2)则显示,蓝细菌 *ksgA* 基因聚集为 2 个大的分支,海洋聚球蓝细菌和

海洋原绿蓝细菌仍位于同一分支,而在 16S rDNA 进化树上的第 2 支蓝细菌和嗜热蓝细菌在 *ksgA* 树中同属一个大分支下。除这部分差异外,2 个进化树拓扑结构高度相似,相互作用的基因之间的进化关联是蛋白质进化过程中对结构和功能进行自然选择的结果,拓扑结构的相似说明在蓝细菌中 *ksgA* 基因基本与 16S rRNA 在进化上存在着关联。此外,蓝细菌 *Synechococcus elongates* PCC 7942 与 *Synechococcus elongates* PCC 6301 在 16S rDNA 序列进化树上与海洋蓝细菌接近,而在 *ksgA* 基因进化树中与淡水嗜热蓝细菌同属一个分支。

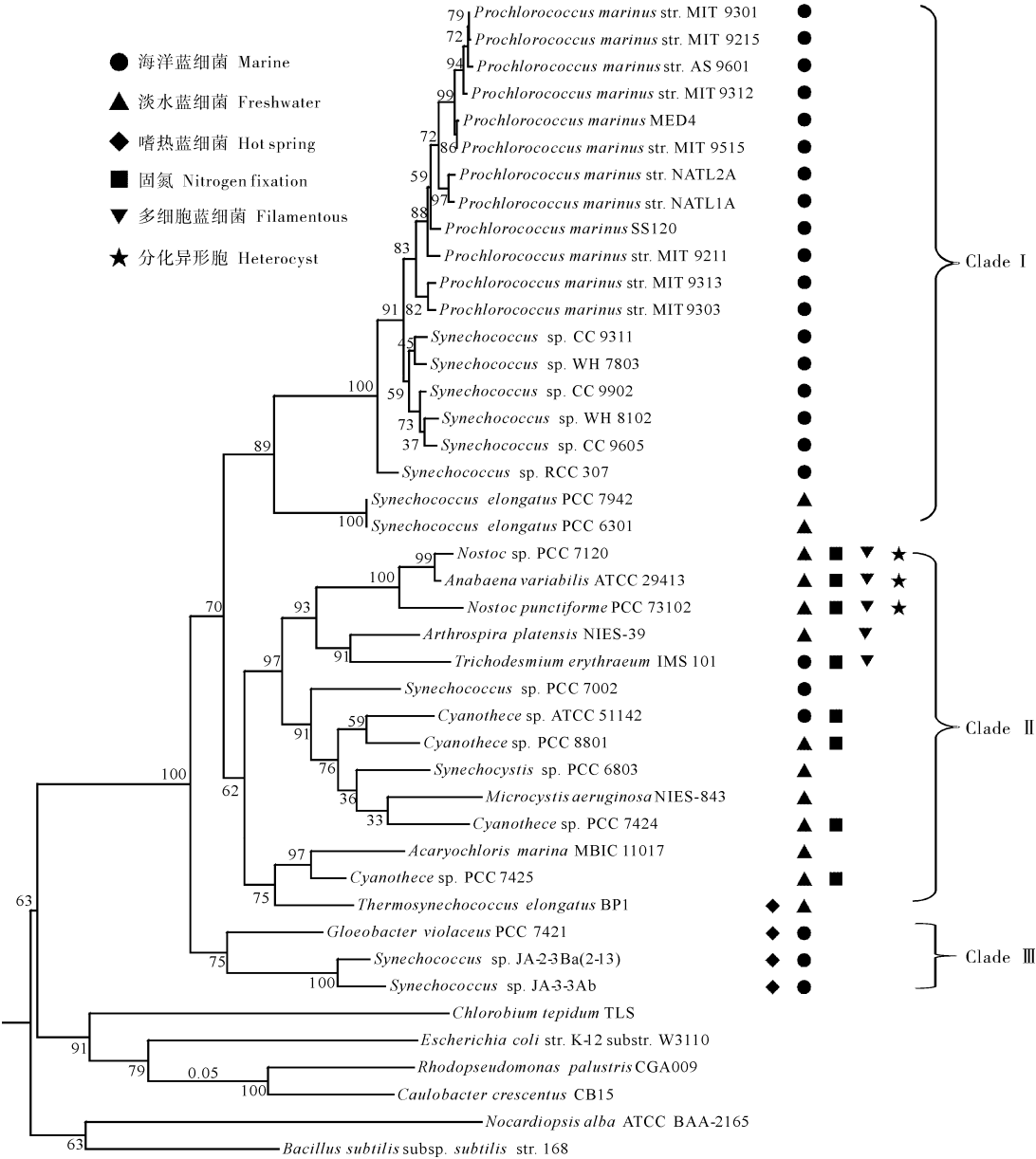


图 1 蓝细菌 16S rDNA 的最大似然进化树

Fig.1 Maximum-likelihood tree of 16S rDNA of cyanobacteria

2.2 蓝细菌 KsgA 蛋白结构特征

为更好地理解蓝细菌 KsgA 双甲基化酶的结构和功能在进化中的关联性,本研究通过同源建模构建了 37 种蓝细菌 KsgA 双甲基化酶的三维结构模型。以 *Anabaena* sp. strain PCC 7120 的 KsgA 双甲基化酶为例(图 3A),蛋白三维结构模型的模板为 *E. coli* KsgA(SMTL id 3tpz.2.A),序列相似度为 34.63%。蓝细菌 KsgA 双甲基化酶拥有保守的结构(图 3B),包括 9~12 个 α 螺旋和 7 个 β 折叠片,其中第 6 个 β 折叠片与第 7 个 β 折叠片之间为逆向平行,其他为正向平行。整体上,蛋白可分为 N

端和 C 端 2 个结构域(图 3A)。N 端含有可能的催化活性位点,与 RNA 甲基化酶 ErmC、NSUN4、Trm14 存在相同的保守结构域,即 3 个 α 螺旋围绕 7 个 β 折叠片,其他 α 螺旋位于第 6、第 7 个 β 折叠片之间围绕着活性位点区域。蛋白质的 C 端结构域整体均略小于 N 端,可能参与决定 RNA 底物特异性。鉴于 KsgA 蛋白质不存在明显的特异性核酸结合区域,因此,C 端可能通过与核糖体蛋白质的互作来识别修饰区域^[17]。

蓝细菌 N 端结构域中的 SAM 结构域也十分保守,且同其他 SAM 依赖的甲基化酶的 SAM 结构域

结构相似。在 *Anabaena* sp. strain PCC 7120 的 KsgA 双甲基化酶中,Gln 10、Glu 36、Glu 59、Ile 60、Asp 84、Asn 109 等 6 个氨基酸残基参与了与 SAM 的结合,在其他蓝细菌的 KsgA 双甲基化酶中,这 6 个氨基酸及其相对空间位置都是保守的(图 3C、3D)。

对 rRNA 甲基化酶 ErmC 和小鼠 KsgA 家族蛋白 mmTFB1M 的结构解析表明,N 端结构域的底物结合区域的 1 个芳香族氨基酸残基非常保守,且

很可能是这些甲基化酶的催化核心^[18]。根据蛋白质相应的空间结构和化学特性,在蓝细菌中,KsgA 双甲基化酶的 1 个位于第 4、第 5 个 β 折叠片之间的 α 螺旋末端的酪氨酸残基很可能是其催化位点(图 3A),这个酪氨酸(Tyr)残基亦保守存在于大肠杆菌及一些原核生物 ksgA 双甲基化酶中,但不同于哺乳动物中 KsgA 家族蛋白在同一结构位置的苯丙氨酸(Phe)残基。

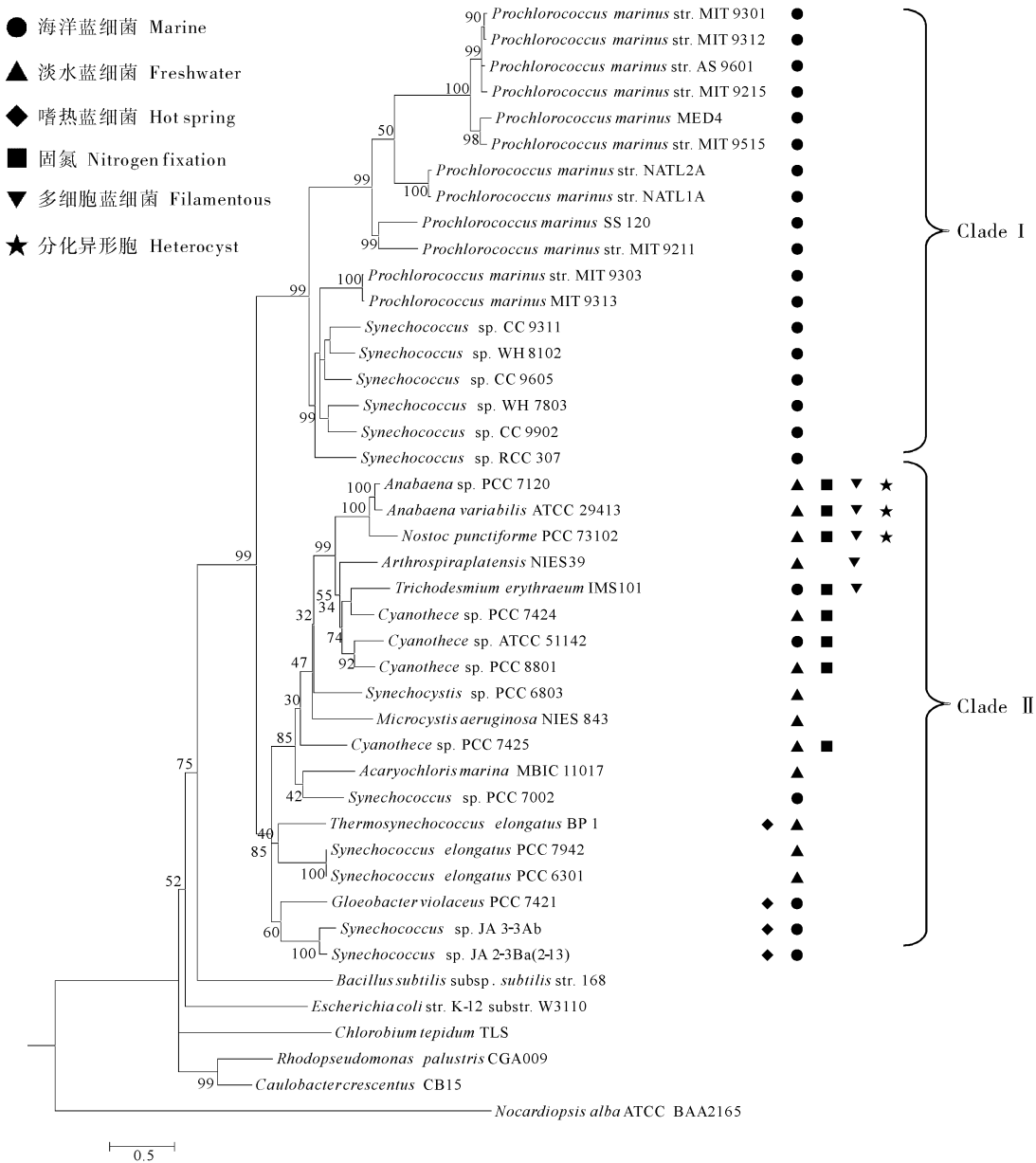
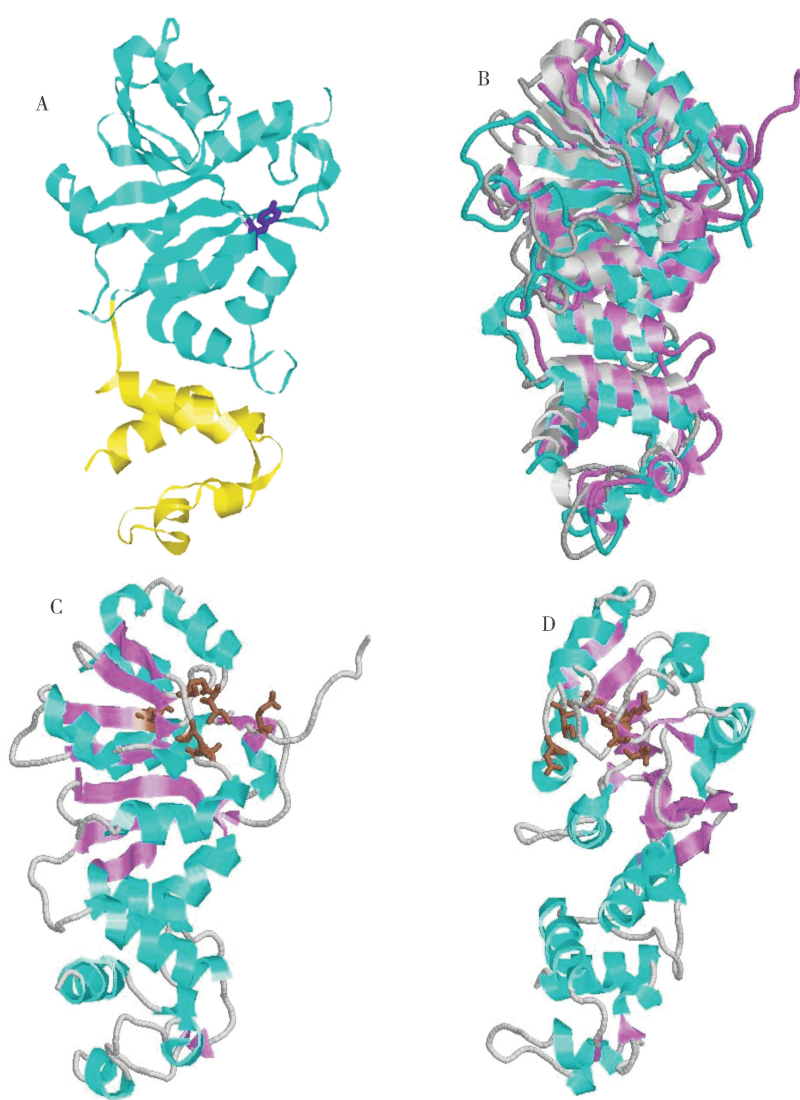


图 2 蓝细菌 ksgA 的最大似然进化树

Fig.2 Maximum-likelihood tree of putative ksgA of cyanobacteria



A: *Anabaena* sp. strain PCC 7120 的 KsgA 蛋白三维结构模型,蓝色为 N 端结构,黄色为 C 端结构,紫色显示的分子为对催化非常重要的酪氨酸残基 Tyr; B:*Anabaena* sp. strain PCC 7120(紫色)、*Synechocystis* sp. strain PCC 6803(青色)和 *Synechococcus elongates* sp. strain PCC 7942(灰色)3 种蓝细菌 KsgA 蛋白三维结构模型的对比; C: *Anabaena* sp. strain PCC 7120 KsgA 蛋白的 SAM-Pocket 参与结合的保守氨基酸残基(褐色),蓝色为 α 螺旋结构,紫色为 β 折叠片; D:图 C 水平顺时针旋转 90° 。A:3D structure models of KsgA dimethyltransferase of *Anabaena* sp. strain PCC 7120. It adopts a methyltransferase fold with two-domain architecture. C-terminal lobe is yellow. N-terminal domain is blue. Tyr112 (violet) is important for catalysis; B:*Anabaena* sp. strain PCC 7120 is structurally similar to other cyanobacteria. Overlay between *Anabaena* sp. strain PCC 7120 (violet), *Synechocystis* sp. strain PCC 6803 (cya) and *Synechococcus elongates* sp. strain PCC 7942 (gray); C: SAM-binding pocketin the N-terminal domain of KsgA of *Anabaena* sp. strain PCC 7120(brown); C 90° rotation is shown in D.

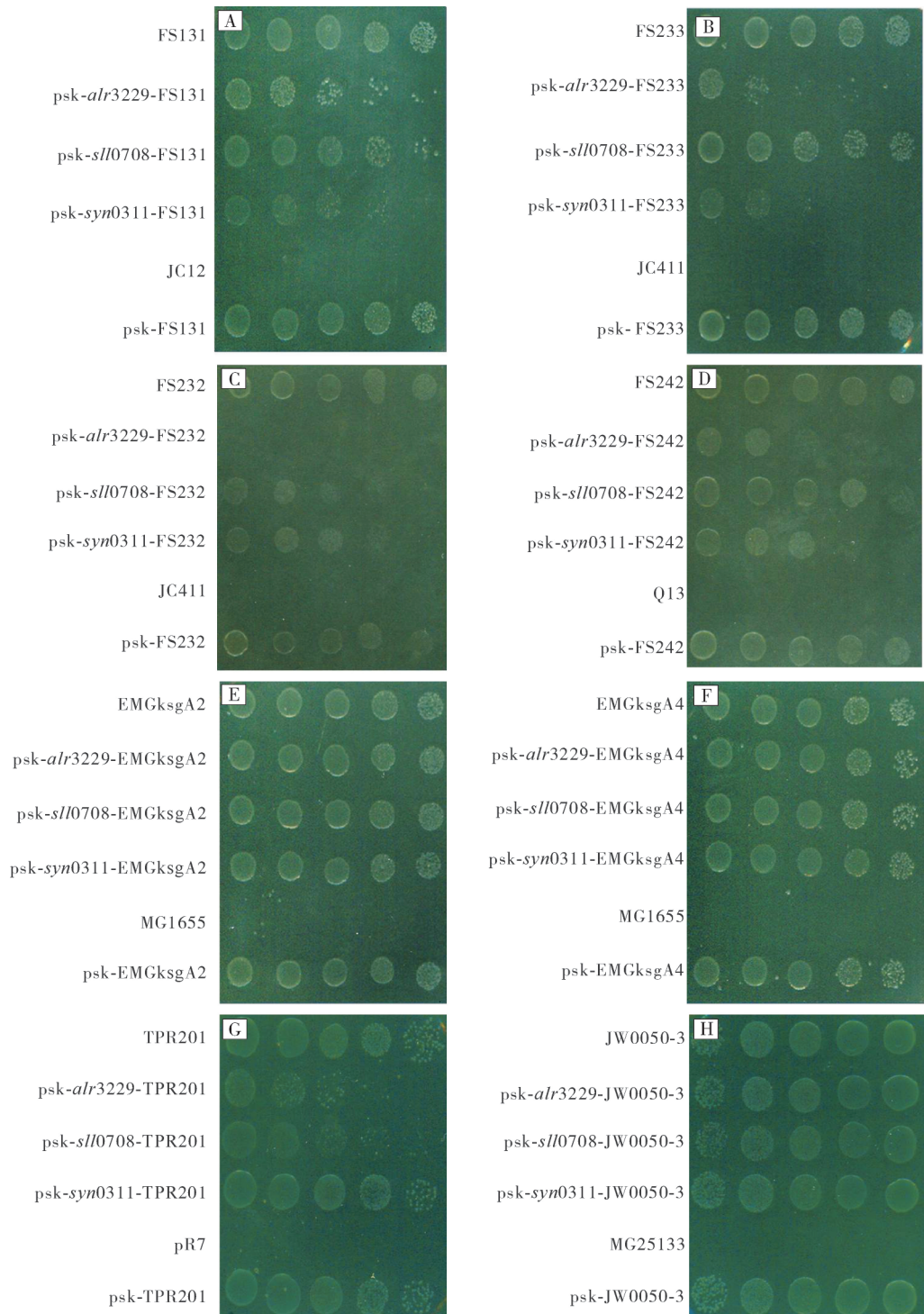
图 3 蓝细菌 KsgA 蛋白的结构模型

Fig.3 3D structure models of KsgA dimethyltransferase from cyanobacteria strains

2.3 蓝细菌 *ksgA* 突变体异源互补检测结果

为了探索蓝细菌 *ksgA* 基因的功能及其在进化过程中是否在功能上存在演化差异,本研究选取了单细胞的集胞蓝细菌 *Synechocystis* sp. strain PCC 6803、丝状能分化异形胞固氮的鱼腥蓝细菌 *Anabaena* sp. strain PCC 7120,以及在 16S rRNA 和 *ksgA* 系统进化树中聚类分支差异较大的单细胞棒

状细长聚球蓝细菌 *Synechococcus elongates* sp. strain PCC 7942,克隆它们的 *ksgA* 基因,在大肠杆菌的 *ksgA* 突变体中进行异源互补实验。大肠杆菌 *ksgA* 突变体对抗生素春雷霉素存在抗性,可以在含春雷霉素 400 mg/L 的 LB 培养基中生长。本研究使用了 8 株大肠杆菌 *ksgA* 突变株,包括缺失突变和 7 个不同位点的点突变。如果蓝细菌 *ksgA* 基因



A-H 依次为突变体 FS131、FS233、FS232、FS242、EMGksgA2、EMGksgA4、TPR201、JW0050-3 的大肠杆菌互补结果。图 4A 第一排为突变体 $10^{-1} \sim 10^{-5}$ 的梯度稀释结果，2 到 4 排依次为 *Anabaena* sp. strain PCC 7120、*Synechocystis* sp. strain PCC 6803、*Synechococcus elongates* sp. strain PCC 7942 的 *ksgA* 基因互补菌株的 $10^{-1} \sim 10^{-5}$ 的梯度稀释结果。5 到 6 排为负对照和空载对照的 $10^{-1} \sim 10^{-5}$ 的梯度稀释结果。图 B-H 的样品与图 A 排序一致。The functional complementation assay was used to express *ksgA* from cyanobacteria *Synechocystis* sp. strain PCC 6803, *Anabaena* sp. strain PCC 7120 and *Synechococcus elongates* sp. strain PCC 7942 in *Escherichia coli* mutant FS131(A), FS233(B), FS232(C), FS242 (D), EMGksgA2(E), EMGksgA4(F), TPR201(G), JW0050-3(H). From left to right is gradient dilution from 10^{-1} to 10^{-5} .

图 4 蓝细菌 *ksgA* 基因对大肠杆菌突变体的互补

Fig.4 Functional complementation of *E. coli* *ksgA* mutanats by cyanobacteria *ksgA*

能够互补成功,则这些大肠杆菌 *ksgA* 突变菌株将不再具有春雷霉素抗性。

互补实验结果显示, *Synechocystis* sp. strain PCC 6803 的 *ksgA* 同源基因 *sll0708* 可以互补点突变菌株 TPR201; 部分互补点突变菌株 FS131、FS232、FS233、FS242; 不能互补点突变菌株 EMGksgA2、EMGksgA4 和缺失突变菌株 JW0050-3。 *Anabaena* sp. strain PCC 7120 的 *ksgA* 同源基因 *alr3229* 可以互补点突变菌株 TPR201、FS131、FS232、FS233、FS242; 不能互补点突变菌株 EMGksgA2、EMGksgA4 和缺失突变菌株 JW0050-3。 *Synechococcus elongates* sp. strain PCC 7942 的 *ksgA* 同源基因可以互补点突变菌株 FS131、FS232、FS233、FS242; 不能互补点突变菌株 TPR201、EMGksgA2、EMGksgA4 和缺失突变菌株 JW0050-3(图 4)。

3 讨 论

在本研究中, 37 种蓝细菌在 16S rDNA 和 *ksgA* 的进化树中均大体反映出了蓝细菌的聚类特征, 体现出与生境、形态、光合系统、生理功能等相关属性的显著关联性, 显示出典型的生物分子和功能结构共同进化的特征。海洋原绿蓝细菌属和聚球蓝细菌属蓝细菌在 16S rRNA 和 *ksgA* 进化树中均独立分支, 清晰地显示其由同一祖先在进化早期独立分离, 在海洋中繁衍, 并进一步演化为原绿蓝细菌属和聚球蓝细菌 2 个分支^[19]。在淡水蓝细菌 16S rRNA 和 *ksgA* 聚类中, 与前人的报道相一致, 已知的多细胞蓝细菌有着共同的起源, 尽管固氮蓝细菌似乎存在着不同的起源, 但是发育异形胞的固氮蓝细菌之间有着很近的亲缘关系并起源于同一祖先^[19]。

不同于其他的甲基化酶, KsgA/Dim1 家族的双甲基化酶需要识别 2 种底物, 腺嘌呤和 N6-单甲基腺嘌呤, 分 2 步使 2 个腺嘌呤先后进入催化活性位点, 一共催化 4 个甲基基团转移到 2 个腺嘌呤上。尽管部分真核生物的 KsgA/Dim1 家族的同源蛋白线粒体转录因子 B 已被证实仍具有双甲基化酶的活性并能够互补大肠杆菌 *ksgA* 缺失突变体的表型^[10], 但本研究发现蓝细菌中的 *ksgA* 基因并不能很好地实现大肠杆菌 *ksgA* 缺失突变体的表型互补。这或许是由于 2 种蛋白在结构上存在局部细节的差异, 导致蓝细菌 KsgA 蛋白对大肠杆菌 16S

rRNA 底物无法识别。而大肠杆菌 *ksgA* 点突变体的互补结果则显示蓝细菌 *ksgA* 基因应该具有甲基化酶的功能。两者结果的差异可能源于大肠杆菌突变体的点突变蛋白并非完全失去活性, 而蓝细菌的异源蛋白并不能在大肠杆菌中正常地实现功能, 它们可能分别交叉参与了复杂的多步催化反应的不同步骤, 从而实现了部分功能和表型的互补。

一般认为, 功能越重要的蛋白质在进化过程中越趋于保守。KsgA/Dim1 蛋白质家族作为少数在古菌、细菌、真核生物中高度保守的蛋白质家族是研究蛋白质结构和功能进化的理想材料^[20]。但是, 目前对于 KsgA/Dim1 功能的研究并不能很好地解释其在三界进化过程中极端的保守性。因而, 对于这一问题的解答仍需对蛋白质结构和功能进化机制更深入的了解。

参 考 文 献

- [1] PAL C, PAPP B, LERCHER M J. An integrated view of protein evolution[J]. Nat Rev Genet, 2006, 7(5): 337-348.
- [2] O'FARRELL H C, MUSAYEV F N, SCARSDALE J N, et al. Crystallization and preliminary X-ray diffraction analysis of KsgA, a universally conserved RNA adenine dimethyltransferase in *Escherichia coli*[J]. Acta Crystallogr D Biol Crystallogr, 2003, 59: 1490-1492.
- [3] HELSER T L, DAVIES J E, DAHLBERG J E. Change in methylation of 16S ribosomal RNA associated with mutation to kasugamycin resistance in *Escherichia coli* [J]. Nat New Biol, 1971, 233(35): 12-14.
- [4] CONNOLLY K J, RIFE P, CULVER G. Mechanistic insight into the ribosome biogenesis functions of the ancient protein KsgA [J]. Mol Microbiol, 2008, 70(5): 1062-1075.
- [5] VAN BUUL C P, HAMERSMA M, VISSER W, et al. Partial methylation of two adjacent adenosines in ribosomes from *Euglena gracilis* chloroplasts suggests evolutionary loss of an intermediate stage in the methyl-transfer reaction [J]. Nucleic Acids Res, 1984, 12(23): 9205-9208.
- [6] O'FARRELL H C, RIFE J P. *Staphylococcus aureus* and *Escherichia coli* have disparate dependences on KsgA for growth and ribosome biogenesis [J]. BMC Microbiol, 2012, 12: 244.
- [7] DUFFIN P M, SEIFERT H S. KsgA mutations confer resistance to kasugamycin in *Neisseria gonorrhoeae* [J]. Int J Antimicrob Agents, 2009, 33(4): 321-327.
- [8] TOKUHISA J G, VIJAYAN P, FELDMANN K A, et al. Chloroplast development at low temperatures requires a homolog of DIM1, a yeast gene encoding the 18S rRNA dimethylase [J]. Plant cell, 1998, 10(5): 699-711.
- [9] SHADEL G S, CLAYTON D A. A *Saccharomyces cerevisiae* mitochondrial transcription factor, sc-mtTFB, shares features

- with sigma factors but is functionally distinct [J]. *Mol Cell Biol*, 1995, 15(4): 2101-2108.
- [10] FALKENBERG M, GASPARI M, RANTANEN A, et al. Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA [J]. *Nat Genet*, 2002, 31(3): 289-294.
- [11] STANIER R Y, COHEN-BAZIRE G. Phototrophic prokaryotes: the cyanobacteria [J]. *Annu Rev Microbiol*, 1977, 31: 225-274.
- [12] CHI X Q, YANG F, ZHAO S, et al. Comparative analysis of fatty acid desaturases in cyanobacterial genomes [J]. *Comp Funct Genomics*, 2008; 284508 [2015-05-24]. <http://dx.doi.org/10.1155/2008/284508>.
- [13] LARKIN M A, BLACKSHIELDS G, BROWN N P, et al. Clustal W and Clustal X version 2.0 [J]. *Bioinformatics*, 2007, 23(21): 2947-2948.
- [14] HALL B G. Building phylogenetic trees from molecular data with MEGA [J]. *Mol Biol Evol*, 2013, 30(5): 1229-1235.
- [15] ARNOLD K, BORDOLI L, KOPP J, et al. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling [J]. *Bioinformatics*, 2006, 22(2): 195-201.
- [16] BENKERT P, KUNZLI M, SCHWEDE T. QMEAN server for protein model quality estimation [J]. *Nucleic Acids Res*, 2009, 37: 510-514.
- [17] BOEHRINGER D, O'FARRELL H C, RIFE J P, et al. Structural insights into methyltransferase KsgA function in 30S ribosomal subunit biogenesis [J]. *J Biol Chem*, 2012, 287(13): 10453-10459.
- [18] GUJA K E, VENKATARAMAN K, YAKUBOVSKAYA E, et al. Structural basis for S-adenosylmethionine binding and methyltransferase activity by mitochondrial transcription factor B1 [J]. *Nucleic Acids Res*, 2013, 41(16): 7947-7959.
- [19] URBACH E D, SCANLAN J, DISTEL D J, et al. Rapid diversification of marine picophytoplankton with dissimilar light-harvesting structures inferred from sequences of *Prochlorococcus* and *Synechococcus* (Cyanobacteria) [J]. *J Mol Evol*, 1998, 46(2): 188-201.
- [20] PARK A K, KIM H, JIN H J. Comprehensive phylogenetic analysis of evolutionarily conserved rRNA adenine dimethyltransferase suggests diverse bacterial contributions to the nucleus-encoded plastid proteome [J]. *Mol Phylogenet Evol*, 2009, 50(2): 282-289.

Phylogenetic analysis and structure prediction of rRNA adenine dimethyltransferase(*ksgA*) in cyanobacteria

WANG Jinglan HU Sheng WANG Li CHEN Wenli

State Key Laboratory of Agricultural Microbiology, Huazhong Agricultural University, Wuhan 430070, China

Abstract Our research paid attention to molecular evolution and structure of KsgA/Dim1 family in cyanobacteria, which catalyzes the dimethylations of two neighboring adenines in small ribosomal subunit rRNA. The phylogenetic analysis of 16S rDNA and *ksgA* gene from MEGA 5.2 showed evolutionary conservation of KsgA/Dim1 family in cyanobacteria. The homology model of KsgA protein constructed by SWISS-MODEL indicated that the fold of KsgA of cyanobacteria is two-domain architecture and relatively well conserved. The functional complementation assay was used to express KsgA from cyanobacteria *Synechocystis* sp. strain PCC 6803, *Anabaena* sp. strain PCC 7120 and *Synechococcus elongates* sp. strain PCC 7942 in *Escherichia coli* mutant to detect dimethyltransferase activity. The results indicated that the functions of *E. coli ksgA* mutanats was complemented by cyanobacteria *ksgA* to some extent.

Keywords *ksgA*; cyanobacteria; dimethylations; sequence analysis; protein structure prediction; SWISS-MODEL

(责任编辑:张志钰)