非核糖体多肽合成过程中腺苷化结构域特异性识别 Ala、Leu 机制研究

赵鹏，王思宇，叶景润，陈伟*
(青岛农业大学海洋科学与工程学院，山东青岛 266000)

表 S1 所有 A 结构域 NCBI 基因号汇总

Table S1 Summary of NCBI Gene Numbers for All A Domain Structures

A domain	NCBI ID
Ala	O94205 P0A398 P35854 P39581 P59591 P68876 P68878 Q01886 Q09164 Q50857 Q53526 Q83VS0 Q83VS1 Q88VM6 Q8K744 Q8NJX1 Q8P0J9 Q8VM67 Q8Y8D4 Q92D47 Q99ZA6 Q9CG49 Q9FB26 Q9FDB3 Q9RLP6 Q9RNB1 Q9S1A9 Q9X2N4 A9LJA2 Q0VZ71 AAM78457.1 XP_065971804 CAB39315.1 AAA33023.2 AN2545 WP_013308170
Leu	O30409 O52819 O66069 O66070 O68006 O69246 P27206 Q04747 Q08787 Q09164 Q44928 Q45295 Q45675 Q84BQ4 Q84BQ5 Q84BQ6 Q8G982 Q8GAQ6 Q8NJX1 Q939Z1 Q9RAH2 Q9RAH4 Q9RNB0 Q9S1A8 P0C063 Q50JA3 Q0W980 B6SF67 D1FVF0 AAM78457.1 XP_065971804 AN2545

表 S2 特异性识别不同氨基酸的 A 结构域的数目

Table S2 The number of A-domains that specifically recognize different amino acids

Amino acid	The number of A domains	Amino acid	The number of A domains
Ala	38	Leu	36
Val	29	Thr	27
Ser	25	Gly	23
Pro	22	Cys	19
Tyr	18	Asn	17
Phe	16	Hpg	14
Gln	13	Dhb	12
Asp	12	Glu	11
Orn	11	Ile	11
Trp	10	Aad	10
Dab	8	Pip	7
Bht	7	Iva	7
Dhpg	6	Arg	5
Lys	4		

表 S3 特异性识别 Ala、Leu 的 A 结构域的十个特异性差异位点

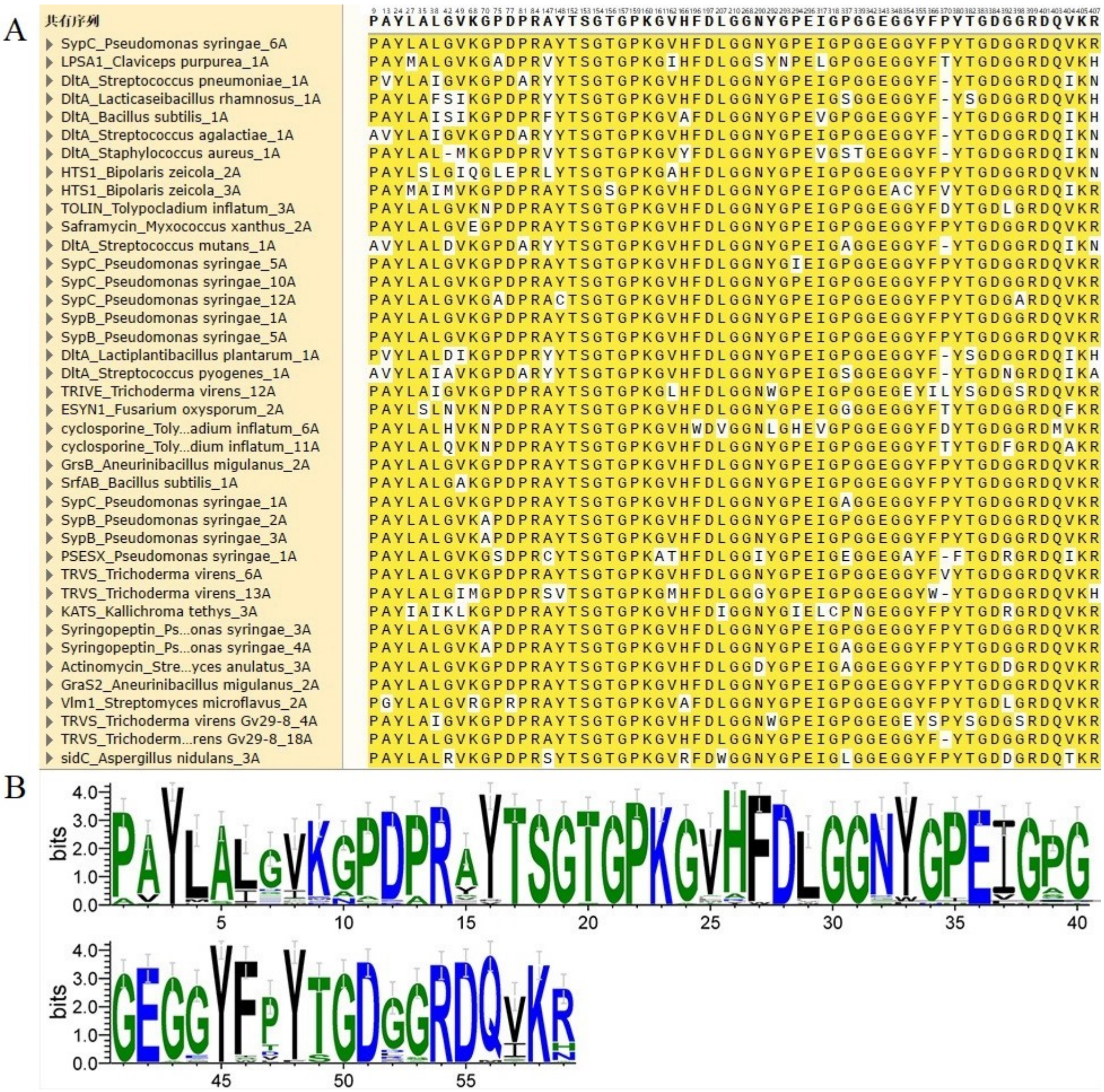
Table S3 Ten specific difference sites from the A domain of Ala and Leu

氨基酸位置信息	特异性识别 Ala 的 A 结构域	特异性识别 Leu 的 A 结构域
56	45%Ser	89%Ser
98	50%Leu	92%Leu
199	87%Ser	61%Ser
243	95%Pro	44%Pro
267	39%Ala	95%Gly
269	95%Glu	64%Asp
299	55%Val	83%Val
302	63%Thr	89%Thr
351	66%Val	81%Leu
407	61%Arg	97%Arg

表 S4 通过三级结构模型合理性评估的结构

Table S4 Structures evaluated by the rationality assessment of the tertiary structure model

A domain	Verify 3D	ERRAT	PROCHECK
DltA_*Bacillus subtilis*_1A	81.91%	91.2	93.9633%
DltA_*Lacticaseibacillus rhamnosus*_1A	85.18%	93.2	94.087%
DltA_*Lactiplantibacillus plantarum*_1A	86.22%	91.5	91.667%
DltA_*Listeria innocua*_1A	85.86%	91.7	95.09%
DltA_*Listeria monocytogenes*_1A	86.85%	91.7	91.969%
DltA_*Streptococcus mutans*_1A	86.00%	90.5	91.795%
DltA_*Streptococcus pneumoniae*_1A	80.75%	91.0	85.974%
easA_*Aspergillus nidulans* FGSC A4_5A	82.75%	90.1	92.969%
McyA_*Microcystis aeruginosa* PCC 7806_2A	81.42%	91.7	95.812%
MYCSM_*Mycolicibacterium smegmatis*_3A	91.24%	94.4	89.125%
PAPX_*Paenibacillus polymyxa*_6A	94.00%	91.7	89.231%
Saframycin_*Myxococcus xanthus*_2A	87.18%	92.0	91.053%
SypA_*Pseudomonas syringae*_5A	88.07%	90.3	93.229%
SypB_*Pseudomonas syringae*_1A	86.92%	91.6	92.105%
SypB_*Pseudomonas syringae*_5A	91.05%	93.1	94.180%
SypC_*Pseudomonas syringae*_5A	94.32%	93.0	90.909%
SypC_*Pseudomonas syringae*_6A	96.81%	92.1	93.939%
SypC_*Pseudomonas syringae*_12A	93.16%	91.7	91.623%
TRVS_*Trichoderma virens* Gv29-8_3A	80.89%	92.3	87.817%
TRVS_*Trichoderma virens* Gv29-8_8A	81.46%	90.6	91%
AMBA_*Amycolatopsis balhimycina*_1A	90.77%	90.7	91.351%
arthrofactinA_*Pseudomonas* sp. MIS38	94.87%	91.5	92.857%
arthrofactinB_*Pseudomonas* sp. MIS38	89.63%	92.1	89.059%
arthrofactinB_*Pseudomonas* sp. MIS38	86.34%	92.4	94.118%
arthrofactinC_*Pseudomonas* sp. MIS38	95.13%	90.7	91.421%
BacA_*Bacillus licheniformis*_3A	84.11%	92.9	88.503%
cyclosporine_*Tolypocladium inflatum*_10A	81.16%	90.2	92.612%
grsB_*Brevibacillus brevis*_4A	87.06%	91.8	94.531%
LchAA_*Bacillus licheniformis*_3A	93.26%	91.9	94.920%
LchAB_*Bacillus licheniformis*_3A	94.60%	91.3	93.298%
LicA_*Bacillus paralicheniformis*_3A	94.82%	92.2	95.2%
LicB_*Bacillus licheniformis*_3A	96.92%	91.3	93.583%
McyB_*Microcystis aeruginosa*_1A	89.09%	91.2	85.45%
PCZA363.3_*Amycolatopsis orientalis*_1A	90.18%	91.3	96.267%
PLAI_*Planktothrix agardhii*_1A	86.03%	91.7	89.514%
SrfAA_*Bacillus subtilis*_2A	85.18%	92.8	94.805%
SrfAA_*Bacillus subtilis*_3A	95.10%	94.3	94.149%
SrfAC_*Bacillus subtilis*_1A	89.45%	93.1	93.995%
STAA_*Stigmatella aurantiaca*_2A	89.49%	92.0	86.772%
TRVS_*Trichoderma virens*_3A	82.63%	93.4	87.056%
SrfAB_*Bacillus subtilis*_3A	91.24%	91	87.733%



共有序列	PAYLALGVKGPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFPYTGDGGRDQVKR
SypC_Pseudomonas syringae_6A	PAYLALGVKGPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFPYTGDGGRDQVKR
LPSA1_Claviceps purpurea_1A	PAYMALGVKGADPRVYTSGTGPKGIHFDLGGSYNPELGPGGEGGYFTYTGDGGRDQVKH
DltA_Streptococcus pneumoniae_1A	PVYLAIGVKGPDARYYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYF-YTGDGGRDQIKN
DltA_Lacticaseibacillus rhamnosus_1A	PAYLAFSIKGPDPRYYTSGTGPKGVHFDLGGNYGPEIGSGGEGGYF-YSGDGGRDQVKH
DltA_Bacillus subtilis_1A	PAYLAISIKGPDPRFYTSGTGPKGVAFDLGGNYGPEVGPGGEGGYF-YTGDGGRDQIKH
DltA_Streptococcus agalactiae_1A	AVYLAIGVKGPDARYYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYF-YTGDGGRDQIKN
DltA_Staphylococcus aureus_1A	PAYLAL-MKGPDPRVYTSGTGPKGVYFDLGGNYGPEVGSTGEGGYF-YTGDGGRDQIKN
HTS1_Bipolaris zeicola_2A	PAYLSLGIQGLEPRLYTSGTGPKGAHFDLGGNYGPEIGPGGEGGYFPYTGDGGRDQVKN
HTS1_Bipolaris zeicola_3A	PAYMAIMVKGPDPRAYTSGSGPKGVHFDLGGNYGPEIGPGGEACYFVYTGDGGRDQIKR
TOLIN_Tolypocladium inflatum_3A	PAYLALGVKNPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFDYTGDLGRDQVKR
Saframycin_Myxococcus xanthus_2A	PAYLALGVEGPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFPYTGDGGRDQVKR
DltA_Streptococcus mutans_1A	AVYLALDVKGPDARYYTSGTGPKGVHFDLGGNYGPEIGAGGEGGYF-YTGDGGRDQIKN
SypC_Pseudomonas syringae_5A	PAYLALGVKGPDPRAYTSGTGPKGVHFDLGGNYGIEIGPGGEGGYFPYTGDGGRDQVKR
SypC_Pseudomonas syringae_10A	PAYLALGVKGPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFPYTGDGGRDQVKR
SypC_Pseudomonas syringae_12A	PAYLALGVKGADPRACTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFPYTGDGARDQVKR
SypB_Pseudomonas syringae_1A	PAYLALGVKGPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFPYTGDGGRDQVKR
SypB_Pseudomonas syringae_5A	PAYLALGVKGPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFPYTGDGGRDQVKR
DltA_Lactiplantibacillus plantarum_1A	PVYLALDIKGPDPRYYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYF-YSGDGGRDQIKH
DltA_Streptococcus pyogenes_1A	AVYLAIAVKGPDARYYTSGTGPKGVHFDLGGNYGPEIGSGGEGGYF-YTGDNGRDQIKA
TRIVE_Trichoderma virens_12A	PAYLAIGVKGPDPRAYTSGTGPKGLHFDLGGNWGPEIGPGGEGEYILYSGDGSRDQVKR
ESYN1_Fusarium oxysporum_2A	PAYLSLNVKNPDPRAYTSGTGPKGVHFDLGGNYGPEIGGGGEGGYFTYTGDGGRDQFKR
cyclosporine_Toly...adium inflatum_6A	PAYLALHVKNPDPRAYTSGTGPKGVHWDVGGNLGHEVGPGGEGGYFDYTGDGGRDMVKR
cyclosporine_Toly...dium inflatum_11A	PAYLALQVKNPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFTYTGDFGRDQAKR
GrsB_Aneurinibacillus migulanus_2A	PAYLALGVKGPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFPYTGDGGRDQVKR
SrfAB_Bacillus subtilis_1A	PAYLALGAKGPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFPYTGDGGRDQVKR
SypC_Pseudomonas syringae_1A	PAYLALGVKGPDPRAYTSGTGPKGVHFDLGGNYGPEIGAGGEGGYFPYTGDGGRDQVKR
SypB_Pseudomonas syringae_2A	PAYLALGVKAPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFPYTGDGGRDQVKR
SypB_Pseudomonas syringae_3A	PAYLALGVKAPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFPYTGDGGRDQVKR
PSESX_Pseudomonas syringae_1A	PAYLALGVKGSDPRCYTSGTGPKATHFDLGGIYGPEIGEGGEGAYF-FTGDRGRDQIKR
TRVS_Trichoderma virens_6A	PAYLALGVKGPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFVYTGDGGRDQVKR
TRVS_Trichoderma virens_13A	PAYLALGIMGPDPRSVTSGTGPKGMHFDLGGGYGPEIGPGGEGGYW-YTGDGGRDQVKH
KATS_Kallichroma tethys_3A	PAYIAIKLKGPDPRAYTSGTGPKGVHFDIGGNYGIELCPNGEGGYFPYTGDRGRDQVKR
Syringopeptin_Ps...onas syringae_3A	PAYLALGVKAPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFPYTGDGGRDQVKR
Syringopeptin_Ps...onas syringae_4A	PAYLALGVKAPDPRAYTSGTGPKGVHFDLGGNYGPEIGAGGEGGYFPYTGDGGRDQVKR
Actinomycin_Stre...yces anulatus_3A	PAYLALGVKGPDPRAYTSGTGPKGVHFDLGGDYGPEIGAGGEGGYFPYTGDDGRDQVKR
GraS2_Aneurinibacillus migulanus_2A	PAYLALGVKGPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYFPYTGDGGRDQVKR
Vlm1_Streptomyces microflavus_2A	PGYLALGVRGPRPRAYTSGTGPKGVAFDLGGNYGPEIGPGGEGGYFPYTGDLGRDQVKR
TRVS_Trichoderma virens Gv29-8_4A	PAYLAIGVKGPDPRAYTSGTGPKGVHFDLGGNWGPEIGPGGEGEYSPYSGDGSRDQVKR
TRVS_Trichoderm...rens Gv29-8_18A	PAYLALGVKGPDPRAYTSGTGPKGVHFDLGGNYGPEIGPGGEGGYF-YTGDGGRDQVKR
sidC_Aspergillus nidulans_3A	PAYLALRVKGPDPRSYTSGTGPKGVRFDWGGNYGPEIGLGGEGGYFPYTGDDGRDQTKR

注：A-A 结构域集序列比对结果；B-A 结构域集序列标识图。

图 S1　A 结构域集序列保守性分析结果

Fig.S1　The result of conserved analysis of A-domain set sequences

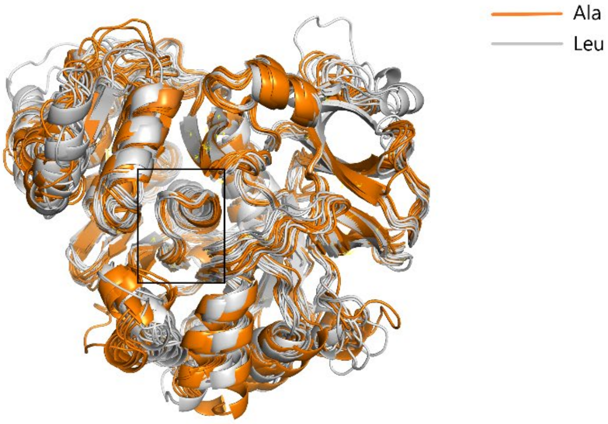


图 S2　特异性识别 Ala 的 A 结构域与特异性识别 Leu 的 A 结构域三级结构比对图

Fig. S2　Comparison map between A domain tertiary structure of Ala with A domain for specific identification of Leu