

微生物学报 *Acta Microbiologica Sinica*
50(10):1293–1304; 4 October 2010
ISSN 0001–6209; CN 11–1995/Q
<http://journals.im.ac.cn/actamicrocn>

放线菌模块型聚酮合酶的系统发育组学分析及其在聚酮类化合物筛选中的应用

王浩, 刘宁, 黄英*

(中国科学院微生物研究所, 微生物资源前期开发国家重点实验室, 北京 100101)

摘要:【目的】通过分析模块型聚酮合酶 (polyketide synthase, PKS) 的系统进化关系, 阐明酮基合成酶 (ketosynthase, KS) 和酰基转移酶 (acyltransferase, AT) 序列与聚酮产物之间的关系, 为放线菌天然产物的筛选提供指导。【方法】从 PKSDB 数据库的 20 个模块型 PKS 基因簇中调取所有 KS (190 个) 和 AT (195 个) 氨基酸序列, 利用 MEGA 4.0 软件分别构建 KS、AT、KS + AT 3 种序列模式的系统发育树, 并计算 KS 序列的簇内和簇间平均进化距离。设计了一对 KS 结构域的引物, 通过 PCR 方法对 20 株活性放线菌分离菌株进行了筛选, 测定了阳性菌株的 KS 序列, 和已知的相关 KS 序列构建系统发育树, 并对阳性菌株进行了发酵培养和代谢产物分析。【结果】放线菌来源的同一 PKS 的 KS 序列倾向于聚成一个进化枝, 且按照其产物结构聚类; 同一 PKS 的 KS 簇内平均进化距离小于 0.300, 不同 PKS 的 KS 簇间平均进化距离一般大于 0.300。AT 系统发育树按照其底物特异性聚成两个大的分枝; 同一 PKS 的部分 AT 分别处于两个分枝, 其余 AT 散在分布。KS + AT 系统发育树则综合了 KS 树和 AT 树的拓扑结构特点。获得 13 株 KS 阳性分离菌株, 它们的多数 KS 序列按照菌株分别聚类, 其中 4 株菌的大部分 KS 各自聚成独特的簇, 5 株菌的大部分 KS 分别处在已知 PKS 进化枝内。从 3 株阳性菌中分离到预期的聚酮类产物。【结论】放线菌中 KS 的进化方式以垂直进化为主, 而 AT 则以水平进化为主; KS 序列与产物结构相关, 且 KS 簇间平均进化距离可作为不同 PKS 的判定标准; 相对于 AT 和 KS + AT, KS 系统发育组学分析更适用于指导放线菌聚酮类产物的筛选。

关键词: 放线菌; 进化; 聚酮合酶; 酮基合成酶; 酰基转移酶; 系统发育组学; 产物筛选

中图分类号: Q939 **文献标识码:** A **文章编号:** 0001-6209 (2010) 10-1293-12

放线菌是多种天然生物活性物质的主要来源, 至今报道的数万种微生物来源的生物活性物质中, 约 50% 由放线菌产生^[1]。尽管如此, 已知的放线菌活性物质的数量仍然只是冰山一角^[2–3]。

放线菌的大部分活性次级代谢产物由聚酮合酶 (polyketide synthase, PKS) 途径和非核糖体多肽合成酶 (non-ribosomal peptide synthase, NRPS) 途径合成。PKS 基因广泛存在于细菌、真菌以及植物基因组之中, 其产物聚酮化合物 (polyketide) 是一大类具

有广谱生物活性的化合物, 如抗真菌、抗细菌、抗肿瘤活性, 免疫抑制剂活性, 酶抑制剂活性等。PKS 根据酶结构和产物类型主要分为三类: I 型 (模块型) PKS、II 型 (重复型) PKS 和 III 型 (查尔酮型) PKS。然而, 这种分类仍并不完善, 越来越多新类型的 PKS 被发现, 如重复型 I 型 PKS、*cis*–/*trans*-AT 型 I 型 PKS^[4–5]。放线菌中发现的主要是 *cis*-AT 型 I 型和 II 型 PKS^[6–8], 其中又以模块型 I 型 PKS 研究最为透彻。

基金项目: 国家自然科学基金 (30770002); 国家“863 计划” (2007AA09Z420)

* 通信作者。Tel: +86-10-64807311; E-mail: huangy@im.ac.cn

作者简介: 王浩 (1985–), 男, 湖北孝感人, 硕士研究生, 主要从事放线菌次级代谢基因及产物筛选。E-mail: wanghao0310@126.com

收稿日期: 2010-04-11; 修回日期: 2010-05-06

模块型 PKS 是由多个结构域以模块形式组成的复杂多肽链,主要有酮基合成酶(ketosynthase, KS),酰基转移酶(acyltransferase, AT)、酰基载体蛋白(acyl-carrier-protein, ACP)、脱氢酶(dehydratase, DH)、烯酰还原酶(enoyl reductase, ER)和酮基还原酶(ketoreductase, KR)等结构域,产物主要为大环内酯类聚酮化合物。其中 KS 和 AT 是最为保守的两个相邻的结构域,近年来较多的研究围绕它们的进化关系以及它们与相应 PKS 产物之间的关系展开。PKS 由脂肪酸合成酶(fatty acid synthase, FAS)进化而来,基因或模块复制、基因横向转移(horizontal gene transfer, HGT)以及同源重组是其进化的主要方式^[6,9-10],模块型 PKS 可能是由细菌 FAS 和原始的重叠型 PKS 的基因复制进化而来^[6]。Yadav 等人开发了预测模块型 PKS 各结构域的软件与数据库^[11-12],极大方便了对 PKS 的研究。Nguyen 等人发现 HGT 和重组是 *trans*-AT 模块型 PKS 进化的主要方式,并根据 *trans*-AT 模块型 PKS 的 KS 结构域底物特异性的特点成功预测了 thailandamide A 的结构^[7]。Haydock 等人在模块型 PKS 的 AT 结构域氨基酸序列中发现了与底物特异性相关的基序(motif)^[13-14],Ridley 等人随后补充了其他几种底物特异性的 AT 结构域^[15]。

模块型 PKS 由于酶自身结构的复杂性及其进化方式的多样性,不同拷贝的 KS 或 AT 基因不能简单视为同一基因,单个 KS 或 AT 基因往往不能正确显示整个基因簇的进化关系,因此需要运用系统发育组学分析的方法来研究 PKS 的进化^[6,16-18]。系统发育组学分析由于使用了更多的同源特征,能构建更精细、准确的系统发育树^[19]。Jenke-Kodama 等人对真细菌的 KS 和 AT 进行了系统发育组学分析,发现不同细菌类群的 I 型 PKS 主要进化方式也不同^[6]。

面对当今天然产物尤其是聚酮化合物筛选过程中的大量重复发现的问题,亟需发展新的筛选技术和筛选模型。PKS 的进化关系以及和产物结构之间的联系可用于预测相应产物的结构,本文期望通过分析完整的模块型 PKS 的系统进化关系,结合便捷的基因筛选技术,构建简便易行的筛选模型,为放线菌聚酮类产物的筛选提供指导。

1 材料和方法

1.1 材料

1.1.1 菌株与培养基:随机选取 20 株由本实验室

分离的活性放线菌,活化和发酵培养基均采用 GYM 培养基[配方如下:Glucose 4.0 g;Malt extract 10.0 g;Yeast extracts (Difco) 4.0 g;CaCO₃ 2.0 g;Agar 15.0 g;Distilled water to 1000 mL。调整 pH 至 7.0-7.3]。

1.1.2 主要试剂和仪器:PCR 扩增相关试剂、DNA 回收试剂盒和测序服务均由北京博迈德科技发展有限公司提供,克隆载体 pMD-18T 和连接试剂盒均购自宝生物工程有限公司,引物由上海生物工程技术服务有限公司合成,其余试剂均为国产或进口分析纯。PCR 仪和凝胶成像仪分别为 Bio-rad PTC-200 和 Gel-Doc2000;高压液相色谱仪为 Shimadzu SPD-M20A,色谱柱为 YMC ODS-A,4.6 × 150-mm。

1.2 PKS 系统发育组学分析

1.2.1 KS 和 AT 氨基酸序列获取:从 PKSDB 数据库(<http://linux1.nii.res.in/~pkfdb/DBASE/pageALL4.html>)中下载了 20 个 *cis*-AT 模块型 PKS 完整基因簇的氨基酸序列,其中 14 个来自放线菌;并根据该数据库的注释调取其中的 KS 和 AT 序列,共获得 190 个 KS 全长氨基酸序列、195 个 AT 全长氨基酸序列(见表 1)。

1.2.2 系统发育树的构建:对所获取的所有 KS 和 AT 氨基酸序列,按照 KS、AT、KS + AT(相邻序列串联)3 种序列模式分别进行系统发育组学分析。通过软件 MEGA 4.0 中 ClustalW 进行序列比对,手动调整后再次比对,用 Neighbour-Joining 法分别构建 3 种序列模式的系统发育树,自展重复次数为 2000,其余参数为默认设置。

1.2.3 进化距离计算:将放线菌来源的同一 PKS 中除起始模块外的绝大部分 KS 归为一簇,利用 MEGA 4.0 中的 Poisson correction 模型分别计算同一 PKS 的 KS 簇内平均进化距离,和不同 PKS 的 KS 簇间平均进化距离。产物结构极为类似、KS 树上聚成混合分枝的 PKS 在计算进化距离时仅保留一个 PKS,即 amphotericin、nystatin、pimaricin 三者仅保留 pimaricin,erythromycin、megalomycin、oleandomycin 三者仅保留 oleandomycin。

1.3 分离菌株 KS 序列测定与分析

1.3.1 菌株培养与基因组 DNA 提取:20 株活性分离菌株在 GYM 平板上 28℃ 培养 5 d 后,按照文献^[20]提取基因组 DNA。

1.3.2 引物设计、PCR 扩增及克隆测序:分析表 1 中所有放线菌来源的 KS 氨基酸序列,根据其中两个保守位点 DPQQR 和 VNQDG 设计了一对兼并引

表 1 20 个 PKS 的相关信息

Table 1 Information of the 20 PKSs retrieved from PKSDB

PKS pathway	Abbreviation	No. of KS	No. of AT	Producer
amphotericin	ampho	19	19	<i>Streptomyces nodosus</i>
ascomycin *	ascom	10	10	<i>Streptomyces hygroscopicus</i>
avermectin *	averm	12	13	<i>Streptomyces avermitilis</i>
erythromycin	eryth	6	7	<i>Saccharopolyspora erythraea</i>
megalomycin	megal	6	6	<i>Micromonospora megalomicea</i>
niddamycin *	nidda	8	8	<i>Streptomyces caelestis</i>
nystatin	nysta	19	19	<i>Streptomyces noursei</i>
oleandomycin *	olean	7	7	<i>Streptomyces antibioticus</i>
pikromycin *	pikro	7	6	<i>Streptomyces venezuelae</i>
pimaricin *	pimar	12	13	<i>Streptomyces natalensis</i>
rapamycin *	rapam	14	14	<i>Streptomyces hygroscopicus</i>
rifamycin *	rifam	10	10	<i>Amycolatopsis mediterranei</i>
spinosad *	spino	11	11	<i>Saccharopolyspora spinosa</i>
tylactone *	tylac	8	8	<i>Streptomyces fradiae</i>
epothilone	epoth	9	9	<i>Sorangium cellulosum</i>
myxalamide	myxal	7	8	<i>Cystobacter fuscus</i>
myxothizol	myxot	5	6	<i>Myxococcus fulvus</i>
stigmatellin	stigm	9	10	<i>Stigmatella aurantica</i>
soraphen	sorap	8	9	<i>Sorangium cellulosum</i>
pyoluteorin	pyolu	3	2	<i>Pseudomonas fluorescens</i>

* PKS pathways from actinomycetes used to compute evolutionary distances in 1. 2. 3.

物; KSF (5'-GARGCSCTSGARAYSGAYCCVC-3') 和 KSR (5'-CCGTTSGACGCRCCGTCCTGGTTSAC-3')。用这对引物对提取的基因组 DNA 进行 PCR 扩增, PCR 体系为 50 μL, 包含 0.4 μmol KSF, 0.4 μmol KSR, 0.2 mmol dNTPs, 1 μL DNA, 5 单位 *Taq* 酶, 1 × buffer(含 Mg²⁺) 和 3 μL DMSO。PCR 程序为: 95℃ 5 min; 95℃ 45 s, 64℃ 45 s, 72℃ 50 s, 循环 28 次; 72℃ 10 min。PCR 产物经 1.5% 琼脂糖凝胶电泳检测(电压为 100 V), 回收约 570 bp 的目标片段, 试剂盒纯化后克隆测序。每株菌随机挑取 10 – 12 个克隆进行测序。

1.3.3 KS 序列分析:测得序列通过 BLAST 比对分析以验证序列功能。搜寻比对结果中序列相似性较高且报道有产物和已发表完整基因簇的 PKS, 获取 8 个 PKS (chlorothricin、lasalocid、lipomycin、meridamycin、nanchangmycin、nigericin、salinomycin、tetrinomycin) 的所有 KS 序列, 并与 10 个用于计算 KS 进化距离的放线菌 PKS 的 KS 序列及所测分离菌株序列一同构建系统发育树。测得的 KS 序列 GenBank 注册号见表 2。

1.4 菌株发酵及产物的分离纯化与鉴定

1.4.1 菌株发酵:小量发酵:将 13 株 PKS 阳性分离菌株及一株参照菌株 JCM 3373 的新鲜 GYM 斜面菌种分别接种于 100 mL 的 GYM 平板培养基, 28℃ 静置培养 7 d。大量发酵:从菌株 E3 的新鲜 GYM

平板挑取小块接种入装有 100 mL GYM 琼脂块的 500mL 三角瓶中, 共接 250 瓶合计 25 L 发酵量, 28℃ 静置培养 7 d。

1.4.2 产物提取及分离纯化:小量提取:将发酵完毕的琼脂平板捣碎后, 置于 500 mL 三角瓶中并加入 100 mL 乙醇震荡抽提过夜; 滤取乙醇部分并将其减压浓缩至干, 用 2 mL 甲醇溶解浓缩物后取甲醇部分待用。大量提取: 25 L 菌株 E3 的固体发酵产物用等体积乙醇反复震荡抽提过夜 3 次, 滤取乙醇部分并将其减压浓缩后用乙酸乙酯萃取 3 次; 收集、蒸干有机相后得粗提物; 粗提物经反复硅胶 (100-200 目) 层析及 LH-20 凝胶层析后得 1 个主要纯化合物。

1.4.3 产物结构鉴定:小量提取产物使用 HPLC-UV/VIS-ESI-MS/MS 分析。HPLC 采用 Shimadzu SPD-M20A (PDA 检测器, 190-800 nm), 梯度程序采用 15 min 内甲醇浓度从 50% 升至 100%, 色谱条件是: YMC ODS-A, 4.6 mm × 150 mm 色谱柱; 流速 1.0 mL/min; 柱温为室温; 进样量 20 μL。ESI-MS/MS 分析采用 Thermo-Finnigan LCQ DECA XP 质谱仪, 正/负离子模式检测。大量发酵产物的 ESI-MS 分析采用 Thermo-Finnigan LCQ DECA XP 质谱仪, 正/负离子模式检测; NMR 测定采用 Bruker DRX-600/400 型核磁共振仪, 四甲基硅烷 (TMS) 为内标。

表 2 本实验鉴定的 KS 基因序列 GenBank 注册号

Table 2 GenBank accession numbers of ketosynthase genes identified in this study

SeqID	Accession No.	SeqID	Accession No.	SeqID	Accession No.
FXJ1. 264-1	GU363593	FXJ1. 066-7	GU363624	FXJ1. 172-6	GU363654
FXJ1. 264-2	GU363594	FXJ6. 141-1	GU363625	FXJ1. 172-7	GU363655
FXJ1. 264-3	GU363595	FXJ6. 141-2	GU363626	FXJ1. 172-8	GU363656
FXJ1. 264-4	GU363596	FXJ6. 141-3	GU363627	FXJ7. 346-1	GU363657
FXJ1. 264-5	GU363597	FXJ6. 141-4	GU363628	FXJ7. 346-2	GU363658
FXJ1. 264-6	GU363598	FXJ6. 141-5	GU363629	FXJ7. 346-3	GU363659
FXJ1. 264-7	GU363599	FXJ6. 141-6	GU363630	FXJ7. 346-4	GU363660
FXJ1. 264-8	GU363600	FXJ6. 141-7	GU363631	FXJ7. 346-5	GU363661
FXJ1. 264-9	GU363601	FXJ6. 141-8	GU363632	FXJ1. 068-1	GU363664
E3-1	GU363603	FXJ6. 141-9	GU363633	FXJ1. 068-2	GU363665
E3-2	GU363604	FXJ6. 050-1	GU363634	FXJ1. 068-3	GU363666
E3-3	GU363605	FXJ6. 050-2	GU363635	FXJ1. 068-4	GU363667
E3-4	GU363606	FXJ6. 050-3	GU363636	FXJ1. 068-5	GU363668
E3-5	GU363607	FXJ6. 050-4	GU363637	FXJ1. 262-1	GU363669
E3-6	GU363608	FXJ6. 050-5	GU363638	FXJ1. 262-2	GU363670
E3-7	GU363609	FXJ6. 050-6	GU363639	FXJ1. 262-3	GU363671
E3-8	GU363610	FXJ6. 050-7	GU363640	FXJ1. 262-4	GU363672
E3-9	GU363611	FXJ6. 050-8	GU363641	FXJ1. 262-5	GU363673
E3-10	GU363612	FXJ6. 050-9	GU363642	FXJ1. 262-6	GU363674
A33-1	GU363613	FXJ6. 050-10	GU363643	FXJ1. 262-7	GU363675
A33-2	GU363614	FXJ6. 196-1	GU363644	FXJ1. 121-1	GU363676
A33-3	GU363615	FXJ6. 196-2	GU363645	FXJ1. 121-2	GU363677
A33-4	GU363616	FXJ6. 196-3	GU363646	FXJ1. 121-3	GU363678
A33-5	GU363617	FXJ6. 196-4	GU363647	FXJ1. 121-4	GU363679
FXJ1. 066-1	GU363618	FXJ6. 196-5	GU363648	FXJ7. 329-1	GU363684
FXJ1. 066-2	GU363619	FXJ1. 172-1	GU363649	FXJ7. 329-2	GU363683
FXJ1. 066-3	GU363620	FXJ1. 172-2	GU363650	FXJ7. 329-3	GU363682
FXJ1. 066-4	GU363621	FXJ1. 172-3	GU363651	FXJ7. 329-4	GU363681
FXJ1. 066-5	GU363622	FXJ1. 172-4	GU363652	FXJ7. 329-5	GU363680
FXJ1. 066-6	GU363623	FXJ1. 172-5	GU363653		

2 结果

2.1 PKS 系统发育组学分析

2.1.1 KS、AT 及 KS + AT 系统发育组学分析:KS 序列系统发育树见图 1。9 个 PKS 起始模块 (initiative)^[21] 的 KS 聚成最外围进化枝,6 个来源于非放线菌 PKS 的 KS 均位于发育树外围(图 1),而 14 个放线菌 PKS 的 KS,除部分起始模块外,均位于发育树内层。来自同一 PKS 的绝大部分 KS 都处在同一进化枝内,尤其是放线菌的 KS。Ascomycin 和 rapamycin 的进化枝稳定地聚为一群,形成较大分枝;amphotericin、nystatin 和 pimaricin 这 3 个 PKS 的 KS 聚在一起形成一个大的混合分枝。

AT 序列系统发育树分成了两个大的进化枝

Clade A 和 Clade B(图 2),非放线菌来源的 AT 均分别聚在两个大进化枝的外围,放线菌来源的 AT 则分别聚在两个大进化枝的内部。来源于放线菌的同一 PKS 的部分 AT 聚成两簇分别处于两个大进化枝内,其余 AT 则弥散分布在两大分枝内,如 avermectin 和 rapamycin。Amphotericin、nystatin 和 pimaricin 这 3 个 PKS 的 AT 同样聚成混合分枝。

190 个 KS + AT 序列构建的系统树明显兼具了 KS 和 AT 树的拓扑结构特点(图 3):非放线菌来源的序列聚在了外围,放线菌来源的序列则聚在内层,起始模块序列在最外层,这一点与 KS 树一致;内层的序列形成了两个大的进化枝,这种拓扑结构又与 AT 树一致。Ascomycin 和 rapamycin 的进化枝,以及 amphotericin、nystatin 和 pimaricin 的混合分枝在 Clade A 和 Clade B 中均有分布。

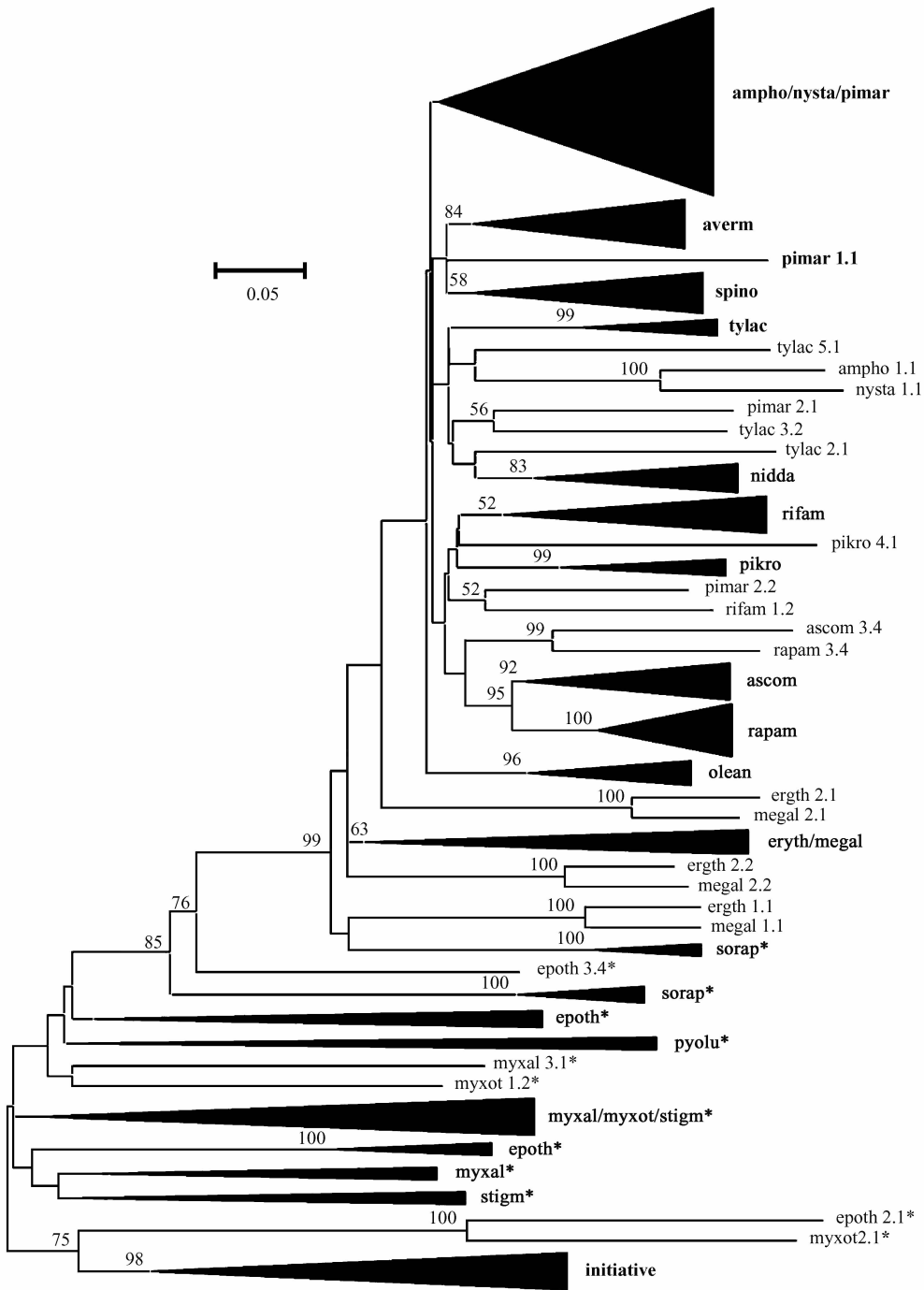


图 1 根据 20 个不同来源 PKS 的 190 个 KS 氨基酸序列构建的 Neighbour-joining 树

Fig. 1 Neighbour-joining tree of 190 KSs from 20 PKSs published. Clades containing KSs from the same PKS pathways are marked with black triangles. KSs from non-actinomycetes are marked with symbol *. The first number of each KS represents the number of module, and the second number represents the order of KS in the module. Significant bootstrap values (>50%) are indicated at the nodes. The scale bar represents 0.05 mutational events per site.

2.1.2 进化距离:10 个放线菌来源的 PKS 分别计算 KS 簇内平均进化距离和簇间平均进化距离。KS 簇内平均进化距离除 pimaricin 为 0.302 外,其余 9 个 PKS 的簇内平均进化距离均小于 0.300,最小值为 0.140(rapamycin),总平均值 0.225。除 avermectin 与 oleandomycin、avermectin 与 spinosad、ascomycin 与

rapamycin 这 3 组 KS 簇间平均进化距离分别为 0.292、0.272、0.265 外,其他 42 组 KS 簇间平均进化距离均大于 0.300,最大值为 0.367,总平均值为 0.325。

2.2 分离菌株的 KS 序列分析

用 KSF/KSR 引物对 20 株分离菌株进行 PCR 检测,结果 13 株为阳性。经克隆测序和 BLAST 比对鉴

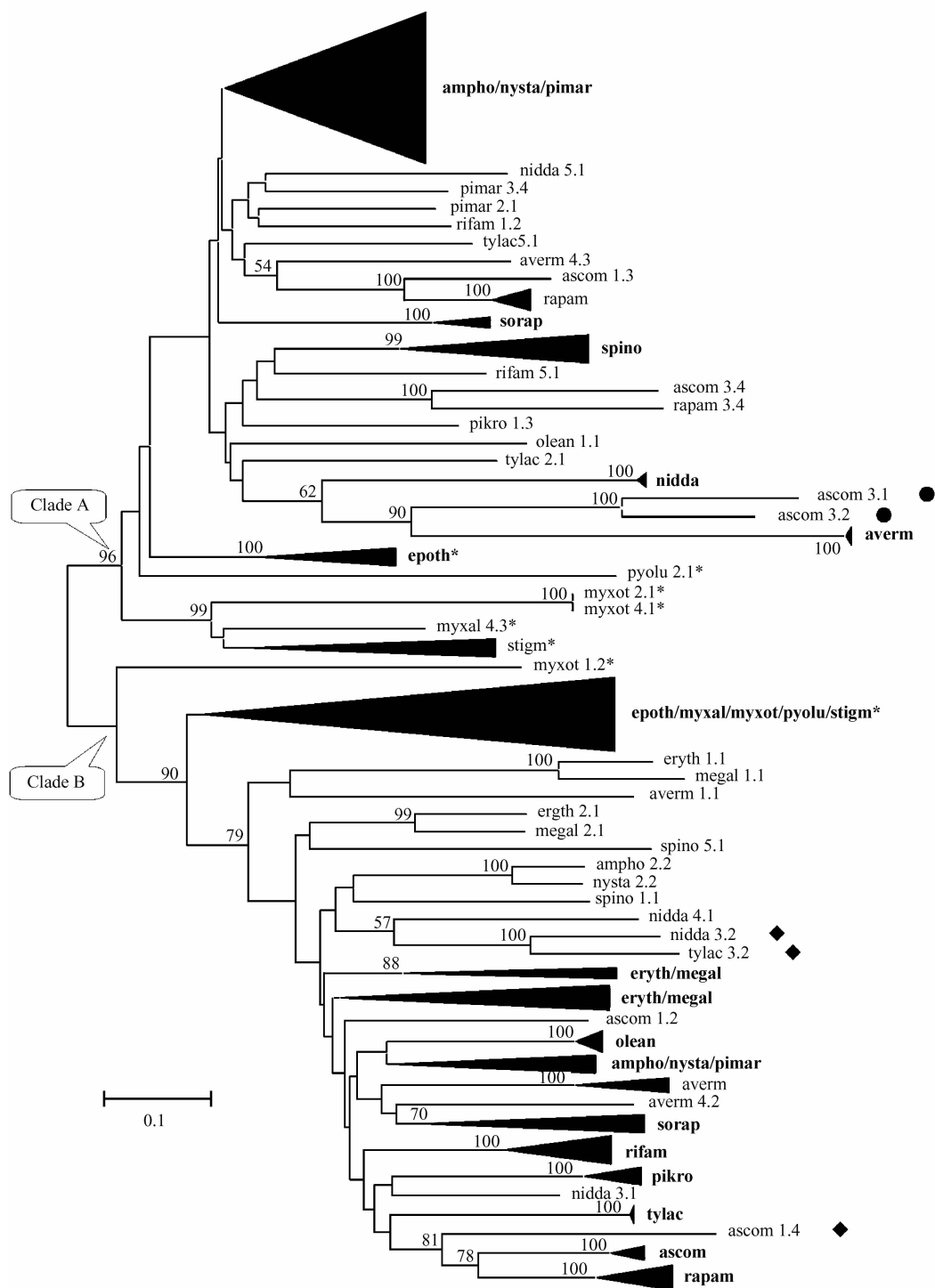


图 2 根据 20 个不同来源的 PKS 的 195 个 AT 氨基酸序列构建的 Neighbour-joining 树

Fig. 2 Neighbour-joining tree of 195 ATs from 20 PKSs published. Clades containing ATs from the same PKS pathways are marked with black triangles. ATs from non-actinomycetes are marked with symbol *. A couple of methoxymalonyl-specific ATs and ethylmalonyl-specific ATs are marked with symbols ● and ◆, respectively. The first number of each AT represents the number of module, and the second number represents the order of AT in the module. Significant bootstrap values (>50%) are indicated at the nodes. The scale bar represents 0.1 mutational events per site.

定后,每株阳性菌分别获得了 4 – 10 个不同 KS 拷贝的序列,共获得 89 个 KS 拷贝的序列。用这些序列和 18 个已知 PKS 的 KS 序列共 249 条构建的系统发育

树见图 4。分离菌株的大部分 KS 按照菌株分别聚在一起,每个菌株也都有 KS 弥散分布。4 株分离菌株 E3、A33、FXJ6.050 和 FXJ1.066 的大部分 KS 各自聚

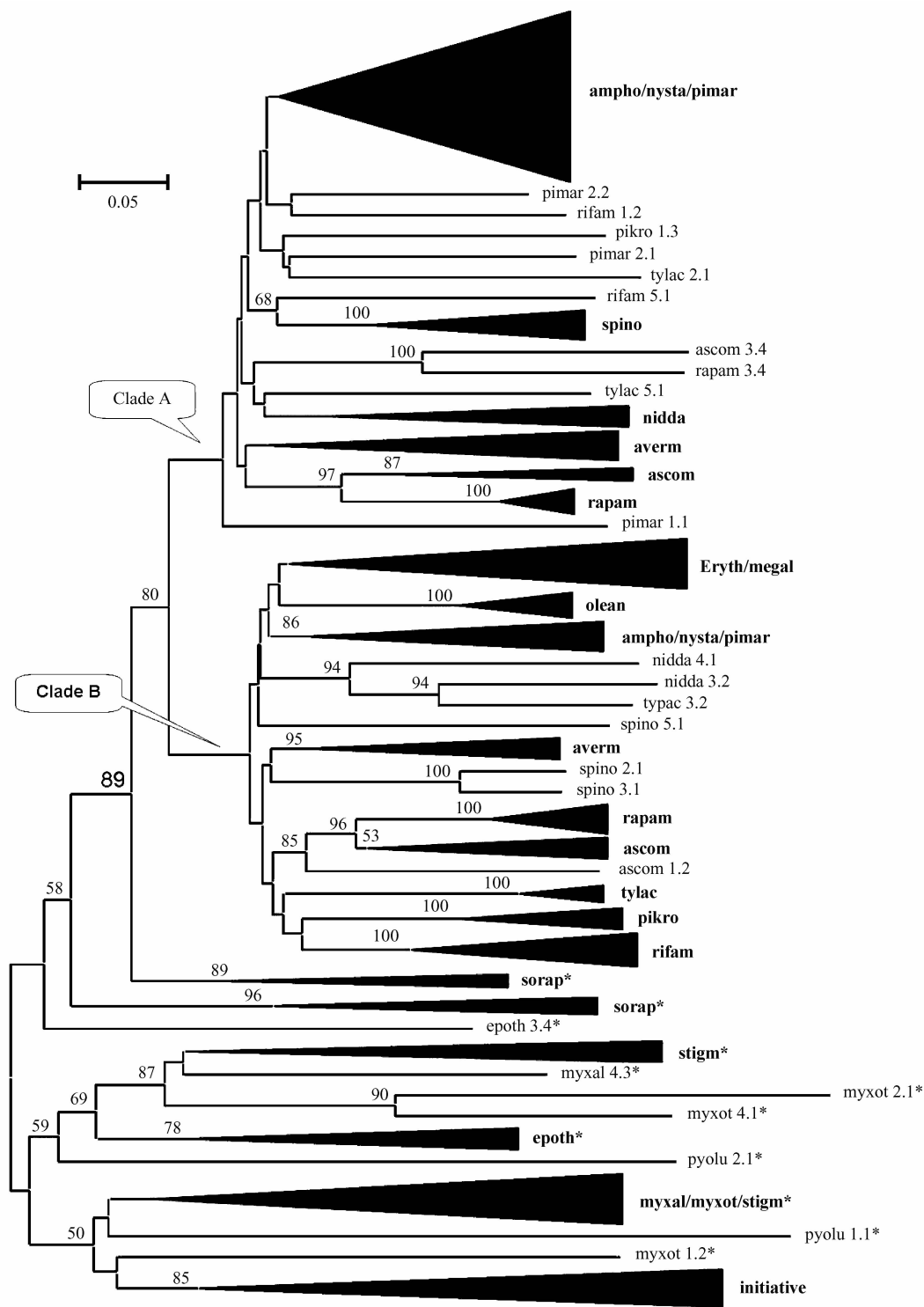


图 3 根据 20 个不同来源的 PKS 的 190 个 KS + AT 氨基酸序列构建的 Neighbour-joining 树

Fig. 3 Neighbour-joining tree of 190 KS + ATs from 20 PKSs published. Clades containing domains from the same PKS pathways are marked with black triangles. Domains from non-actinomycetes are marked with symbol *. The first number of domains represents the number of module, and the second number represents the order of domains in the module. Significant bootstrap values (>50%) are indicated at the nodes. The scale bar represents 0.05 mutational events per site.

簇,不与任何已知 PKS 有明显聚类关系;5 株分离菌株的多个 KS 处在已知 PKS 分枝内,如 FXJ1. 172 的 8 个 KS 中有 5 个处在 lasalocid 分枝,再如 FXJ6. 196 与 tetrinomycin, FXJ1. 121、FXJ1. 264 与 nigericin,

FXJ6. 141 与 salinomycin; FXJ7. 346、FXJ7. 329 和 FXJ6. 141 的大部分 KS 聚成混合分枝;FXJ1. 068 和 FXJ1. 262 的 KS 则弥散分布在树上。

分别计算图 4 中分离菌株所在的分枝簇内和簇间

平均进化距离,同样发现 KS 的簇内平均进化距离均小于 0.300,而簇间平均进化距离绝大多数大于 0.300。

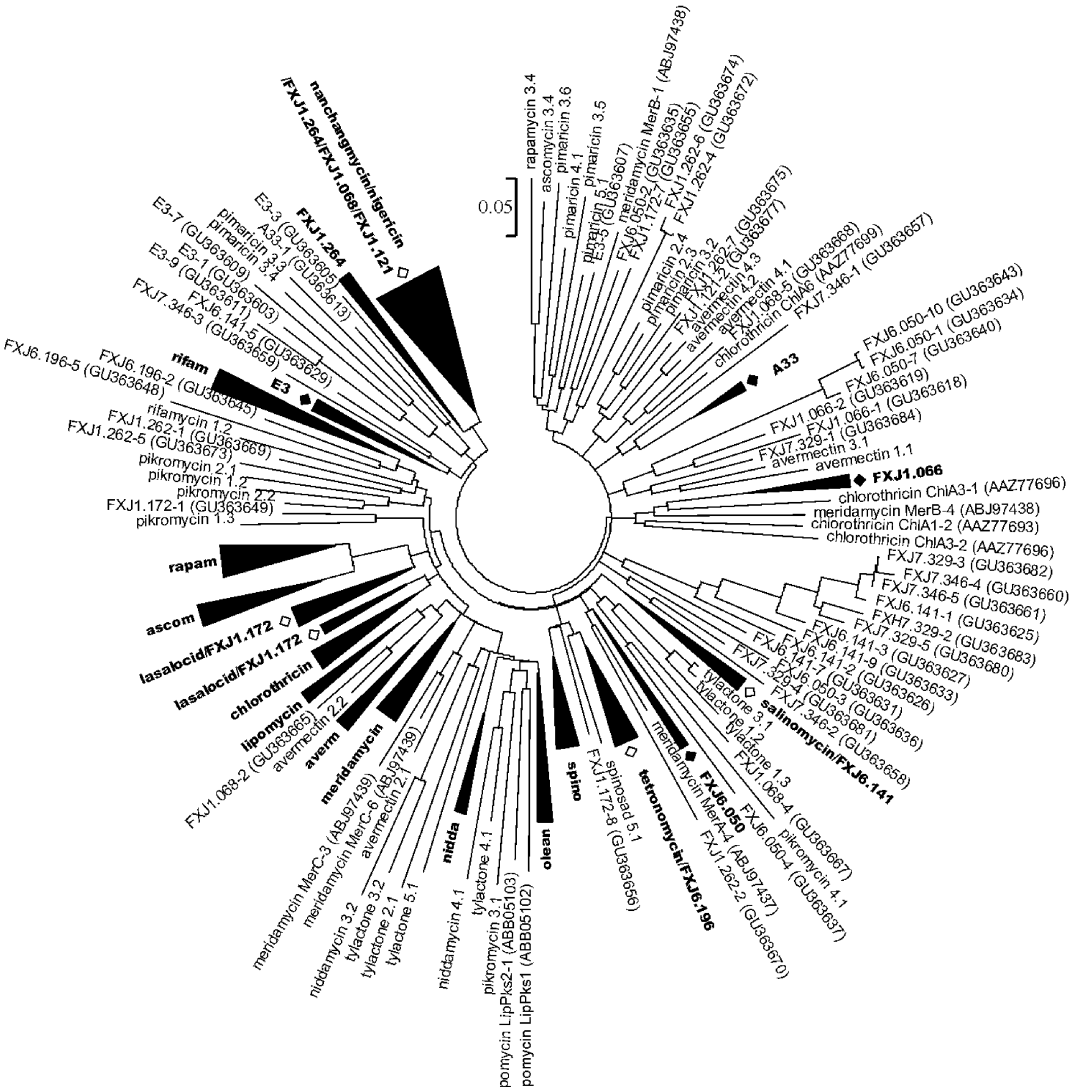


图 4 根据 18 个已知 PKS 和 13 株分离菌株的 KS 构建的 Neighbour-joining 树

Fig. 4 Neighbour-joining tree of KSs from known PKSs and isolates. Clades containing KSs from the same PKS pathways are marked with black triangles. KSs from isolates fell into the clades of known PKSs are marked with symbol ◇. KSs from isolates clustered with none of known PKSs are marked with symbol ◆. The first number of KSs represents the number of module, and the second number represents the order of KSs in the module. The scale bar represents 0.05 mutational events per site.

2.3 聚酮类代谢产物分析

13株阳性分离菌株的小量发酵产物经 HPLC-UV/VIS-ESI-MS/MS 分析后获得 2 株产生聚醚类聚酮化合物的菌株(FXJ1.172 和 FXJ1.068)。菌株 JCM 3373 是已发表的产生 lasalocid A 的菌株^[22],从表 3 可以看出,FXJ1.172 产物的紫外吸收、质谱数据及质谱二级断裂数据均与菌 JCM 3373 产生的 lasalocid A 保持一致,由此推定其产物应为 lasalocid A;FXJ1.068 能够产生 2 个分子量皆为 866 D 的化合物,其紫外吸收、质谱数据及二级质谱断裂数据均高度相似,且与已知化合物南昌霉素(nanchangmycin)保持一致^[23],由此推测二者应为南昌霉素及其立体异构体(19-epimer)。菌株 E3 小量发酵产物中的目标化合物无紫外吸收,无法经由 HPLC-UV 检测,然而经 TLC-ESI-MS 分析后可初步推定其分子量为 568 D。进一步放

大发酵并反复纯化后得到的相应化合物 Na⁺盐的¹³C-NMR 波谱(CDCl₃,100MHz)数据如下:10.4(q, C-29),10.8(q, C-27),11.0(q, C-31),12.6(q, C-30),14.2(q, C-25),15.5(q, C-33),17.4(q, C-32),17.5(q, C-28),20.6(t, C-24),20.9(q, C-26),24.9(t, C-4),27.1(t, C-5),28.4(t, C-15),31.8(d, C-22),32.2(d, C-6),32.6(d, C-8),35.0(t, C-14),35.9(d, C-10),36.0(d, C-12),39.9(t, C-23),41.0(d, C-2),41.6(d, C-18),68.3(d, C-13),75.6(d, C-7),76.7(d, C-3),81.8(d, C-19),83.4(d, C-11),86.3(d, C-9),131.5(d, C-16),133.5(s, C-20),134.3(d, C-17),135.3(d, C-21),182.9(s, C-1)。其分子量及¹³C-NMR 波谱数据与 griseochelin^[24]基本一致,经 2D-NMR 波谱法进一步确认为 griseochelin。

表 3 菌株产物的化学分析数据及结果

Table 3 Chemical data and products identified from the positive isolates			
Strain No.	UV	Pseudomolecular ions of compounds (m/z)	Compound identified
JCM 3373 *	304	MS ₁ : [M + H] ⁺ = 591, [M + NH ₄] ⁺ = 608, [M + Na] ⁺ = 613, [M + K] ⁺ = 629, [M-H] ⁻ = 589; MS ₂ (613.5): 359.2, 377.3 , 395.0, 569.4 , 577.4 , 595.3 , 613.3;	lasalocid A
FXJ1.172	304	MS ₁ : [M + H] ⁺ = 591, [M + NH ₄] ⁺ = 608, [M + Na] ⁺ = 613, [M + K] ⁺ = 629, [M-H] ⁻ = 589; MS ₂ (613.6): 359.2, 377.4 , 394.9, 569.4 , 577.4 , 595.3 , 613.2;	lasalocid A
FXJ1.068	234	MS ₁ : [M + NH ₄] ⁺ = 884, [M + Na] ⁺ = 889; MS ₂ (889.7): 447.3 , 491.3 , 591.4 , 631.3, 725.4 , 743.4 , 761.4 , 853.4 , 871.4 , 889.5	nanchangmycin
	234	MS ₁ : [M + NH ₄] ⁺ = 884, [M + Na] ⁺ = 889; MS ₂ (889.7): 447.0 , 491.1 , 591.3 , 631.3, 725.3 , 743.4 , 761.4 , 853.4 , 871.4 , 889.4	
E3	none	[M + H] ⁺ = 569, [M-H] ⁻ = 567	griseochelin

* Strain JCM 3373 is reported to produce lasalocid A.

3 讨论

随着测序技术的不断发展,近 100 个放线菌全基因组序列测定已完成,加之 PKS 各个结构域预测工具的开发,为研究 PKS 的系统进化提供了极大的方便。越来越多的研究表明,不同类型、不同来源的 PKS,其进化的主要机制也不尽相同。

在本研究所得的 KS 系统发育树中,对来源于放线菌的 PKS,同一 PKS 的绝大部分 KS 聚成一个进化枝。如果几个 PKS 的进化枝聚成大的分枝,那么其产物化学结构也类似,例如 amphotericin、nystatin 和 pimaricin,这 3 个聚酮化合物的主链结构尤其是前两者基本一致(图 5),而且 pimaricin 的多个模块由 amphotericin 或 nystatin 整体基因横向转移而来^[10],这 3 个 PKS 的 KS 也完全聚在一起形成一个较大的混合分枝。再如 ascomycin 和

rapamycin,这两个聚酮的化学结构相似,在发育树上相应的 KS 也分别聚在一起;erythromycin、megalomycin 和 oleandomycin 在树上位置也相近。与放线菌的情况不同,非放线菌来源的 myxalamide、myxothiazol 和 stigmatellin 三者尽管在化学结构上没有多大一致性,但这 3 个 PKS 的 KS 却聚在一起形成了外围混合进化枝。由此推断放线菌中的 KS 主要由垂直进化而来,其祖先基因在进化过程中经过了多次基因复制事件形成多个 KS,因此,同一 PKS 的 KS 序列能保证较高的序列相似性。但我们也注意到, pimar2.1、nysta1.1 和 ampho1.1 聚在 niddamycin 和 tylactone 的进化枝内, pimar1.1 聚在 avermectin 和 spinosad 进化枝内,以及类似的少数 KS 聚类在其它 PKS 进化枝中的现象,说明在放线菌 KS 的进化过程中可能也普遍存在 HGT 事件。

放线菌中 AT 结构域有着和 KS 结构域明显不同的进化方式。AT 树的两大进化枝分别代表两大

类型的 AT, Clade A 为 malonyl-CoA 底物特异性的 AT, Clade B 为 methylmalonyl-CoA 底物特异性的 AT^[6] (图 2)。在这两大分枝内还散在分布有少数几个其他底物特异性的 AT, 如 ethylmalonyl-CoA 和 methoxymalonyl-CoA^[15]。AT 结构域在树上散在分布, 且多个不同 PKS 的 AT 结构域具有较高的序列相似性, 这表明 AT 在放线菌内的基因交流现象十

分普遍, HGT 可能是 AT 结构域进化的主要方式。但是, 仍然可以看到同一 PKS 的部分 AT 序列相似性较高, 在发育树上聚在一起, 说明基因复制可能也是 AT 结构域进化的一种方式。此外, 不管是 KS 树还是 AT 树, 非放线菌来源的序列都聚在外围, 提示放线菌中的 KS 和 AT 与其他微生物之间的基因交流并不常见。

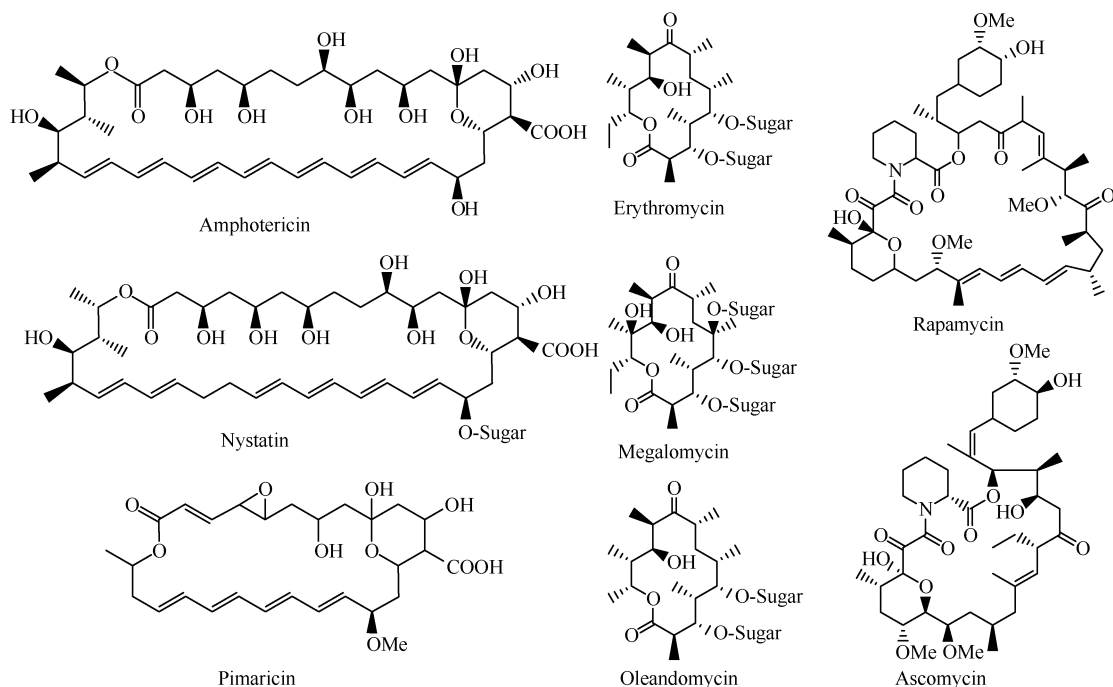


图 5 化学结构类似的几个聚酮化合物

Fig. 5 Polyketide compounds with similar chemical structures involved in this study.

KS + AT 树由于综合了 KS 和 AT 两个系统发育因素, 在拓扑结构上兼具了二者的特点, 反而不能正确反应产物信息。有些学者在研究实验菌株的 PKS 时, 会使用 PCR 片段横跨 KS 和 AT 两个结构域的引物, 如 K1F/M6R, 这样的序列在做系统发育分析时又引入了 KS 和 AT 之间连接序列的信号, 更不能准确反应产物信息。尽管 AT 结构域底物特异性的特点十分有利于产物结构的预测, 但其限制性条件是必须有 PKS 基因簇的全序列。由于 KS 的 HGT 现象也广泛存在, 而且同一放线菌基因组往往有多个 PKS 途径^[25-26], 一株分离菌株的一两个 KS 序列不足以根据 KS 以产物结构聚类的特点来预测相应产物的结构。我们对 KS 簇内和簇间平均进化距离的分析表明, 如果有足够多的已知 PKS 的 KS 序列, 通过分析菌株的多个 KS 拷贝序列, 能够大致预测实验菌株聚酮产物的主结构, 即: 若多个 KS 拷贝序列与已知序列的平均进化距离均小于 0.300, 菌株可能产生相应的已知聚酮化合物的类似物, 反之可

能产生不同的化合物。

为了验证这一推测, 我们设计了 KS 结构域的特异引物, 筛选出 13 株阳性菌株并测定了其 KS 序列。分离菌株的 KS 结构域的发育树和进化距离分析得到了同样的结果。4 株分离菌株 A33、E3、FXJ6.050 和 FXJ1.066 其 KS 不与 18 个已知 PKS 聚成簇, 且簇间平均进化距离均大于 0.300, 因此, 推测这 4 株菌的产物主结构可能不同于这 18 个已知聚酮。FXJ1.172 的 KS 落在 lasalosid 进化枝内且平均进化距离仅为 0.202, 推测 FXJ1.172 的聚酮产物与 lasalosid 结构类似; 同理推测 FXJ1.068、FXJ1.121 产物与 nanchangmycin 结构类似, FXJ6.141 产物与 salinomycin 结构类似, FXJ16.196 产物结构与 tetronomycin 结构类似。随后菌株发酵产物的分离鉴定工作证实, 菌株 E3 的主产物为 griseochelin, 结构与 18 个已知聚酮相差较远; FXJ1.172 产物为 lasalocid A, 与推测一致; FXJ1.068 尽管 5 个 KS 中仅 1 个落在 nanchangmycin 进化枝

内,但是其产物证实为 nanchangmycin。虽然这种方法在新结构的预测上有一定限制,但随着越来越多的 PKS 序列发表,这种限制会不断减弱。重要的是,这种方法能够排除某些已知化合物(有 PKS 序列信息的聚酮)的产生菌,从而提高筛选效率。

综上所述,放线菌中模块型 PKS 的系统进化不能一概而论,KS 结构域和 AT 结构域有着明显不同的进化方式。KS 基因复制可能在 KS 结构域的进化过程中起了主要作用,但 HGT 现象也广泛存在;而 AT 结构域的进化方式可能以 HGT 以为主,基因复制为辅。KS 以产物结构聚类,而且 KS 平均进化距离 0.300 可以作为不同 PKS 的初步判定标准,因此,KS 的系统发育组学分析比 AT 和 KS + AT 更适用于指导放线菌聚酮类化合物的筛选。

参考文献

- [1] Bérdy J. Bioactive microbial metabolites. *The Journal of Antibiotics*, 2005, 58 (1) : 1-26.
- [2] Watve MG, Tickoo R, Jog MM, Bhole BD. How many antibiotics are produced by the genus *Streptomyces*? *Archives of Microbiology*, 2001, 176 (5) : 386-390.
- [3] Baltz RH. Renaissance in antibacterial discovery from actinomycetes. *Current Opinion in Pharmacology*, 2008, 8 (5) : 557-563.
- [4] Shen B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current Opinion in Chemical Biology*, 2003, 7 (2) : 285-295.
- [5] Cheng YQ, Tang GL, Shen B. Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proceedings of the National Academy of Sciences*, 2003, 100 (6) : 3149.
- [6] Jenke-Kodama H, Sandmann A, Muller R, Dittmann E. Evolutionary implications of bacterial polyketide synthases. *Molecular Biology and Evolution*, 2005, 22 (10) : 2027-2039.
- [7] Nguyen T, Ishida K, Jenke-Kodama H, Dittmann E, Gurgui C, Hochmuth T, Taudien S, Platzer M, Hertweck C, Piel J. Exploiting the mosaic structure of trans-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nature Biotechnology*, 2008, 26 (2) : 225.
- [8] Wawrik B, Kutliev D, Abdivasievna UA, Kukor JJ, Zylstra GJ, Kerkhof L. Biogeography of actinomycete communities and type II polyketide synthase genes in soils collected in New Jersey and Central Asia. *Applied and Environmental Microbiology*, 2007, 73 (9) : 2982-2989.
- [9] Jenke-Kodama H, Dittmann E. Evolution of metabolic diversity: Insights from microbial polyketide synthases. *Phytochemistry*, 2009, 70 : 1858 - 1866.
- [10] Ginolhac A, Jarrin C, Robe P, Perrière G, Vogel TM, Simonet P, Nalin R. Type I polyketide synthases may have evolved through horizontal gene transfer. *Journal of Molecular Evolution*, 2005, 60 (6) : 716-725.
- [11] Yadav G, Gokhale RS, Mohanty D. SEARCHPKS: a program for detection and analysis of polyketide synthase domains. *Nucleic Acids Research*, 2003, 31 (13) : 3654.
- [12] Yadav G, Gokhale RS, Mohanty D. Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *Journal of Molecular Biology*, 2003, 328 (2) : 335-363.
- [13] Haydock SF, Aparicio JF, Molnar I, Schwecke T, Khaw LE, König A, Marsden AFA, Galloway IS, Staunton J, Leadlay PF. Divergent sequence motifs correlated with the substrate specificity of (methyl) malonyl-CoA:acyl carrier protein transacylase domains in modular polyketide synthases. *FEBS Letters*, 1995, 374 (2) : 246-248.
- [14] Reeves CD, Murli S, Ashley GW, Piagentini M, Hutchinson CR, McDaniel R. Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. *Biochemistry*, 2001, 40 (51) : 15464-15470.
- [15] Ridley CP, Lee HY, Khosla C. Evolution of polyketide synthases in bacteria. *Proceedings of the National Academy of Sciences*, 2008, 105 (12) : 4595.
- [16] Castoe TA, Stephens T, Noonan BP, Calestani C. A novel group of type I polyketide synthases (PKS) in animals and the complex phylogenomics of PKSs. *Gene*, 2007, 392 (1 - 2) : 47-58.
- [17] John U, Beszteri B, Derelle E, Van de Peer Y, Read B, Moreau H, Cembella A. Novel insights into evolution of protistan polyketide synthases through phylogenomic analysis. *Protist*, 2008, 159 (1) : 21-30.
- [18] Kroken S, Glass NL, Taylor JW, Yoder OC, Turgeon BG. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proceedings of the National Academy of Sciences*, 2003, 100 (26) : 15670-15675.
- [19] Delsuc F, Brinkmann H, Philippe H. Phylogenomics and the reconstruction of the tree of life. *Nature Reviews Genetics*, 2005, 6 (5) : 361-375.
- [20] Hopwood DA, Bibb MJ, Chater KF, Kieser T, Bruton CJ, Kieser HM, Lydiate DJ, Smith CP, Ward JM, Schrepf H. Genetic manipulation of *Streptomyces*: a laboratory manual. 1985.
- [21] Kao CM, Pieper R, Cane DE, Khosla C. Evidence for two catalytically independent clusters of active sites in a

- functional modular polyketide synthase. *Biochemistry*, 1996, 35 (38) : 12363-12368.
- [22] Berger J, Rachlin AI, Scott WE, Sternbach LH, Goldberg MW. The isolation of three new crystalline antibiotics from *Streptomyces*¹. *Journal of the American Chemical Society*, 1951, 73 (11) : 5295-5298.
- [23] 孙宇辉, 邓子新. 南昌链霉菌中抗生素生物合成的遗传基础与分子机理研究. 上海交通大学的博士学位论文, 2003.
- [24] Grafe U, Schade W, Roth M, Radics L, Incze M, Ujszaszy K. Griseochelin, a novel carboxylic acid antibiotic from *Streptomyces griseus*. *The Journal of Antibiotics*, 1984, 37 (8) : 836-846.
- [25] Omura S, Ikeda H, Ishikawa J, Hanamoto A, Takahashi C, Shinose M, Takahashi Y, Horikawa H, Nakazawa H, Osonoe T. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. *Proceedings of the National Academy of Sciences*, 2001, 98 (21) : 12215.
- [26] Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3 (2). *Nature*, 2002, 417 : 141-147.

Phylogenomic analysis of modular polyketide synthases in actinomycetes and its application in product screening

Hao Wang, Ning Liu, Ying Huang*

(State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China)

Abstract: [**Objective**] To elucidate phylogenetic relationships of modular polyketide synthases (PKS) in actinomycetes, and to provide a guide for screening and discovery of polyketides. [**Methods**] We retrieved amino acid sequences of 190 ketosynthases (KS) and 195 acyltransferase (AT) of 20 modular polyketide synthases from database PKSDB, and constructed Neighbor-Joining trees based on amino acid sequences of KS, AT and KS + AT respectively using MEGA 4.0. We computed the mean distances within and between groups of KSs. We designed a pair of degenerate primers based on two conserved regions of KS, and screened 20 bioactive isolates by PCR. After sequencing the KS genes of positive strains, we constructed a Neighbor-joining tree of the 89 KSs identified in this study with the 160 known ones. We also fermented 13 KS-positive isolates and carried out chemical analysis of the fermentation extracts. [**Results**] KSs from the same PKSs of actinomycetes tended to group respectively into clades, and KSs responsible for synthesis of products with similar structures tended to cluster together. The mean distances within groups of KSs from each PKS pathway were less than 0.300, and the mean distances between groups of KSs from different PKS pathways were generally more than 0.300. All the ATs grouped into two big clades according to the substrate specificity. Some ATs from the same pathway were grouped within the two big clades respectively, and the rest were scattered. The tree of KS + AT integrated the topological structures of both KS tree and AT tree. The majority KSs from individual isolates tended to group into clades. Most KSs from four of the isolates grouped into four distinct clades, and most KSs from another five isolates respectively fell into four clades of known pathways. We isolated and identified three predicted compounds as expected. [**Conclusion**] Vertical evolution is likely to be the dominant evolutionary mode of KS genes in actinomycetes, while AT genes may have evolved mainly through horizontal gene transfer. Considering that KSs cluster according to the corresponding product structures, and that the mean distances 0.300 between groups of KSs could serve as an evaluation criteria for different PKS pathways, the phylogenomic analysis of KS is more suitable than that of AT or KS + AT for guiding the screening of polyketide compounds from actinomycetes.

Keywords: actinomycetes; evolution; polyketide synthase; ketosynthase; acyltransferase; phylogenomics; product screening

(本文责编:王晋芳)

Supported by the (Natural Science Foundation of China (30770002) and the National Programs for High Technology Research and Development of China (2007AA09Z420)

* Corresponding author. Tel: +86-10-64807311; E-mail: huangy@im.ac.cn

Received: 11 April 2010/ Revised: 6 May 2010