微生物学通报 Microbiology China tongbao@im.ac.cn



细菌萜类合成酶的生物信息学分析

程甜 郝志强 魏强 李广林^{*} (陕西师范大学 生命科学学院 陕西 西安 710119)

摘 要:【目的】目前对于萜类合成酶(Terpenoid synthase, TPS)的研究主要集中在植物和真菌中, 而对细菌 TPS 的系统研究尚少。建立在大量已经被测序的细菌基因组基础上,利用生物信息学 方法,对细菌 TPS 在全基因组范围内进行识别、分类和功能分析。【方法】利用 TPS 的隐马尔科 夫模型(Pfam 编号为 PF03936)搜索自建的细菌蛋白质组数据库,预测出细菌 TPS。对这些候选 TPS 的蛋白序列用 MAFFT 7.130b 进行多序列比对,并利用 MEGA 6.0 对多序列比对结果进行进 化分析。利用 MEME 和 PredictProtein 分别进行细菌 TPS 的基序(Motifs)和点突变分析。【结果】 建立在生物信息学分析的基础上,1423条细菌 TPS 被识别,它们分布在 8 个门中,即放线菌门 (Actinobacteria)、变形菌门(Proteobacteria)、蓝藻门(Cyanobacteria)、拟杆菌门(Bacteroidetes)、厚 壁菌门(Firmicutes)、绿弯菌门(Chloroflexi)、酸杆菌门(Acidobacteria)和衣原体门(Chlamydiae)。进 化分析表明细菌 TPS 可分为 4 大类, Motifs 分析表明除了各类之间保守的基序(Motifs)外,还有 特异的 Motifs,这暗示着细菌 TPS 在不同类别之间的功能分化。点突变分析表明,细菌 TPS 不 同位点的氨基酸突变对 TPS 功能的影响不同。【结论】细菌 TPS 主要分布于 8 个门中,其中在 2 个门中细菌 TPS 尚未见报道,即厚壁菌门(Firmicutes)与酸杆菌门(Acidobacteria)。基于进化分析, 可以把细菌 TPS 分为 4 类,各类之间的差异可能是由类特异的 Motifs 决定的,另外细菌 TPS 不 同氨基酸位点的突变分析为今后验证 TPS 的功能提供了很好的理论基础。

关键词:细菌,萜类合成酶,生物信息学

Analysis of bacterial terpenoid synthase by bioinformatics

CHENG Tian HAO Zhi-Qiang WEI Qiang LI Guang-Lin*

(College of Life Sciences, Shaanxi Normal University, Xi'an, Shaanxi 710119, China)

Abstract: [Objective] At present, studies on terpenoid synthase mainly focus on plants and fungi. The systematic study of bacterial terpenoid synthase is still rare. Based on a large number of bacterial genomes, we identified bacterial terpenoid synthase in genome-wide and predicted their function by bioinformatics. **[Methods]** Local bacterial proteome database was first built based on the bacterial terpenoid synthase. Then multiple sequence alignment for candidate terpenoid synthase was aligned by the tool of MAFFT 7.130b, and phylogenetic tree was constructed by MEGA 6.0. Finally, analysis of motifs was conducted by MEME and point mutation of TPS by PredictProtein. **[Results]** In total 1 423

基金项目: 国家自然科学基金项目(No. 31370329); 中央高校基本科研业务费(No. GK201403004)

^{*}通讯作者: ⊠: glli@snnu.edu.cn

收稿日期: 2014-12-06; 接受日期: 2015-03-20; 优先数字出版日期(www.enki.net): 2015-04-10

terpenoid synthases were identified; they are distributed in 8 phylum, including: Actinobacteria, Proteobacteria, Cyanobacteria, Bacteroidetes, Firmicutes, Chloroflexi, Acidobacteria and Chlamydiae. Phylogenetic analysis revealed that bacterial terpenoid synthases can be divided into 4 groups. Motifs analysis of each group by MEME showed that in addition to conserved motifs among groups, there were some specific motifs within each group, which implies functional differentiation among different groups. Point mutations analysis showed that mutations of amino acid sites in different positions of terpenoid synthase could have different effect on TPS function. **[Conclusion]** Bacterial terpenoid synthases are mainly distributed in 8 phylum, of which 2 phylum have not been reported before; they are Firmicutes and Acidobacteria. Phylogenetic analysis showed that the difference among 4 groups of terpenoid synthase is mainly caused by group-specific motifs. In addition, mutation analysis of amino acids in different positions of terpenoid synthase function.

Keywords: Bacteria, Terpenoid synthase, Bioinformatics

萜类化合物(Terpenoid)是自然界中结构最为多 样、种类最为丰富的天然产物^[1]。萜类合成酶 (Terpenoid synthase, TPS)是萜类化合物合成的关键 酶。根据异戊二烯(Isoprene, C5)单元数的不同, 萜 类化合物可分为单萜(Monoterpene, C10)、倍半萜 (Sesquiterpene, C15)、二萜(Diterpene, C20)、三萜 (Triterpene, C30)和多萜等。在生物体内, 萜类化 合物可通过两条不同的途径合成,即甲羟戊酸 (Mevalonate, MVA)途径与甲基赤藓糖磷酸 (2-C-methyl-D-erythritol-4-phosphate, MEP)途径。 除了植物可同时利用 MVA 和 MEP 途径外, 大部分 生物体只能利用两条途径中的一种^[2]。细菌主要利 用 MEP 途径进行萜类产物的合成^[3],即首先合成异 戊烯基焦磷酸(Isopentenylallyl diphosphate, IPP)前 体和二甲基丙烯基焦磷酸 (Dimethylallyl diphosphate, DMAPP); 然后 IPP 和 DMAPP 在聚 异戊二烯焦磷酸合成酶(Polyisoprenyl diphosphate synthase)的作用下缩合成长度不同的非环化前体, 即香叶基焦磷酸(Geranyl diphosphate, GPP)、法尼 基焦磷酸(Farnesyl diphosphate, FPP)、香叶基香叶 基焦磷酸(Geranylgeranyl diphosphate, GGPP); 最 后这些前体在 TPS 的催化下环化重排形成萜类化 合物。

萜类化合物具有重要的生理、生态和药用价 值,它们作为激素、味道与气味成分以及抗生素被 广泛应用。长期以来人们在各种物种中不断探究

http://journals.im.ac.cn/wswxtbcn

TPS 的功能。在植物中, TPS 已经在拟南芥 (Arabidopsis thaliana)^[4]、水稻(Oryza sativa)^[5]、杨树 (Populus trichocarpa)^[5]、葡萄(Vitis vinifera)^[5]和番茄 (Solanum lycopersicum)^[6]等物种中被报道。在真菌 中, TPS 已经在娄地青霉(Penicillium roqueforti)^[7]、 土曲霉(Aspergillus terreus)^[8]和葡萄孢菌(Botrytis cinerea)^[9]等物种中被发现。以前人们认为 TPS 只存 在于植物和真菌中^[10],但随着人们对细菌气味与挥 发性物质代谢的探索^[11-13], 细菌 TPS 的研究逐渐引 起人们的重视。Dairi 等^[14]于 2001 年首次在 Kitasatospora griseola MF730-N6 中发现了二萜合 成酶。目前发现的细菌 TPS 主要集中在革兰氏阳性 菌如链霉菌属的细菌以及其他一些放线菌中^[15], 如从链霉菌属 Streptomyces lasaliensis NRRL 3382^[16] 与 Streptomyces coelicolor A3(2)^[17]分别发现能够合 成 2-Methylisoborneol 的基因 tpc 和 sco7700; 从 *Streptomyces coelicolor* A3(2)^[18] 发现能够合成 Epi-isozizaene 的基因 sco5222; 从 Streptomyces melanosporofaciens MI614-43F2^[19]中发现了1个二 萜合成酶基因簇,由 cotB1、cotB2、cotB3、cotB4 组成。

目前对于 TPS 的研究主要集中在植物和真菌中,而对细菌 TPS 的研究尚少。随着越来越多的细菌基因组被测序,单靠传统的实验方法去鉴定细菌 TPS 的功能,已经远远落后于数据增长的速度,迫切需要开发和利用新的手段对细菌 TPS 进行系统

研究。建立在已经被大量测序的细菌基因组基础 上,本文利用生物信息学方法,对细菌 TPS 在全基 因组范围内进行了识别、分类和功能分析,这为系 统理解细菌 TPS 的物种分布、分类、功能和进化奠 定了很好的基础。

1 材料与方法

1.1 数据来源

Pfam 数据库中^[20]标号为 PF03936 的结构域是 TPS 蛋白的特征结构域。首先从 Pfam 数据库中 (http://pfam.xfam.org/)下载此结构域包含的信息, 然 后 从 NCBI (National Central for Biotechnology Information, http://www.ncbi.nlm.nih.gov/genome/ ?term=bacteria)下载4705个完成基因组测序的细菌 的蛋白质序列(共44322195条)构建本地细菌蛋白 质组数据库。

1.2 细菌 TPS 的识别

利用基于隐马尔科夫模型的 HMMER2.3.2^[21] 程序,以 PF03936 为检索模型,搜索本地细菌蛋白 质组数据库中含有 TPS 结构域的候选序列,*E*-value 设为<10^{-5[22-23]}。

1.3 多序列比对与分子系统进化树的构建

利用 MAFFT 7.130b^[24]中的 L-INS-i 方法对搜 集到的细菌 TPS 序列进行多序列比对分析。比对结 果用 MEGA 6.0^[25]软件构建细菌 TPS 蛋白的分子系 统进化树。建树方法为最大似然法(Maximum likelihood)。具体设置参数 Test of Phylogeny: Bootstrap method; Replications: 500; Model: Jones-Taylor-Thornton (JTT) model; Rates among Sites: 4; Gamma Distributed (G); Gaps: Complete deletion。其余参数为默认。

1.4 Motifs 的分析

利用 MEME^[26] (http://meme.nbcr.net/ meme/)对 TPS 蛋白的保守基序(Motifs)进行了图示,具体参数 设置如下: The occurrences of a single motif are distributed among the sequences: Zero or one per sequence; Minimum width (≥2): 6; Maximum width (≤300): 50; Maximum number of motifs to find: 20。 其余参数为默认。

1.5 点突变影响的分析

为了研究各个点突变对整条蛋白序列功能的 影响,选择代表性的细菌 TPS 序列,使用 PredictProtein (http://ppopen.rostl-ab.org/)对其进行 点突变分析,此预测基于 SNAP2,在分析点突变对 功能的影响时准确性能达到 82%^[27]。

2 结果与分析

2.1 细菌萜类合成酶的全基因组范围识别和 分布

为了系统探索细菌 TPS,以 TPS 的隐马尔科夫 模型 HMM (Hidden Markov models) PF03936 为查 找模型,利用 HMMER 软件搜索本地自建的细菌蛋 白质数据库,共得到具有统计学显著性的 1 423 条 匹配序列(*E*-value<10⁻⁵),即为预测的细菌 TPS。

分析表明,这1423条 TPS 来自553个细菌基 因组中,它们主要分布于细菌8个门中(图1A),在 放线菌门(Actinobacteria)的 405 个基因组中发现了 1 206 条 TPS, 占总数的 84.751%, 在变形菌门 (Proteobacteria)的 95 个基因组中发现了 141 条, 在 蓝藻门(Cyanobacteria)的 21 个基因组中发现了 29 条, 在拟杆菌门(Bacteroidetes)的 16 个基因组中发 现了 27 条, 在厚壁菌门(Firmicutes)的 8 个基因组 中发现了9条,在绿弯菌门(Chloroflexi)的5个基因 组中发现了8条,在酸杆菌门(Acidobacteria)的2个 基因组中发现了2条,在衣原体门(Chlamydiae)的1 个基因组中发现了1条,它们所占TPS总数的比例 依次为 9.909%、2.038%、1.897%、0.632%、0.562%、 0.141%和 0.07%。研究结果显示, 细菌 TPS 主要集 中在放线菌中,这与所报道的结果是一致的[15]。与 前人^[10]的研究结果相比,厚壁菌门(Firmicutes)和酸 杆菌门(Acidobacteria)中的细菌 TPS 是我们基于预 测模型发现的。对放线菌门(Actinobacteria) TPS 的 物种分布进一步分析表明,链霉菌科 (Streptomycetaceae)所含序列数目最多(图 1B),达到



图 1 细菌萜类合成酶的分布情况 Figure 1 Distribution of terpenoid synthase in bacteria 注: A: 在各个门中的分布情况; B: 在放线菌门各个科中的分布情况.

Note: A: Distribution for each phylum; B: Distribution for each family of Actinobacteria.

902 条之多,占放线菌门 TPS 总数的 74.793%,其 余科所含序列数目均在 100 条以内。

通常单种细菌含有的 TPS 数在 10 个以内(表 1),但我们的研究发现一些细菌含有更多数目的 TPS,它们的物种分布是:棒状链霉菌(*Streptomyces clavuligerus* ATCC 27064,拷贝数为 26 个)、砂栖海 水小单胞菌(*Salinispora arenicola* CNS-205,拷贝数 为 27 个),分别属于放线菌门(Actinobacteria)的链霉 菌 科 (Streptomycetaceae)和小单胞菌科 (Micromonosporaceae)。这些编码多拷贝的 TPS 基 因可能是由细菌 TPS 基因倍增或者基因水平转移 引起的,它们在细菌 TPS 的产物多样性中扮演着重 要角色,其潜在的分子机制还需要进一步研究。

2.2 细菌萜类合成酶的分子系统进化与功能 预测

为了研究细菌 TPS 的进化关系,根据细菌萜类 合成酶在各个门类中的分布情况选取 109 个典型细

表 1 各个物种中 TPS 基因拷贝数统计分析 Table 1 Statistic analysis of TPS gene copy number in each species										
物种数										
Number of species										
199										
128										
94										
74										
26										
13										
10										
4										
3										
1										
1										

注:拷贝数为单个细菌含有的 TPS 数目;物种数为含有该拷贝数的细菌数目.

Note: Copy numbers mean the numbers of TPS in signal bacteria; Species numbers mean the numbers of bacteria with this copy number.

菌 TPS 以及6个已被实验证实的二萜合成酶进行蛋 白序列分析。首先利用多序列比对软件 MAFFT 对 细菌 TPS 蛋白序列进行比对, 然后利用 MEGA 软 件中的最大似然法构建细菌 TPS 蛋白序列的分子 进化树。根据进化树的拓扑结构,将细菌分为4大 类(图 2)。第1类(Group I)包含 41个序列,分布在 放线菌门 (Actinobacteria)、变形菌门 (Proteobacteria)、蓝藻门(Cyanobacteria)中; 第2类 (Group II)包含 33 个序列,分布在放线菌门 (Actinobacteria)、变形菌门(Proteobacteria)、蓝藻门 (Cyanobacteria)、绿弯菌门(Chloroflexi)、拟杆菌门 (Bacteroidetes)、厚壁菌门(Firmicutes)、酸杆菌门 (Acidobacteria)、衣原体门(Chlamydiae)中; 第3类 (Group III)包含 19 个序列,分布在放线菌门 (Actinobacteria)和变形菌门(Proteobacteria)中; 第 4 类(Group IV)包含 9 个序列分布在放线菌门 (Actinobacteria)、拟杆菌门(Bacteroidetes)、厚壁菌 门(Firmicutes)中;未能分类的包含 13 个序列,分 布在放线菌门 (Actinobacteria) 和酸杆菌门 (Acidobacteria)中。其中放线菌门(Actinobacteria)的 TPS 分布最广,在各大类中都存在,Group II 所含 细菌种类最多,包含了预测到的 TPS 所在的所有门 类。绿弯菌门(Chloroflexi)的 TPS 仅在 Group II 中 出现,这一方面暗示这些 TPS 具有功能特异性,另 一方面可能是由绿弯菌门发现的包含 TPS 的细菌 种类较少造成的(仅有 5 种细菌), 但对其进一步的 探索需要实验验证。

建立在细菌 TPS 分类的基础上(图 2),可以在 每个类别中根据已知功能的细菌 TPS 去预测未知 TPS 的功能。因为在 Group III 中,已知的 TPS 功 能主要为单萜合成酶(Monoterpene synthase),所以 我们推测 Group III 中的未知 TPS 也可能为单萜合 成酶。同样的道理,我们可以推测 Group I 和 Group II 中未知功能的 TPS 主要为倍半萜合成酶 (Sesquiterpene synthase)。Group IV包含了 6 条被实 验证实的二萜合成酶(Diterpene synthase)及 3 条本 研究预测的 TPS, 推测此类可能为二萜合成酶。

2.3 细菌萜类合成酶序列的特征

虽然植物、真菌和细菌中 TPS 的氨基酸序列差 异很大,但是所有的 TPS 序列都具有 2 个高度保守 的金属离子结合基序:第一个基序为-DDxx[D/E]-或-DDxxx[D/E]-,此基序富含酸性氨基酸,在细菌 TPS 中经常位于 N 末端下游 80-120 个氨基酸处, 本研究发现此基序在 Group II中为-DDxx[D/E]-, 而在 Group I 与 Group III中为-DDxxx[D/E]-;第二 个基序为-xxNDxxSxxxE-,通常位于第一基序下游 140±5 aa 处,这与以前报道的结果是一致的^[28]。

值得注意的是我们发现-DDxx[D/E]-或-DDxxx[D/E]-基序的氨基酸组成在不同类别的细菌TPS中具有很大差异(图3,红色方框标记):Group I与Group II的上游基序富含酸性氨基酸并含有高比例的芳香族氨基酸,这个保守基序又分为-FxFDDHFLE-(图3A)和-FFxxDDxxD-(图3B)两种情况。Group III的上游基序为-xVDDxxx[D/E]-(图3C),虽然同样富含酸性氨基酸,但芳香族氨基酸的含量要低于Group I与Group II。相对而言, 下游的保守基序-xxNDxxSxxE-在这3个类群中都 很保守(图3,蓝色方框标记)。虽然Group IV的6 条蛋白序列已被实验证实是二萜合成酶^[29](Diterpene synthase),但序列比对结果表明-DDxx[D/E]-和-xxNDxxSxxxE-的保守性较低。

2.4 细菌萜类合成酶的基序(Motifs)分析

为了进一步分析细菌 TPS 序列在各个类别之间的差异性,利用 MEME 软件对 TPS 序列的 Motifs 进行了系统分析,并得到 20 个基序(图 4)。研究发现,这 20 个基序在不同类群细菌 TPS 中的分布差异很大。Group I 所含基序数目最多,为 17-19 个。Group II 和 Group III所含基序数目分别为 5-7 个和 4-5 个。Group IV所含基序数目最少,为 1-3 个。分析发现 Group I 、Group II、Group III所共有的



图 2 115 个典型细菌 TPS 蛋白序列的系统进化树 Figure 2 Phylogenetic tree of 115 typically bacterial TPS protein sequences

注: 括号中的序号表示 GenBank 蛋白登录号,分支上的数字表示 Bootstrap 值(<30 的隐藏). 进化树中不同颜色的序列名称代表不同 类别的 TPS: 绿色代表 Group II,红褐色代表 Group II,紫色代表 Group III,水绿色代表 Group IV,黑色为无法分类的 TPS 序列. Note: The numbers in parenthesis represent GenBank accession numbers, the numbers on each branch indicate the bootstrap values (hide if <30). Sequences names with different colors in the tree represent different groups of TPS: green represent Group II, reddish brown represent Group II, purple represent Group III, aqua represent Group IV, black represent sequences that can't be classified. A

			_								_			_		
YP 483306.1	75	DLVTD	WYVWVF	YFDDHF	LE YK	RSHD	MAGAR	AYLD	109	.224	ADAVH	lfndi	FSYQRE	/EEF	EGEL	246
YP 003950745.1	83	DLVTD	WYVWVF	YFDDHF	'LD /YK	RIQD	QVGAR	FYLD	117	.233	SDGVH	lfndi	FSYQRE	. QEF	EGEL	255
YP 634376.1	83	DLVTD	WYVWVF	YFDDHF	'LE LYK	RPQD	QVGAK	AYLD	117	.233	SDAVH	lfndi	FSYERE	LEF	EGEL	255
YP 001866236.1	77	DLVTD	WYVWVF	FFDDHF	'LE IYKI	RIQDI	MAGAK	EYLG	111	.227	ADGVH	lfndi	FSYQRE	7DDF	EGEN	249
CBN59287.1	77	DLVTD	WYVWVF	FFDDHF	'LE IYKI	RIQDI	MIGAK	EYLH	111	.227	ADGVH	LFNDI	FSYQRE	7DDE	EGEN	249
YP 001612078.1	75	NLITD	WYVWVF	FFDDHF	'LE LYKI	RT <mark>K</mark> DI	MRGAK	AYLD	109	.225	SDGVH	LFNDI	FSYQRE	VEVE	EGEN	247
NP 823339.1	72	SLITD	WYVWVF	FFDDHE	LE FK	RIQDI	RECCK	AYLD	106	.221	SDGVH	LFNDI	FSYQRE	/DEF	EGEL	243
в												_		_	_	
NP_824174.1	67	DLGV	LMSWF	FLFDDL	FDGPR	GE-D	PQETF	KLTI	99	.212	TDTNI	II ND.	IASLEK	e ear	RGEL	233
YP_004967928.1	67	DLGV	LMSWF	FLFDDL	FDGPR	GE-N	PEETK	QITI	99	.212	VDVNI	LLND	IASLEK	e ear	۶GEQ	233
NP_629369.1	86	QAIA	YSAWF	FVWDDR	HDRDI	VHGR	AGAWF	RIRO	119	.233	QEFAA	WY NDI	LCSIPK	EIAG	DEV	254
WP_003992851.1	86	QAIA	YSAWF	FVWDDR	HDRDI	VHGR	PVAWF	RLRI	119	.233	QEFAA	WY NDI	LCSLPK	EIAG	DEV	254
WP_004925880.1	86	QAIA	YSAWF	FVWDDR	HDRDI	VHGR	HAAWI	RIKI	119	.233	QEFAA	WY NDI	LCSIPK	E LAG	SDV	254
WP_007384601.1	86	QAIA	YSAWF	FVWDDR	HDRDI	VHRR	PGAWF	RIRI	119	.233	QEFAA	WY NDI	LCSIPK	EIAG	DDV	254
WP_003950672.1	85	SAIS	FSTWF	FVWDD <mark>O</mark>	HDRDA	VHGR	RTAWY	RISZ	118	.232	QDFSA	WY ND	ICSIPK	D LAG	DDQ	253
С									016	000	<u></u>		VOVTRE	NOD	- DU	200
NP_733742.1	184	MLATR	LMVAEN	IA VDDCY	CEDHG	G	SEVGI	GGRL	215	.338			VOVTRE	DCD	CILL	300
YP_003486275.1	184	MLATR	LMVADN	IA VDDCY	CEDHG	G	SEVGI	GGRL	215	.338	SN A MM		VOVTER	DOP	CHH	200
YP_001822781.1	181	MLATR	LMVAPN	IAVDDCY	CEDHG	G	SEVGI	GERL	212	.335			VOVTRE	LAGD	CHH	357
YP_004815441.1	176	MLATR	LMVADN		CEDHG	G	SEIGI	GGRI	207	.330			VOVTRE	LASP	CHH	332
YP_004967206.1	52	MLATR	LMVADN		CEDHG	G	SEIGI	GGRI	83 01	.200	2NV ST		VOVTER	ASP	CRI	220
WP_007267885.1	50	MIAAR	LMVADN		GEDHG	G	SEIGI	GGRL	250	201			V SVTRF	ASD	KH	404
YP_006882656.1	228	MLATE	LMVADN		GEDHG	G	SEVGL	GGRL	239	.364		TUNDT	VOVTER	ASD	RH	224
WP_008743709.1	58	MIATE	LMVADN		GEDHG	G	SEVGL	EGEL	115	220			VGAVKE	IT.NE	TCD	204
YP_347573.1	82	LLVAQ	CFAALE	AVDDHY		-LGG	REETV	AESL	144	268			VOLEKE	FIN	FD	200
YP_003510780.1	111	LAAAK	CAVADW		LDBDS	-AEP	EFERI	GPRL	144	261	29397		VSMCKE	D		270
YP_003680543.1	195	TABAK	CALABW		VDSEA	-EEA		GQRL	124	257			ASCEPT	rppu	T.D	370
YP_001105919.1	100	TADGE	MIVABI		QE INS.	KKUA.	NDÖJT	EPNI	134	.201				inen	Ent	219
D				EDHV	F		1 106	TET			T.ATT	?	RERAE	D 21	6	
BAD86798.1 6	4	WT	YLLIG	RUDDEE	HYGTRE	DD o	1190	ASI	AWR		TRTW		-KLAFF		0	
AC031279.1 7	1	GR	TVSTE	WISYV	GVVLWS		ZZIZ	 	CIT	TUNE	FYSY)	-REVSL	G 23	2	
BAI44338.1 6	5	WTACI	IVNDD	WD-YV	0		6 211	VVI	REGO	SNE	YYS		-BEKKA	≝ 23 V 22	1	
BAB39207.1 6	8	NAG		ADTHT	SGR		0211 7 100	TC1	CIT	FGR	STWA	TPDT	TGDAKA	- 23 - 21	2	
AC031274.1 7	0	NAG	T	ADTHT.	SCK		7188	TG1	C T	FGF	STW	TDDT	TCDAKA		3	
ADD83014.1 7	0	- ANG	10210	THE THIT	561	2 8	/188	101	10 III	rGr	STW W	TEDT	TODAVA	21	5	

图 3 细菌 TPS 在各个类群的序列比对结果

Figure 3 Alignment of amino acid sequences of bacterial TPS in each group

注: A: Group I 的序列比对结果; B: Group II 的序列比对结果; C: Group III的序列比对结果; D: Group IV的序列比对结果. 红 色方框中为-DDxx[D/E]-或-DDxxx[D/E]-基序,蓝色方框中为-xxNDxxSxxxE-基序.

Note: A: Alignment result in Group I ; B: Alignment result in Group II ; C: Alignment result in Group III; D: Alignment result in Group IV. The sequence in red box is -DDxx[D/E]- or -DDxx[D/E]- motif, the sequence in blue box is -xxNDxxSxxE- motif.

Motifs 为 Motif 2 和 Motif 5 (图 5A); Group I 特有 的 Motifs 为 Motif 15 和 Motif 19 (图 5B); Group III 所特有的 Motif 为 Motif 18 (图 5C), Group II和 Group IV所含的 Motifs 均在其他类群中出现过,无 特异 Motifs。

2.5 点突变对细菌萜类合成酶功能的影响分析

萜类产物是由其 TPS 决定的,而 TPS 的重要 氨基酸位点的点突变通常可导致 TPS 功能(主要指 TPS 的酶活和 TPS 的萜类产物)的改变。为了识别 和理解细菌 TPS 中重要氨基酸位点突变对 TPS 功 能的影响,选取 Group III中的一条序列(登录号为 NP_733742.1)进行点突变分析(图 6)。由于-xxNDxxSxxxE-这个基序在不同细菌 TPS 之间是非常保守的,并且 D 的保守性最强,所以无论 D 被替换成 20 个氨基酸中的哪一个,都会对 TPS 的酶活产生严重的影响;而 D 后第 6 位的 E 替换为 D 时,预测结果为白色,证明此处 E 可与 D 互换,但当 E 替换为除 D 外的其他氨基酸时,则会对 TPS 的酶活产生严重的影响。除-xxNDxxSxxxE-外,其他位置的氨基酸突变也会对细菌 TPS 的功能产生



图 4 典型细菌 TPS 的 Motifs Figure 4 The motifs of typical bacterial TPS

注: 左半部分为 TPS 的分子进化关系; 中间为 TPS motifs 组成结构的 Combined *P*-value; 右半部分为 20 个 Motifs 在 TPS 中的分布. Note: Left part illustrates the molecular phylogenetic relationships among TPS. Middle part indicates combined *P*-value of motifs structure of TPS. Right part shows 20 motifs distribution in TPS protein sequence of bacteria.

http://journals.im.ac.cn/wswxtbcn



图 5 细菌 TPS 在各类群中共有及特有的 Motifs Figure 5 Common and group-specific motifs of bacteria TPS in each group 注: A: Group I、Group II、Group III共有的基序(Motifs); B: Group I 特有的基序(Motifs); C: Group III特有的基序(Motif). Note: A: In common motifs of Group I, Group II and Group III; B: Motifs specific to Group I; C: Motif specific to Group III.



图 6 分析点突变对 NP_733742.1 的功能影响 Figure 6 Analysis of effect of point mutations of NP_733742.1

注: 深红色(>50): 替换后对功能影响很大; 绿色和白色(<50): 替换后对功能比较小; 黑色: 野生型. Note: Deep red (>50): Have great influence on function if replaced; Green and White (<50): Have relatively small influence on function if replaced; Black: Wild type. 很大影响(详见图 6 中红色箭头位置所示的氨基酸 位点 E、A、G、A、N和L),由于这些氨基酸位点 在不同细菌之间是不保守的,所以这些氨基酸位点 的突变可能导致萜类产物的改变。细菌 TPS 的重要 氨基酸位点的识别,为今后定向改变萜类合成酶的 功能,利用代谢工程方法改变目标萜类产物提供了 理论基础。同时,我们的点突变方法也可以应用到 其他细菌 TPS 的功能分析中。

3 结论

鉴于 TPS 具有重要的生理、生态和药用价值, 长期以来人们一直对不同的物种进行 TPS 的识别 和功能研究,并在拟南芥、水稻、杨树、葡萄、番 茄、土曲霉、娄地青霉、葡萄孢菌等许多物种中进 行了较为系统的研究。但相对于植物和真菌 TPS 的 识别和功能研究,细菌 TPS 的系统研究尚少,本研 究通过对细菌 TPS 的系统分析,在厚壁菌门 (Firmicutes)和酸杆菌门(Acidobacteria) 2 个门中发 现了新的 TPS,这为今后利用这些新发现 TPS 进行 功能分析和揭示不同细菌门类之间 TPS 的进化关 系打下了很好的基础。

不同物种 TPS 的功能探究一直是 TPS 研究领 域的热点,为了系统探究细菌 TPS 的特征和功能, 本文利用 MAFFT、MEGA、MEME、PredictProtein 等生物信息学方法,从不同角度对细菌 TPS 进行了 预测和分析,通过分子进化、Motifs 及点突变影响 分析,我们对细菌 TPS 功能差异的可能原因进行了 多层次研究(不同类别之间的差异、Motifs 差异和 TPS 中不同位点氨基酸的差异),这为今后深入研究 细菌 TPS 的功能提供了理论支撑和数据来源。

参考文献

- Daum M, Herrmann S, Wilkinson B, et al. Genes and enzymes involved in bacterial isoprenoid biosynthesis[J]. Current Opinion in Chemical Biology, 2009, 13(2): 180-188
- [2] Vranová E, Coman D, Gruissem W. Network analysis of the MVA and MEP pathways for isoprenoid synthesis[J]. Annual Review of Plant Biology, 2013, 64(4): 665-700
- [3] Rohmer M, Knani M, Simonin P, et al. Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate[J]. The Biochemical Journal, 1993, 295(Pt 2): 517-524

http://journals.im.ac.cn/wswxtbcn

- [4] Aubourg S, Lecharny A, Bohlmann J. Genomic analysis of the terpenoid synthase (AtTPS) gene family of *Arabidopsis thaliana*[J]. Molecular Genetics and Genomics, 2002, 267(6): 730-745
- [5] Chen F, Tholl D, Bohlmann J, et al. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom[J]. The Plant Journal, 2011, 66(1): 212-229
- [6] Falara V, Akhtar TA, Nguyen TT, et al. The tomato terpene synthase gene family[J]. Plant Physiology, 2011, 157(2): 770-789
- [7] Kimura M, Tokai T, Takahashi-Ando N, et al. Molecular and genetic studies of *Fusarium* trichothecene biosynthesis: pathways, genes, and evolution[J]. Bioscience, Biotechnology, and Biochemistry, 2007, 71(9): 2105-2123
- [8] Brase S, Encinas A, Keck J, et al. Chemistry and biology of mycotoxins and related fungal metabolites[J]. Chemical Reviews, 2009, 109(9): 3903-3990
- [9] Pinedo C, Wang CM, Pradier JM, et al. Sesquiterpene synthase from the botrydial biosynthetic gene cluster of the phytopathogen *Botrytis cinerea*[J]. ACS Chemical Biology, 2008, 3(12): 791-801
- [10] Yamada Y, Cane DE, Ikeda H. Diversity and analysis of bacterial terpene synthases[J]. Methods in Enzymology, 2012, 515: 123-162
- [11] Berthelot M, André G. Sur l'odeur proper de la terre[J]. Comptes Rendus de l'Académie des Sciences, 1891, 112: 598-599
- [12] Gerber NN. A volatile metabolite of actinomycetes, 2-methylisoborneol[J]. The Journal of Antibiotics, 1969, 22(10): 508-509
- [13] Bentley R, Meganathan R. Geosmin and methylisoborneol biosynthesis in streptomycetes: Evidence for an isoprenoid pathway and its absence in non-differentiating isolates[J]. FEBS Letters, 1981, 125(2): 220-222
- [14] Dairi T, Hamano Y, Kuzuyama T, et al. Eubacterial diterpene cyclase genes essential for production of the isoprenoid antibiotic terpentecin[J]. Journal of Bacteriology, 2001, 183(20): 6085-6094
- [15] Jáchymová J, Votruba J, Víden I, et al. Identification of *Streptomyces* odor spectrum[J]. Folia Microbiologica, 2002, 47(1): 37-41
- [16] Komatsu M, Tsuda M, Omura S, et al. Identification and functional analysis of genes controlling biosynthesis of 2-methylisoborneol[J]. Proceedings of the National Academy of Sciences of the United States of America, 2008, 105(21): 7422-7427
- [17] Wang CM, Cane DE. Biochemistry and molecular genetics of the biosynthesis of the earthy odorant methylisoborneol in *Streptomyces coelicolor*[J]. Journal of the American Chemical Society, 2008, 130(28): 8908-8909
- [18] Lin X, Hopson R, Cane DE. Genome mining in *Streptomyces coelicolor*: molecular cloning and characterization of a new sesquiterpene synthase[J]. Journal of the American Chemical Society, 2006, 128(18): 6022-6023
- [19] Kim SY, Zhao P, Igarashi M, et al. Cloning and heterologous expression of the cyclooctatin biosynthetic gene cluster afford a diterpene cyclase and two p450 hydroxylases[J]. Chemistry & Biology, 2009, 16(7): 736-743
- [20] Finn RD, Bateman A, Clements J, et al. Pfam: the protein families database[J]. Nucleic Acids Research, 2014, 42(D1): D222-D230
- [21] Eddy SR. Hidden Markov models[J]. Current Opinion in Structural Biology, 1996, 6(3): 361-365
- [22] Reis MS, Takita MA, Palmieri DA, et al. Bioinformatics for the Citrus EST Project (CitEST)[J]. Genetics and Molecular Biology, 2007, 30(3): 1024-1029
- [23] Li G, Köllner TG, Yin Y, et al. Nonseed plant *Selaginella moellendorffi* has both seed plant and microbial types of terpene synthases[J]. Proceedings of the National Academy of Sciences

of the United States of America, 2012, 109(36): 14711-14715

- [24] Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability[J]. Molecular Biology and Evolution, 2013, 30(4): 772-780
- [25] Tamura K, Stecher G, Peterson D, et al. MEGA 6: molecular evolutionary genetics analysis version 6.0[J]. Molecular Biology and Evolution, 2013, 30(12): 2725-2729
- [26] Bailey TL, Boden M, Buske FA, et al. MEME SUITE: tools for motif discovery and searching[J]. Nucleic Acids Research, 2009, 37(Web Server issue): W202-W208
- [27] Yachdav G, Kloppmann E, Kajan L, et al. PredictProtein-an

open resource for online prediction of protein structural and functional features[J]. Nucleic Acids Research, 2014, 42(Web Server issue): W337-W343

- [28] Felicetti B, Cane DE. Aristolochene synthase: mechanistic analysis of active site residues by site-directed mutagenesis[J]. Journal of the American Chemical Society, 2004, 126(23): 7212-7221
- [29] Smanski MJ, Yu Z, Casper J, et al. Dedicated ent-kaurene and ent-atiserene synthases for platensimycin and platencin biosynthesis[J]. Proceedings of the National Academy of Sciences of the United States of America, 2011, 108(33): 13498-13503

征稿简则

1 刊物简介与栏目设置

《微生物学通报》是由中国科学院微生物研究所和中国微生物学会主办的,以微生物学应用基础研究及技术创新与应 用为主的综合性学术期刊。刊登内容包括:工业微生物学、海洋微生物学、环境微生物学、基础微生物学、农业微生物学、 食品微生物学、兽医微生物学、药物微生物学、医学微生物学、病毒学、酶工程、发酵工程、代谢工程等领域的最新研究 成果,产业化新技术和新进展,以及微生物学教学研究和改革等。设置的栏目有:研究报告、专论与综述、生物实验室、 高校教改纵横、名课讲堂、教学与科研成果展示、显微世界、专题专栏、专家论坛、书讯、会讯等。

2 投稿方式

投稿时请登陆我刊主页 http://journals.im.ac.cn/wswxtbcn,点击作者投稿区,第一次投稿请先注册,获得用户名和密码,然后依照提示提交稿件,详见主页"投稿须知"。

作者必须在网站投.doc 格式的电子稿,凡不符合(投稿须知)要求的文稿,本部恕不受理。

3 写作要求

来稿要求论点明确,数据可靠,简明通顺,重点突出。

3.1 图表

文中的图表须清晰简明, 文字叙述应避免与图表重复。所有小图的宽度应小于 8 cm (占半栏), 大图的宽度应小于 17 cm (通栏)。

3.2 参考文献及脚注

参考文献按文内引用的先后顺序排序编码,未公开发表的资料请勿引用。我刊的参考文献需要注明著者(文献作者不超 过3人时全部列出,多于3人时列出前3人,后加"等"或"et al.",作者姓前、名后,名字之间用逗号隔开)、文献名、刊名、

年卷期及页码。国外期刊名必须写完整,不用缩写,不用斜体。参考文献数量不限。

参考文献格式举例:

- Marcella C, Claudia E, Pier GR, et al. Oxidation of cystine to cysteic acid in proteins by peroyacids as monitored by immobilized pH gradients[J]. Electrophoresis, 1991, 12(5): 376-377
- Wang BJ, Liu SJ. Perspectives on the cultivability of environmental microorganisms[J]. Microbiology China, 2013, 40(1): 6-17 (in Chinese)

王保军, 刘双江. 环境微生物培养新技术的研究进展[J]. 微生物学通报, 2013, 40(1): 6-17

- [3] Shen T, Wang JY. Biochemistry[M]. Beijing: Higher Education Press, 1990: 87 (in Chinese) 沈同, 王镜岩. 生物化学[M]. 北京:高等教育出版社, 1990: 87
- [4] Liu X. Diversity and temporal-spatial variability of sediment bacterial communities in Jiaozhou Bay[D]. Qingdao: Doctoral Dissertation of Institute of Oceanology, Chinese Academy of Sciences, 2010 (in Chinese) 刘欣. 胶州湾沉积物细菌多样性及菌群时空分布规律[D]. 青岛: 中国科学院海洋研究所博士学位论文, 2010

脚注(正文首页下方): 基金项目: 基金项目(No.) *通讯作者: Tel: ; Fax: ; E-mail: 收稿日期: 2015-00-00; 接受日期: 2015-00-00; 优先数字出版日期(www.enki.net): 2015-00-00

(下转 p.2032)

http://journals.im.ac.cn/wswxtbcn