

DOI:10.12113/202006011

# 土地类芽胞杆菌 (*Paenibacillus terrae*) NK3-4 EsxA 结构与系统发育分析

于文清<sup>1,2,3</sup>, 闫凤超<sup>2</sup>, 刘文志<sup>1,2\*</sup>, 郑桂萍<sup>3</sup>, 肖俊杰<sup>1</sup>

(1. 上饶师范学院 生命科学院, 江西 上饶 334001; 2. 黑龙江省农垦科学院, 哈尔滨 150036;  
3. 黑龙江八一农垦大学, 黑龙江 大庆 163319)

**摘要:** 前期研究在植物根际促生菌土地类芽胞杆菌 (*Paenibacillus terrae*) NK3-4 中发现一个 EsxA 编码基因, 为明确该基因编码的蛋白的性质、结构及系统发生关系, 对该基因进行了生物信息学分析。分析表明, 该 EsxA 含有 91 个氨基酸, 分子质量 10 276.53 Da, 理论 pI 5.29, 分子式为 C<sub>445</sub>H<sub>711</sub>N<sub>125</sub>O<sub>146</sub>S<sub>4</sub>, 弱酸性, 亲水, 具有 WEG 保守基序, 属于 WXG 超级家族成员; 建模预测表明, 自然状态下 EsxA 形成不对称的同源二聚体, 其中每个亚基都由一个  $\beta$  折叠连接两个  $\alpha$  螺旋组成, 两个  $\alpha$  螺旋反向平行排列; 二聚体中两个亚基的肽链呈反向排列, 所有 N 末端和 C 末端均暴露在外, 形成棒状表面形态, 其中一个亚基的 N 端的不规则卷曲形成与棒状二聚体垂直的短柱形凸起; 系统发育分析显示, EsxA 在种内及种间进化关系虽是不保守的, 但类芽胞杆菌源的 EsxA 与病原细菌的 EsxA 同源性较低, 暗示类芽胞杆菌源 EsxA 可能与病原微生物的 EsxA 具有截然不同的功能。结果为包括 NK3-4 菌株在内的类芽胞杆菌属 EsxA 外源分泌表达条件, 活性保持及功能的深入研究提供了理论依据。

**关键词:** 土地类芽胞杆菌; EsxA; 生物信息学分析; 蛋白 3D 结构; 系统发育

中图分类号: Q71 文献标志码: A 文章编号: 1672-5565(2021)04-270-06

## Structure and phylogeny analysis of *Paenibacillus terrae* NK3-4 EsxA

YU Wenqing<sup>1,2,3</sup>, YAN Fengchao<sup>2</sup>, LIU Wenzhi<sup>1,2\*</sup>, ZHENG Guiping<sup>3</sup>, XIAO Junjie<sup>1</sup>

(1. College of Life Science, Shangrao Normal University, Shangrao 334001, Jiangxi, China;  
2. Heilongjiang Academy of Agricultural Reclamation Sciences, Harbin 150036, China;  
3. Heilongjiang Bayi Agricultural University, Daqing 163319, Heilongjiang, China)

**Abstract:** Earlier studies found an EsxA gene in the plant growth-promoting rhizobacteria *Paenibacillus terrae* NK3-4. To clarify the properties, structure, and phylogenetic relationship of the EsxA gene coded protein, a series of analyses were conducted using bioinformatics technology. The analysis of basic properties shows that the protein EsxA contained 91 amino acids with a molecular mass of 10 276.53 Da, the theoretical pI was 5.29, the molecular formula was C<sub>445</sub>H<sub>711</sub>N<sub>125</sub>O<sub>146</sub>S<sub>4</sub>, weakly acidic and hydrophilic, it had a conservative motif of WEG, and was a member of WXG superfamily. Modeling predictions show that EsxA formed an asymmetric homodimer in a natural state, in which each subunit consisted of two  $\alpha$ -helices that were connected by a  $\beta$ -sheet, and the two  $\alpha$ -helices arranged in antiparallel; the peptide chains of the two subunits in the dimer arranged in reverse, and all N- and C-terminuses were exposed to form a rod-shaped surface morphology. The irregular curl at the N-terminus of one subunit formed a short columnar protrusion, perpendicular to the rod-shaped dimer. Phylogenetic analysis shows that EsxAs in *Paenibacillus* had a low homology with that in pathogenic bacteria, although not conservative among *Paenibacillus* species, which indicates that EsxA in *Paenibacillus* may have distinct-different functions with that in the pathogenic bacteria. The results can provide theoretical bases for the studies of the exogenous secretion expression conditions, activity maintenance, and functions of EsxA in *Paenibacillus* genus including *Paenibacillus terrae* NK3-4.

**Keywords:** *Paenibacillus terrae*; EsxA; Bioinformatics analysis; 3D structure of protein; Phylogeny

收稿日期: 2020-06-30; 修回日期: 2020-11-18.

基金项目: 国家自然科学基金项目 (No.32060633); 黑龙江省自然科学基金项目 (No.LH2019C079).

作者简介: 于文清, 女, 副研究员, 研究方向: 植物根际促生菌及其功能基因与植物诱抗型蛋白激发子. E-mail: ywq19801005@163.com.

\* 通信作者: 刘文志, 男, 研究员, 研究方向: 农业微生物研究与应用. E-mail: liuwz19801005@sina.com.

EsxAs (早期分泌抗原靶 6, Early secreting antigenic target-6, ESAT6) 是一类小分子分泌蛋白, 一般含有 100 左右个氨基酸, 具有保守的 WXG 氨基酸基序<sup>[1]</sup>, 因此被称为 WXG 超级家族成员。该蛋白呈螺旋状结构, 与 EsxB (CFP-10) 共同组成 VII 型分泌系统 (T7SS), EsxA 最先在动物病原菌结核分支杆菌 (*Mycobacterium tuberculosis*) 中发现<sup>[2]</sup>, 与其致病力密切相关<sup>[3-5]</sup>, 后陆续在其它动物病原菌中也发现了该类蛋白<sup>[6]</sup>。EsxA 除了与动物致病菌的致病性相关外<sup>[7-8]</sup>, 还能够诱导动物的免疫原活性<sup>[9-11]</sup>, 也有研究发现, 在单核球增多性李斯特菌中 (*Listeria monocytogenes*), EsxA 与宿主病原菌的致病性无关<sup>[12]</sup>。在非致病细菌中也存在 EsxA 编码基因, 如在天蓝色链霉菌中, EsxA 与 EsxB 形成同源二聚体, 与宿主菌成孢相关<sup>[13]</sup>。

前期研究发现, 植物根际促生菌土地类芽胞杆菌 (*Paenibacillus terrae*) NK3-4 基因组中携带着 EsxA 编码基因, 但 EsxA 在类芽胞杆菌中发挥何种功能尚无相关报道, 是否与类芽胞杆菌对植物的促生效应相关也尚未可知。本文采用生物信息学技术分析了该 EsxA 的基本性质、蛋白的 3D 结构及系统发育关系, 以期包括 NK3-4 菌株在内的类芽胞杆菌 EsxA 基因外源分泌表达、活性保持条件及功能的深入研究提供理论依据。

## 1 材料与方法

### 1.1 NK3-4 菌株 EsxA 基本性质分析

将 NK3-4 基因组中注释为毒力因子 (Virulence factor EsxA) 的氨基酸序列提交 ExPASy Protparam tool (<https://web.expasy.org/cgi-bin/protparam/protparam>) 分析其基本性质, 将氨基酸序列提交 NCBI 中的 BLSTp 进行序列相似性比对, 并检索保守域。

### 1.2 NK3-4 菌株 EsxA3D 结构预测

将 NK3-4 菌株 EsxA 氨基酸序列提交在线蛋白质结构预测网站 SWISS-MODEL (<https://www.swissmodel.expasy.org/interactive/rgsxPX/models/>), 通过同源建模预测其 3D 结构<sup>[14]</sup>, 大致步骤如下:

(1) 模板搜索选择及建模: 将序列提交 SWISS-MODEL 后, 会自动匹配相似的模板, 对于每个识别的模板, 都列出了预测模板-目标匹配质量, 选择质量最高的模板建模。将模板与目标序列比对, 模板与序列之间保守匹配的结构被从模板复制建模, 用片段库插入或删除, 再重新建模, 而后重建侧链, 最后利用最小作用力的几何学特性构建模型。

(2) 建模质量评估及配基建模: 利用 QMEAM 评分功能进行全局性和残基模型质量评估。配基建模时, 只有满足以下全部条件后才代表存在配基: (i) 配体在模板库中注释为生物学相关; (ii) 配体与模型联结在一起; (iii) 配体与模型之间不冲突; (iv) 与配体结合的残基在模板与目标蛋白之间是特异保守的。

### 1.3 NK3-4 菌株 EsxA 系统发育树构建

利用 MEGA6 软件对 NK3-4 菌株 EsxA 及其同源较高的蛋白序列进行聚类分析, 这些蛋白的氨基酸序列从 Uniprot 数据库 (<https://www.uniprot.org/>) 中获得。对比序列类型选择“Translated protein Sequences”, 统计方法采用邻接法 (Neighbor-joining), 模板采用最大似然法 (Maximum Composite Likelihood method)。系统发生检测采用自举法 (Bootstrap method), 成对比对参数 (Pairwise Alignment) 为缺口开启罚分 (Gap Opening Penalty): 10, 且缺口延伸罚分 (Gap Extension Penalty): 0.1; 多重比较参数 (Multiple Alignment) 为缺口开启罚分: 10, 缺口延伸罚分: 0.2; 蛋白权值矩阵 (Protein Weight Matrix): Gonnet; 特殊氨基酸罚分 (Residue-specific Penalties): ON, 亲水性罚分 (Hydrophilic Penalties): ON; 缺口分离距离 (Gap Separation Distance): 4; 末端缺口分离 (End Gap Separation): OFF。

## 2 结果与分析

### 2.1 NK3-4 菌株 EsxA 的基本性质

#### 2.1.1 NK3-4 菌株 EsxA 的理化性质

NK3-4 菌株 EsxA 含有 91 个氨基酸, 分子质量 10 276.53 道尔顿, 理论等电点 5.29, 分子式为  $C_{445}H_{711}N_{125}O_{146}S_4$ , 原子数目 1 431, 消光系数 (280 nm) 5 500, Abs 0.1% (= 1 g/L) 0.535, 估计半衰期为 30 h (哺乳动物的网状细胞, 体外), >20 h (酵母, 体内) 和 >10 h (大肠杆菌, 体内)。不稳定指数 46.28, 归为不稳定。脂肪族指数 65.38, 大致亲水性平均值 -0.548。表明 NK3-4 菌株 EsxA 为弱酸性、亲水性蛋白质, 三级结构不稳定。

#### 2.1.2 NK3-4 菌株 EsxA 的保守域

CDD 分析表明, EsxA 含有 WXG100 ESAT6 保守域及 WXG 高度保守基序第 44 到 46 个氨基酸: W (甘氨酸)-E (谷氨酸)-G (色氨酸), 是 WXG 超级家族成员, 该 EsxA 氨基酸序列及其保守域 (见图 1)。

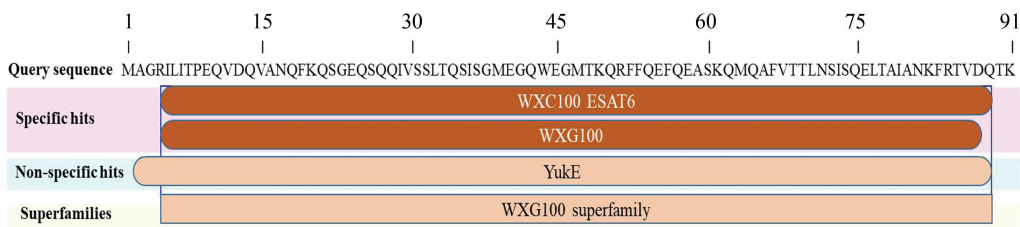


图 1 土地类芽胞杆菌 NK3-4 EsxA 氨基酸序列及其保守域

Fig. 1 Amino acid sequence and conserved domain of EsxA in *Paenibacillus terrae* NK3-4

## 2.2 NK3-4 菌株 EsxA 3D 结构预测结果

### 2.2.1 模板筛选及模型构建结果

将 NK3-4 菌株 EsxA 氨基酸序列提交 SWISS-MODEL 后,共查出 417 个模板,其中一个名为 EsxA 的蛋白与提交序列识别率 30%,相似性 36%,可以该蛋白为模板建模。蛋白模板是同源二聚体,没有配基,A、B 链(见图 2)分别覆盖 NK3-4 菌株 EsxA 1~90 及 7~90 个氨基酸,覆盖率 99%。

### 2.2.2 建模质量评估结果

建模评估表明,分子概率统计表明,分子概率分值 1.22,冲突分数 4.42,拉氏亲和 99.41%,拉氏离群值 0.0%,旋转异构体异常值 0.0%,C-Beta 偏差为 0。全局性模型质量评估显示,QMEAN 为 -1.06(见图 3a);局部模板质量评估表明,预测的局部文库与目标序列预测相似性(见图 3b);预测模型与同源蛋白

质三级结构的相似性分析表明,与 PDB 结构的非冗余集比较,标准 QMEAN4 分值为 0.74(见图 3c),结果表明(散点图中以红色五角星标记),Z 得分 < 1,显示该预测模型结构稳定,是最佳的理想结构。

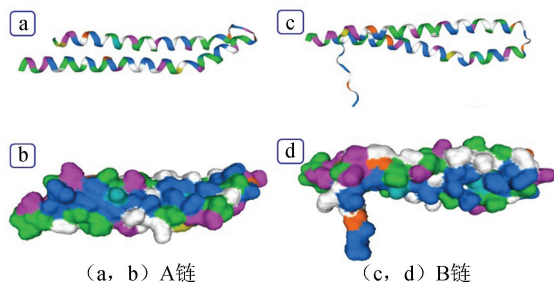


图 2 土地类芽胞杆菌 NK3-4 EsxA 结构预测模板  
Fig.2 Template for *Paenibacillus terrae* NK3-4 EsxA 3D structure prediction

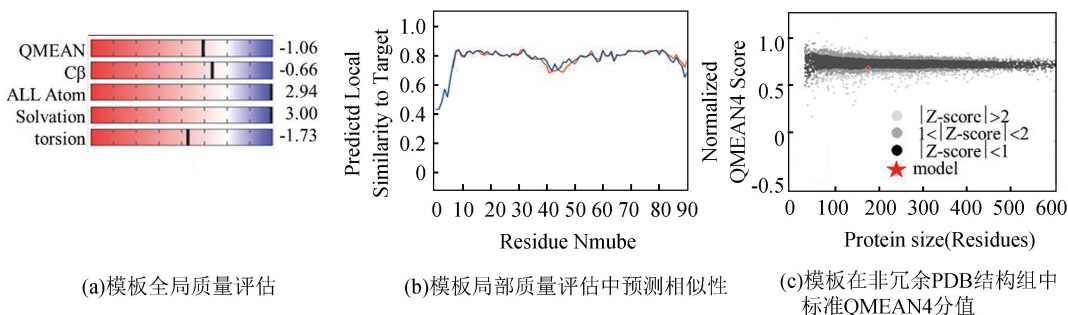


图 3 土地类芽胞杆菌 NK3-4 EsxA 结构预测模板质量评估结果

Fig.3 Assessment results of template for *Paenibacillus terrae* NK3-4 EsxA structure prediction

拉氏构象分析表明,预测模型坏键 0/1 394,坏角 6/1 874,包括 A66 Phe、(B8 Thr-B9 Pro)、A59 Ala、A82 Asn、B11 Gln、B49 Lys,两个亚基中,分别有 97.8%和 95.6%的氨基酸残基位于深绿色最佳允许区域,且全部残基均位于额外允许区域(见图 4)。表明预测模型中全部氨基酸残基均形成了合理的二面角,构成的蛋白质结构稳定。

### 2.2.3 NK3-4 菌株 EsxA 3D 结构预测结果

NK3-4 菌株 EsxA 结构分析显示,EsxA 形成同源二聚体,其中每个亚基都由一个  $\beta$  折叠连接两个  $\alpha$  螺旋组成,这两个  $\alpha$  螺旋反向平行排列。二聚体中两个亚基的肽链反向排列,所有 N 末端和两个 C 末端均暴露在外。这个二聚体是不对称的结构,其中一个亚基(B 链)自由的 N 端肽段暴露出来,形成 3D 结构(见图 5)。

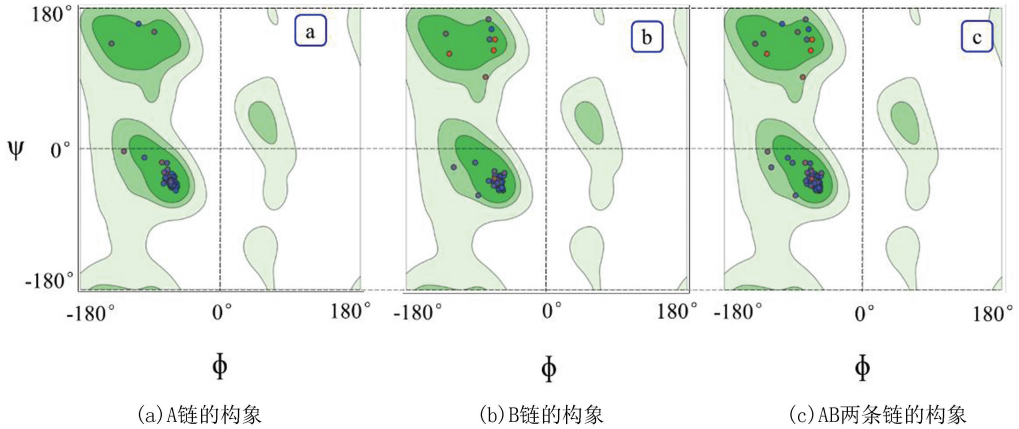


图 4 土地类芽胞杆菌 NK3-4 EsxA 结构预测模板拉氏构象图

Fig. 4 Ramachandran plots of template for *Paenibacillus terrae* NK3-4 EsxA

注:图中深绿色区域代表键角稳定,绿色区域代表键角虽然不稳定,但可在自然界存在,淡绿色区域代表键角不稳定。

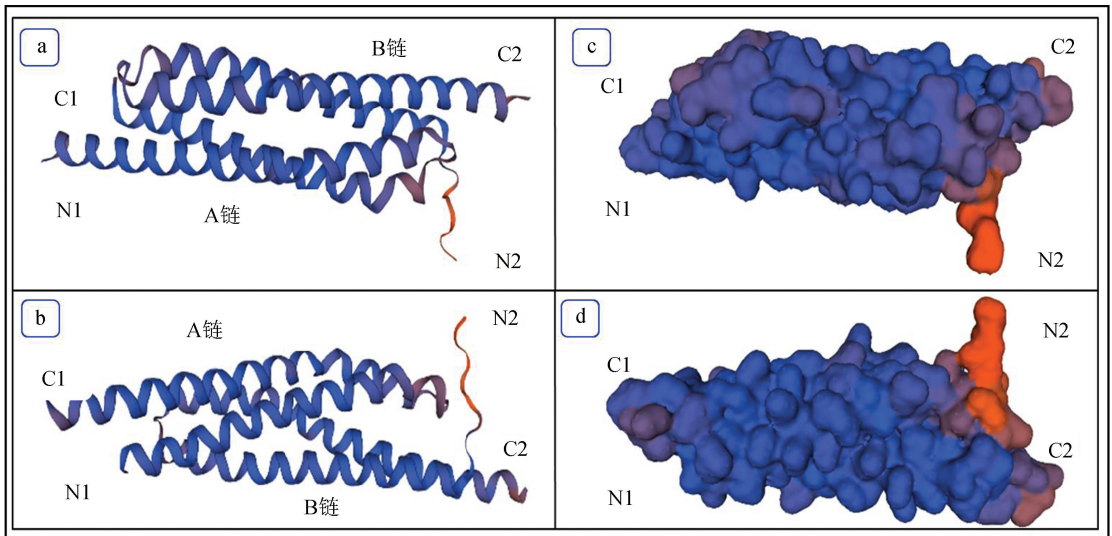


图 5 预测的 NK3-4 菌株 EsxA 激发子结构

Fig.5 Predicted structure of *Paenibacillus terrae* NK3-4 EsxA

注:a, b 为 EsxA 二聚体;c 和 d 为计算出的二聚体表面电势图(蓝色-正电荷,橙色-负电荷,紫色-中性);b 和 d 分别为 a 和 c 旋转 180°后的图象;图中标注了每个肽链的 C 端和 N 端(标注 1 的为 A 链,标注 2 的为 B 链)。

聚类分析表明,EsxA 主要分成两个大簇(见图 6a, 6b),NK3-4 菌株 EsxA 与其中一簇中的两个 *P. terrae* 及两株多粘类芽胞杆菌 (*Paenibacillus polymyxa*) 聚为一个小支,该小支内的氨基酸序列是完全相同的,同时与其它类芽胞杆菌聚为一小簇(见图 6c),这一小簇再依次与侧短芽胞杆菌 (*Brevibacillus laterosporus*),*Paenibacillus yonginensis*, 高温放线菌 (*Thermoactinomyces* sp) 及韩氏芽胞杆菌 (*Bacillus koreans*) 聚为一个大簇(见图 6a);在另一个大簇中,另一些类芽胞杆菌属菌株 EsxAs 聚为一小

簇(见图 6b),这两个大簇与红色梭菌 (*Clostridium roseum*) 和毡毛梭菌 (*Clostridium felsineum* DSM 794) 最终聚到一起。可见,EsxAs 在种间及属间进化关系是不保守的,它们被明显的分为两个进化关系较远的类簇,EsxAs 做为一类分泌型小分子蛋白,广泛存在于包括类芽胞杆菌的微生物中,且包括NK3-4菌株在内的类芽胞杆菌的 EsxAs 与动物致病菌的毒力因子 EsxAs 亲缘关系甚远。亲缘关系较远可能是 EsxA 功能存在差异的原因。

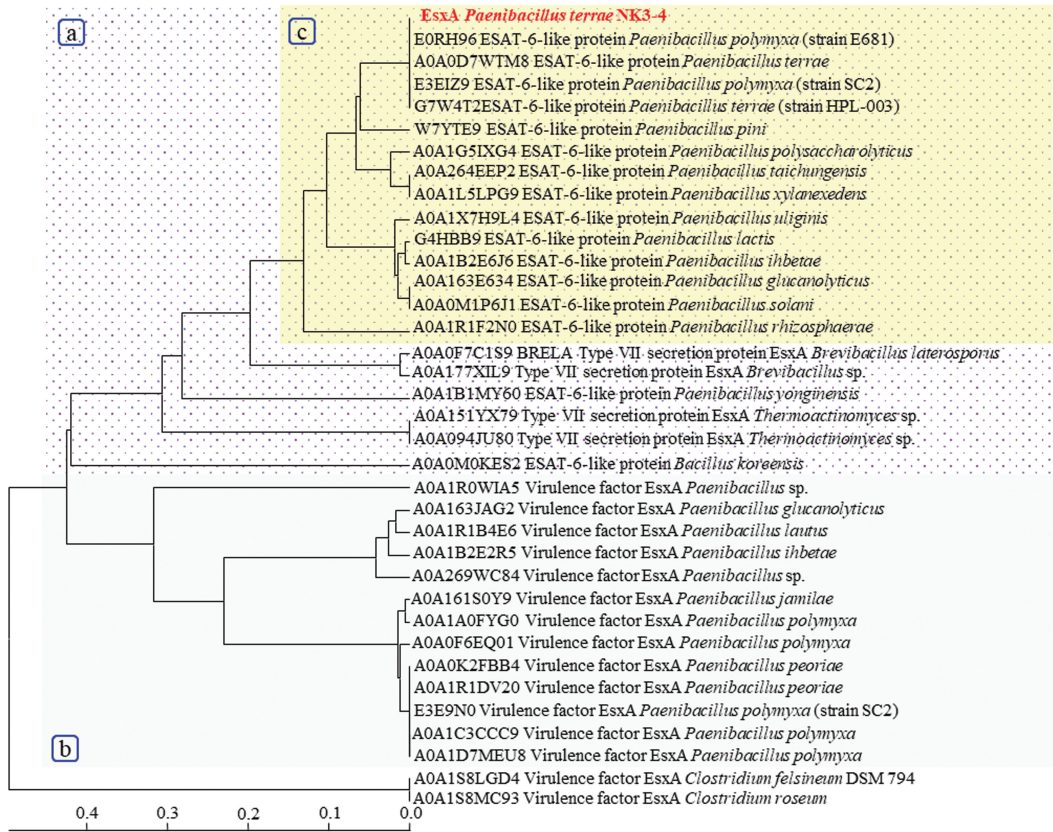


图 6 土地类芽胞杆菌 NK3-4 EsxA 的系统发育树

Fig. 6 Phylogenetic tree of *Paenibacillus terrae* NK3-4 EsxA

注:树上分类单元标注了各蛋白在 Uniport 中的登录号及来源菌种或菌株名称,a,b,c 分别为不同亚分支。

### 3 讨论

本文分析了类芽胞杆菌 NK3-4 EsxA 的基本性质、3D 分子结构及系统发育关系,为其后续研究与应用提供了理论依据,主要表现在明确该蛋白的基本理化性质,可为 EsxA 基因外源分泌表达条件优化提供依据,也可为分泌表达至胞外的蛋白纯化条件优化提供参考,还可为优化蛋白活性保持条件提供依据。3D 结构预测有助于了解 EsxA 的特殊结构及功能,有助于认识该蛋白可能存在的与其它分子之间的相互作用的位点,并指导进行功能确认的生物学实验设计。系统发育关系分析可为判断该 EsxA 在进化过程中的保守性与变异性提供依据,本文将 NK3-4 菌株 EsxA 与其相似性最高的 34 个蛋白质进行聚类分析,表明这些 EsxA 在种内与种间均是不保守的,说明其编码基因变异性强,是一类进化速度较快的蛋白,这也是我们认为其在类芽胞杆菌中与在病原细菌中具有不同功能的依据。在包括 NK3-4 菌株在内的类芽胞杆菌中,在类芽胞杆菌与植物进行互作时,其 EsxA 是否通过分泌到细胞外参与类芽胞杆菌对植物的促生效

应,如参与宿主与植物的互作,其作用的分子机制如何,将是作者下一步研究将要揭示的答案。

### 4 结论

采用生物信息学技术,明确了土地类芽胞杆菌 NK3-4 基因组携带的 EsxA 基因编码的小分子分泌型蛋白的理化性质,分析了保守域并定位了及保守基序 WEG。采用建模的方式,通过对模板的一系列质量评估,预测了蛋白质的 3D 结构,并基于该 EsxA 氨基酸序列构建了系统发育树,系统发育分析表明,包括 NK3-4 菌株在内的类芽胞杆菌的 EsxA 与致病细菌的 EsxA 同源关系较远,这暗示非致病菌分泌的 EsxA 可能具有与病原菌分泌的 EsxA 具有截然不同的功能。结果为 NK3-4 菌株在内的非致病菌 EsxA 基因外源分泌表达条件、蛋白活性保持及功能的深入研究奠定了理论基础。

### 参考文献 (References)

[1] CALLAHAN B, NGUYEN K, COLLINS A, et al. Conservation of structure and protein-protein interactions mediated

- by the secreted mycobacterial proteins EsxA, EsxB, and EspA[J]. *Journal of Bacteriology*, 2010, 192(1): 326-335. DOI:10.1128/JB.01032-09.
- [2] POLLOCK J M, ANDERSEN P. The potential of the ESAT-6 antigen secreted by virulent *Mycobacteria* for specific diagnosis of tuberculosis [J]. *Journal of Infectious Diseases*, 1997, 175(5): 1251-1254. DOI: 10.1086/593686.
- [3] BERTHET F X, RASMUSSEN P B, ROSENKRANDS I, et al. A *Mycobacterium tuberculosis* operon encoding ESAT = 6 and a novel low-molecular-mass culture filtrate protein (CFP-10)[J]. *Microbiology*, 1998, 144(11): 3195-3203. DOI: 10.1099/00221287-144-11-3195.
- [4] ULRICHS T, MUNK M E, MOLLENKOPF H, et al. Differential T cell responses to *Mycobacterium tuberculosis* ESAT6 in tuberculosis patients and healthy donors[J]. *European Journal of Immunology*, 1998, 28(12): 3949-3958. DOI: 10.1002/(SICI)1521-4141(199812)28:12<3949::AID-IMMU3949>3.0.CO;2.
- [5] JAVIER A, HITRA K, LIN L, et al. N $\alpha$ -Acetylation of the virulence factor EsxA is required for *mycobacterial cytosolic* translocation and virulence[J]. *Journal of Biological Chemistry*, 2020, 295(17): jbc.RA119.012497. DOI: 10.1074/jbc.RA119.012497.
- [6] GARUFI G, BUTLER E, MISSIAKAS D. ESAT-6-like protein secretion in *Bacillus anthracis* [J]. *Journal of Bacteriology*, 2008, 190(21): 7004-7011. DOI: 10.1128/JB.00458-08.
- [7] SCHULTHESS B, BLOES D A, BERGER-BÄCHI B. Opposing roles of  $\sigma^B$  and  $\sigma^{B'}$ -controlled SpoVG in the global regulation of *esxA* in *Staphylococcus aureus*[J]. *BMC Microbiology*, 2012, 12: 17. DOI: 10.1186/1471-2180-12-17.
- [8] MA Y, KEIL V, SUN J. Characterization of *mycobacterium tuberculosis* EsxA membrane insertion: roles of N- and C-terminal flexible arms and central helix-turn-helix motif[J]. *Journal of Biological Chemistry*, 2015, 290(11): 7314-7322. DOI: 10.1074/jbc.M114.622076.
- [9] ZHOU H, DU H, ZHANG H, et al. EsxA might as a virulence factor induce antibodies in patients with *Staphylococcus aureus* infection [J]. *Brazilian Journal of Microbiology*, 2013, 44(1): 267-271. DOI: 10.1590/s1517-83822013005000019.
- [10] 马艳平, 梁志凌, 马江耀, 等. 罗非鱼无乳链球菌 EsxA 基因克隆与原核表达[J]. *动物医学进展*, 2016, 37(4): 35-39. DOI: 10.3969/j.issn.1007-5038.2016.04.007.
- MA Yanping, LIANG Zhiling, MA Jiangyao, et al. Cloning and prokaryon expression of EsxA gene of *Streptococcus agalactiae* in tilapia [J]. *Progress in Veterinary Medicine*, 2016, 37(4): 35-39. DOI: 10.3969/j.issn.1007-5038.2016.04.007.
- [11] 易鸳鸯, 赵亚南, 马鹏睿, 等. 牛乳源金黄色葡萄球菌 EsxA 蛋白的表达及免疫原性分析[J]. *生物工程学报*, 2018, 34(5): 694-702. DOI: CNKI:SUN:SHWU.0.2018-05-009.
- YI Yuanyang, ZHAO Yanan, MA Pengrui, et al. Expression and immunogenicity analysis of EsxA protein of *Staphylococcus aureus* isolated from cow milk[J]. *Chinese Journal of Biotechnology*, 2018, 34(5): 694-702. DOI: CNKI:SUN:SHWU.0.2018-05-009.
- [12] WAY S S, WILSON C. The *Mycobacterium tuberculosis* ESAT-6 homologue in *Listeria monocytogenes* is dispensable for growth in vitro and in vivo[J]. *Infection and Immunity*, 2005, 73(9): 6151-6153. DOI: 10.1128/IAI.73.9.6151-6153.2005.
- [13] AKPE S R S, FACEY P D, FERNANDEZ-MARTINEZ L, et al. A heterodimer of EsxA and EsxB is involved in sporulation and is secreted by a type VII secretion system in *Streptomyces coelicolor* [J]. *Microbiology*, 2010, 156(6): 1719-1729. DOI: 10.1099/mic.0.037069-0.
- [14] WATERHOUSE A, BERTONI M, BIENERT S, et al. SWISS-MODEL: homology modelling of protein structures and complexes [J]. *Nucleic Acids Research*, 2018, 46 (Web Server Issue): W296-W303. DOI: 10.1093/nar/gky427.