

综述

# 病原菌 TAL 效应子与寄主靶基因相互识别的分子密码

李岩强, 王春连, 赵开军

中国农业科学院作物科学研究所 农业部作物遗传育种重点实验室 农作物基因资源与基因改良国家重大科学工程, 北京 100081

**摘要:** 黄单胞杆菌属 TAL 效应子类蛋白作为病原菌的毒性因子或无毒因子, 能够与寄主靶基因 DNA 的启动子进行特异性识别, 调控寄主的基因表达, 引起致病或抗病反应。TAL 效应子类蛋白识别靶基因 DNA 的模式, 是 2 个氨基酸决定 1 个核苷酸的识别。这种新型的蛋白质-DNA 互作方式有可能在基因治疗、植物抗病基因发掘、广谱抗病基因构建等生物医学工程和农业工程方面得到广泛应用。文中综述了 TAL 效应子类蛋白的发现及功能, TAL 效应子与寄主靶基因识别的专一性及分子密码, 并对该分子密码当前的应用现状及前景进行了讨论和展望。

**关键词:** TAL 效应子, 分子识别密码, 蛋白质-DNA 互作, 基因治疗, 抗病育种

## Molecular recognition code between pathogenic bacterial TAL-effectors and host target genes: a review

Yanqiang Li, Chunlian Wang, and Kaijun Zhao

Key Laboratory of Crop Genetics and Breeding, Ministry of Agriculture, National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China

**Abstract:** As the pathogenic bacterial virulence and avirulence factors, transcription activator like (TAL) effectors of *Xanthomonas* can result in the host diseases or resistance responses. TAL effectors can specifically bind the target DNA of host plant with a novel protein-DNA binding pattern in which two amino acids recognize one nucleotide. The complexities of TAL-DNA binding have the feasibility in use of gene therapy through homologous recombination and site-specific mutation. By using the molecular recognition code between TAL-effectors and host target genes, we can exploit both the susceptible and resistance genes; broad spectrum resistance induced by multiple TAL effectors could also be manipulated. Deeper insight in the area of protein-DNA binding mechanism will benefit the application in the biomedical engineering and agricultural engineering. This article reviews the findings and functions of TAL effectors, the binding specificity and recognition code between TAL-effectors and host target genes. The possible applications and future prospects of the molecular recognition code have been discussed.

**Keywords:** TAL effector, molecular recognition code, protein-DNA binding, gene therapy, disease resistance breeding

**Received:** December 3, 2010; **Accepted:** March 14, 2011

**Supported by:** Major Project of China on New Varieties of GMO Cultivation (No. 2011ZX08001-002).

**Corresponding author:** Kaijun Zhao. Tel: +86-10-82105852; Fax: +86-10-82105819; E-mail: zhaokj@mail.caas.net.cn

转基因生物新品种培育科技重大专项 (No. 2011ZX08001-002) 资助。

黄单胞杆菌属 (*Xanthomonas*) 的多个致病变种, 在多种作物上造成严重的病害, 包括辣椒和番茄的细菌性斑点病、水稻的白叶枯病和细菌性条斑病<sup>[1]</sup>。TAL 效应子 (Transcription activator Like effector) 作为黄单胞杆菌的蛋白类效应子之一, 能够通过三型分泌系统 (Type III secretion system, TTSS) 进入植物的细胞核, 并与特定基因启动子 DNA 结合, 类似于真核生物转录因子, 启动植物基因的表达, 以控制植物的生理生化进程。病原菌通过进化产生一系列的 TAL 效应子以利于在寄主上的定殖和传播, 而植物亦进化出一系列的抗病策略以抑制该病原菌引起的病害。因此针对不同的寄主植物的基因型, TAL 效应子既有可能是毒性因子, 也有可能是无毒因子。近年来关于黄单胞杆菌的蛋白类效应子 TAL 效应子的研究有许多新进展, 尤其是 TAL 效应子与寄主 DNA 特异性识别的分子密码的破解, 将在生物工程领域和植物抗病育种应用上产生重要的影响。本文在介绍 TAL 效应子类蛋白的发现及功能、TAL 效应子与寄主靶基因识别的专一性及其分子密码的基础上, 讨论了该分子密码的应用前景。

## 1 TAL 效应子的发现及研究现状

### 1.1 TAL 效应子的发现及功能

第一个 TAL 效应子 (AvrBs3) 于 1989 年发现于辣椒斑点病细菌 *Xanthomonas campestris* pv. *vesicatoria* (xcv) 中, 其编码基因 *avrBs3* 的中心区有 17.5 个几乎相同的 102 bp 碱基重复, 每个重复编码 34 个氨基酸, 全基因共编码一个 1 164 aa 的蛋白质 (122 kDa)<sup>[2]</sup>。此后发现的 TAL 效应子类蛋白被称为 AvrBs3 家族蛋白。TAL 效应子类家族蛋白具有结构的相似性 (图 1A, B): 1) N 端高度保守, 其 mRNA 水平上存在着 TTSS 分泌信号 (T3S); 2) C 端含有亮氨酸拉链结构 (Leucine zipper, LZ)、核定位信号 (Nuclear localization signals, NLSs) 和酸性转录激活域 (Acidic activation domain, AD); 3) 中间为多

个串联重复单元组成的串联重复区, 每个重复单元含有 34 个氨基酸, 其中第 12 和 13 位氨基酸高度可变, 被称为重复可变区 (Repeat-variable diresidue, RVD)<sup>[3]</sup>。研究表明 NLS 具有核定位功能, 而保守的 LZ 和 AD 结构都是引起过敏反应 (Hypersensitive reaction, HR) 所必需的。所有的 TAL 效应子都具有相同的 NLS 及 AD, 只是其重复可变区的重复数目和重复类型不同, 并决定 TAL 效应子与寄主识别的特异性。

TAL 效应子对于寄主植物, 既是无毒因子, 又是毒性因子。从最早克隆的 *avrBs3* 基因到最近克隆的 *avrXa7*、*avrXa10*、*AvrXa27*, 分别能使含相应抗病基因的寄主引起 HR 反应, 从而引起寄主的抗病反应, 表现出无毒效应子功能。而对于不含相应抗病基因的寄主, 这些无毒蛋白就表现为毒性因子, 能够引起寄主植物发病。

### 1.2 已经克隆的 TAL 效应子编码基因及其寄主靶基因

TAL 效应子类蛋白广泛存在于黄单胞杆菌属 *Xanthomonas* 和少数的青枯拉尔氏菌属 *Ralstonia solanacearum* 中<sup>[4]</sup>。水稻黄单胞杆菌白叶枯病致病变种 (*Xanthomonas oryzae* pv. *oryzae*) 引起的水稻白叶枯病 (Bacterial blight, BB) 是一种典型的“基因对基因”病害, 植物病原细菌的无毒基因 (Avirulence gene, *avr*) 和寄主植物抗病基因 (Resistance gene, R) 的显性互作决定不亲和反应<sup>[5]</sup>。李平等 2004 年对黄单胞菌无毒基因的研究进展作了综述<sup>[6]</sup>。李玉蓉等于 2007 年对水稻黄单胞杆菌白叶枯病无毒新基因的克隆以及细菌性条斑病 (Bacterial leaf streak, BLS) 的病原菌 (*Xanthomonas oryzae* pv. *oryzicola*) 无毒基因的研究进展进行了综述<sup>[7]</sup>。但随着黄单胞杆菌属许多生理小种的测序<sup>[8-12]</sup>与进一步研究, 该领域的进展日新月异。目前, 已有更多的 TAL 效应子编码基因及其寄主靶基因被克隆。例如, 我们通过构建 PXO99 的 Tn5 转座子插入的突变体库, 筛选使近等基因系 CBB23 致病的菌株<sup>[13]</sup>, 并克隆了水

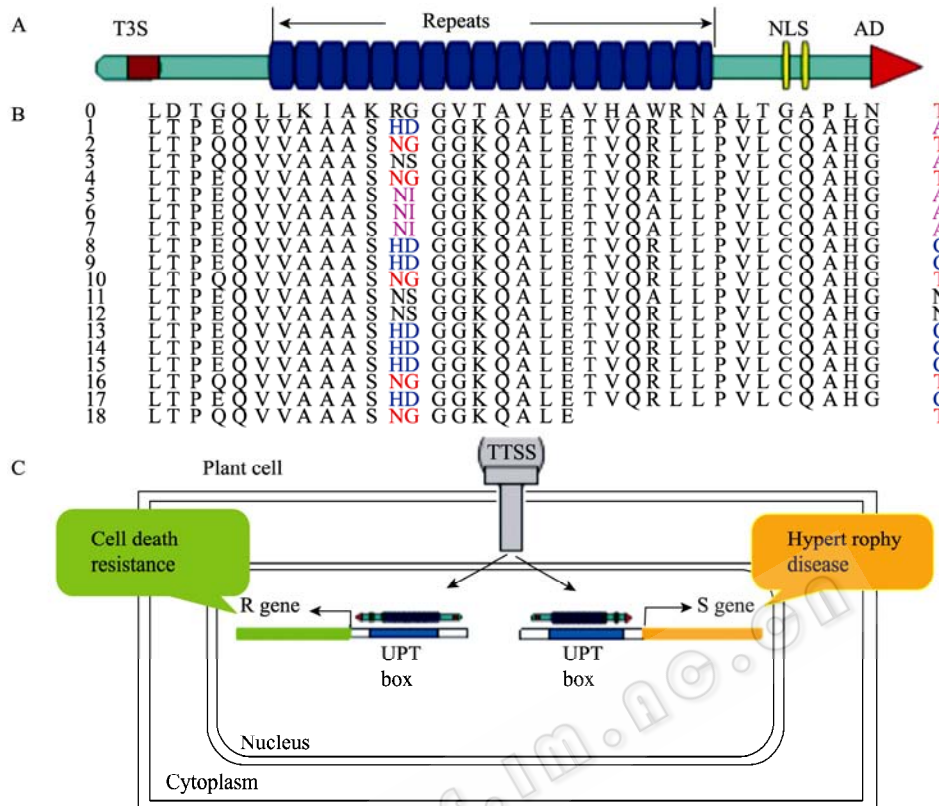


图 1 以 AvrBs3 为例说明 TAL 效应子的结构和功能

Fig. 1 Structure and function of the TAL effector AvrBs3. (A) The TAL effector AvrBs3 contains central tandem repeats (Repeat domain), the type III signal at N-terminal in mRNA level; the nuclear localization signals (NLSs) and an acidic transcriptional activation domain (AD) at C-terminal. (B) The amino acid sequences of the repeat domain of AvrBs3. The amino acids 12 and 13 are hypervariable, the repeats are numbered in the right and the left is the corresponding nucleotide. Because the all known UPT box has an invariable 5' T, so the 0-position is the nucleotide T. (C) Using the nuclear localization signals, the TAL effectors translocate into the plant cell through type III secretion system, bind the UPT Box of target genes: the R genes or S genes, respectively result in the resistance such as HR response and disease symptom like hypertrophy.

稻 *Xa23* 基因对应的无毒基因 *avrXa23*, 发现 AvrXa23 蛋白是一个 TAL 效应子, 它被 *Xa23* 基因的启动子陷阱捕获而激发水稻产生强烈的抗病反应。White 等于 2009 年将已鉴定的 TAL 效应子及其寄主靶基因的信息进行了列表综述<sup>[14]</sup>, 由于该领域进展很快, 笔者在 White 文章基础上将国内外已克隆鉴定的黄单胞菌 TAL 效应子及其相关信息的最新进展整理成表 1。

绝大多数的 TAL 效应子对应的寄主靶基因仍然未知, 研究最清楚的 *avrBs3*, 其寄主靶基因为 *Bs3*、*UPA20* 等; AvrBs3 被注入到植物的细胞核中, 其中

抗病基因 *Bs3* 启动子区存在能够特异性识别该无毒基因的元件, 受其诱导表达, 产生抗性<sup>[15]</sup>。感病品种中受 AvrBs3 诱导的 *UPA20* 等是感病相关的基因, *UPA20* 是一个转录因子, 能够诱导下游基因 *UPA7* 的表达, 进而引起寄主细胞肥大, 有利于病原菌的侵染和定殖<sup>[16]</sup>。AvrBs4 的寄主靶基因是 *Bs4*, 与其他 TAL 效应子不同的是, AvrBs4 是在植物细胞质而不是细胞核中与 *Bs4* 识别并产生抗病的<sup>[17]</sup>。其他研究比较明确的主要为水稻白叶枯病细菌的 TAL 效应子及其对应的寄主靶基因, 如 *avrXa3/Xa3*、*avrXa5/Xa5*、*avrXa7/Xa7* 和 *Os11N3*、*avrXa10/Xa10*、

表 1 已鉴定的 TAL 效应子

Table 1 List of identified TAL effectors

Effector	GenBank Accession No.	Repeats	Original species	Host target genes (host species)
Apl1	BAA37119	17.5	<i>Xac</i> , NA-1	Unkown (citrus)
Apl2	BAA37120	15.5	<i>Xac</i> , NA-1	Unkown (citrus)
Apl3	AB021365	23.5	<i>Xac</i> , NA-1	Unkown(citrus)
AvrBs3	P14727	17.5	<i>Xcv</i> , 71-21	<i>Bs3</i> (pepper), <i>upa20</i> (pepper)
AvrBs4	CAA48680	17.5	<i>Xcv</i> , 82-8	<i>Bs4</i> (tomoto)
AvrHah1	ABP97430	13.5	<i>Xg</i> XV444	<i>Bs3</i> (pepper)
AvrXa3	AY129298	8.5	<i>Xoo</i> , PXO86	<i>Xa3</i> (rice)
Avrxa5	AY377126	19.5	<i>Xoo</i> , JXOIII	<i>xa5</i> (rice)
AvrXa7	AAF98343	25.5	<i>Xoo</i> , PXO86	<i>Xa7</i> (rice), <i>Os11N3</i> (rice)
AvrXa10	AAA92974	15.5	<i>Xoo</i> , PXO86	<i>Xa10</i> (rice)
AvrXa23	unpublished	26.5	<i>Xoo</i> , PXO99A	<i>Xa23</i> (rice)
AvrXa27	AA54168	16.5	<i>Xoo</i> , PXO99A	<i>Xa27</i> (rice)
PthXo1(tal2b)	ACD58243	23.5	<i>Xoo</i> , PXO99A	<i>Os8N3</i> (rice)= <i>Xa13</i> (rice)
PthXo2	AAS46026	21.5	<i>Xoo</i> , PXO71	Unkown (rice)
PthXo3	AAS46027	28.5	<i>Xoo</i> , PXO61	<i>Os11N3</i> (rice)
PthXo6	ACD58920	23.5	<i>Xoo</i> , PXO99A	<i>OsTFX1</i> (rice)
PthXo7(tal1)	ABB70129	21.5	<i>Xoo</i> , PXO99A	Unkown (rice)
PthXo8	ACD60557	19.5	<i>Xoo</i> , PXO99A	Unkown (rice)
PthA	AAC43587	17.5	<i>Xac</i> , Xc270	Unkown (rice)
PthA*	ABO77780	17.5	<i>Xac</i> , Xc270	Unkown (rice)
PthAW	ABO77779	17.5	<i>Xac</i> , X0053	Unkown (rice)
PthB	NP_942641	17.5	<i>Xcau</i> , B96	Unkown (rice)
PthC	ABO77782	17.5	<i>Xcau</i> , C340	Unkown (rice)
PthN	AAB69865	13.5	<i>Xcm</i> , XcmN	Unkown (rice)
Avr/Pth3	AAW59492	17.5	<i>Xoc</i> , RS105	Unkown (rice)
Avr/Pth13	AAW59491	5.5	<i>Xoc</i> , RS105	Unkown (rice)
Avr/Pth14	AAW59493	19.5	<i>Xoc</i> , RS105	Unkown (rice)
Tal2a	YP_001912778	14.5	<i>Xoo</i> , PXO99A	Unkown (rice)
Tal4	YP_001913182	22.5	<i>Xoo</i> , PXO99A	Unkown (rice)
Tal5a	YP_001913457	15.5	<i>Xoo</i> , PXO99A	Unkown (rice)
Tal6a	YP_001913480	19.5	<i>Xoo</i> , PXO99A	Unkown (rice)
Tal7a	YP_001913755	17.5	<i>Xoo</i> , PXO99A	Unkown (rice)
Tal7b	YP_001913759	19.5	<i>Xoo</i> , PXO99A	Unkown (rice)
Tal8a	YP_001913948	17.5 Duplicate of Tal7a	<i>Xoo</i> , PXO99A	Unkown (rice)
Tal8b	YP_001913952	19.5 Duplicate of Tal7b	<i>Xoo</i> , PXO99A	Unkown (rice)
Tal9a	YP_001915089	19.5	<i>Xoo</i> , PXO99A	<i>OsHEN1</i> (rice)
Tal9b	YP_001915093	26.5	<i>Xoo</i> , PXO99A	Unkown (rice)
Tal9d	YP_001915101	12.5	<i>Xoo</i> , PXO99A	Unkown (rice)
Tal9e	YP_001915105	23.5	<i>Xoo</i> , PXO99A	Unkown (rice)
PthC8a	ABY60855	20.5	<i>Xoo</i> , C8	Unkown (rice)
PthC8b	ACD11364	17.5	<i>Xoo</i> , C8	Unkown (rice)

*avrXa23/Xa23*、*avrXa27/Xa27*、*pthXo1/Os8N3*、*pthXo3/Os11N3*、*pthXo6/OsTFX1*、*Tal9a/OsHEN1* 等 (表 1)。抗白叶枯病基因 *Xa27* 能够专一性地受含无毒蛋白 *avrXa27* 菌的诱导, 引起植物的抗病, 而其等位基因 *xa27* 由于启动子区的缺失和突变, 不能受 *avrXa27* 的诱导, 进而发生感病<sup>[18]</sup>。*PthXo1* 是一个毒性因子, 它能够诱导水稻中感病基因 *Os8N3* 的表达, 引起植物感病, 对该感病基因进行沉默可以使水稻对含 *PthXo1* 的菌株产生抗性, *Os8N3* 是 MtN3 家族中的一员, 对水稻的花粉发育有促进作用, 但是 *Os8N3* 的生化功能还未知<sup>[19]</sup>。而隐性抗白叶枯病基因 *xa13* 是 *Os8N3* 基因启动子区发生了变化而不能受 *PthXo1* 的诱导表达<sup>[20]</sup>。目前鉴定的 TAL 效应子类蛋白绝大多数为毒性因子, 少数的为无毒因子, 而目前发现的大部分寄主靶基因为抗病基因, 也有少数的为感病基因如: *Os11N3*、*Os8N3*、*OsTFX1* 及 *OsHEN1* 等。当前研究比较清楚的 TAL 效应子都与靶基因启动子识别进而引起抗病或感病反应, 而仍然有许多 TAL 效应子与靶基因的作用机制仍不明朗, 需要进一步的研究证实。

## 2 TAL 效应子与寄主靶基因识别的专一性

含有 *AvrBs3* 蛋白的黄单胞杆菌能够引起含 *Bs3* 基因的辣椒品种 ECW-30R 的抗病反应, 即产生 HR (图 1C), 对 *avrBs3* 的重复区进行缺失突变后, 发现缺失突变体 *avrBs3Δrep* 不能引起 ECW-30R 产生抗病反应<sup>[2]</sup>。但其中的一些突变体如 *avrBs3Δrep16* (缺失 11~14 之间的重复) 能够引起另一个辣椒品种 ECW 产生 HR 反应。后来的研究发现 *AvrBs3Δrep16* 能够特异性识别 *Bs3* 等位基因 *Bs3-E* 的启动子<sup>[15]</sup>。*AvrXa7* 和 *AvrXa10* 分别含有 25.5 和 15.5 个重复单元, 若它们的重复区互置, 所识别的抗病基因 *Xa7* 和 *Xa10* 也发生相应的变化, 即含有 *AvrXa10* 重复区的 *AvrXa7* 可以识别 *Xa10*, 含有 *AvrXa7* 重复区的 *AvrXa10* 可以识别 *Xa7*<sup>[21]</sup>。水稻抗白叶枯病基因

*Xa27* 专一性地受白叶枯病原菌中无毒蛋白 *AvrXa27* 的诱导表达<sup>[18]</sup>, 感病基因 *Os8N3* (*Xa13*) 受 TAL 效应子 *PthXo1* 的诱导<sup>[19]</sup>, 而当 *Os8N3* 启动子区发生突变, 可以逃避 *PthXo1* 的识别, *xa13* 抗病基因便因此获得对白叶枯病的抗性<sup>[20]</sup>。

深入的研究发现, TAL 效应子与寄主靶基因的特异性识别发生在启动子区域。*AvrBs3* 和 *AvrBs3Δrep16* 分别特异地与 *Bs3* 及 *Bs3-E* 的启动子区的特定序列互作<sup>[22-23]</sup>, Römer 等将这些识别序列称为 UPT((UP regulated by TAL effectors)Box。缺失突变分析和体外的 EMSA 实验表明: *AvrBs3* 和 *AvrBs3Δrep16* 分别与 UPT<sub>AvrBs3</sub> box 和 UPT<sub>AvrBs3Δrep16</sub> box 专一性识别<sup>[22]</sup>。水稻抗白叶枯病基因 *Xa27* 受无毒蛋白 *AvrXa27* 的诱导表达, 序列分析发现, *Xa27* 与其等位基因 *xa27* 的编码区没有明显的差异, 只是在启动子区存在碱基的缺失和变异<sup>[18]</sup>。对 *Xa27* 的启动子进行分析, 并将靠近 TATA box 区的突变位点的相关片段导入到 *xa27* 的启动子中, 发现融合后的 *xa27* 启动子能够受 *AvrXa27* 的诱导, 利用缺失突变及 EMSA 实验最终确定了最短的 UPT<sub>avrXa27</sub> box 序列<sup>[23]</sup>。TAL 效应子不仅能够识别抗病基因的启动子, 它们的靶基因还可能是感病相关的基因。受 *AvrBs3* 诱导的 *Upa20* 基因能够引起植物生理状态上的变化, 例如细胞分裂以及细胞增大, 以利于病原菌在寄主中的定殖和生长 (图 2C)<sup>[16]</sup>。受无毒因子诱导的另外一些典型的感病基因如 *Os8N3* (*Xa13*)、*OsTFX1* 以及 *Os11N3* 等分别受 *PthXo1*、*PthXo6* 及 *AvrXa7* 的特异性诱导<sup>[19,24-26]</sup>。同样地, 利用缺失突变及 EMSA 实验分析的方法, 相应的 UPT Box: UPT<sub>PthXo1</sub>Box、UPT<sub>PthXo6</sub>Box 以及 UPT<sub>AvrXa7</sub>Box 得到确认, 根据 UPT box 单个碱基的突变分析表明, 特定的 TAL 效应子, 专一性诱导特定的 UPT Box<sup>[27]</sup>。上述实验表明, 特定的 TAL 效应子识别相应的 UPT box, 这也是为什么许多 R 基因 (*Bs3*、*Xa27*、*Xa7*、*Xa10*) 仅受特定的 *AvrBs3* 类无毒蛋白诱导的原因, 这从分子水平上验证了“基因对基因”学说。

### 3 TAL 效应子与寄主靶基因识别的分子密码

AvrBs3 直接识别被其诱导的靶基因启动子区的 UPT Box<sup>[22,28]</sup>, 引起受 AvrBs3 诱导的许多基因如 UPA1-11<sup>[29]</sup>、UPA14-25<sup>[30]</sup>的表达, 导致感病植株的叶肉细胞肥大 (图 1C)。Kay 等分析受 AvrBs3 诱导的基因 *Bs3*、*UPA10*、*UPA12*、*UPA14*、*UPA19*、*UPA20*、*UPA21*、*UPA23* 及 *UPA25* 启动子转录起始点上游 100 bp 序列, 发现这些受 AvrBs3 诱导的基因含有相同的 UPA motif 基序, 有 19 个核苷酸组成的 box 具有极强的保守性, 与以前鉴定的 AvrBs3 的 UPT Box 一致<sup>[30]</sup>。由于 UPT Box [18 bp<sup>[29]</sup>和 19 bp<sup>[30]</sup>] 的碱基数与 AvrBs3 的重复数 (17.5) 基本一致, Boch 等推测 AvrBs3 的每一个重复单元识别一个特异的 DNA 碱基对<sup>[31]</sup>。当每一个 AvrBs3 的重复类型 (每个重复区的第 12 和第 13 氨基酸) 对应到 UPT Box 中去时, 发现特定的重复类型与目标 DNA 上的特异碱基相关联 (图 1B)。例如 HD 和 NI 分别对 C 和 A 有很强的偏爱, 而 AvrBs3 $\Delta$ rep16 (缺失 11~14 之间的重复), 能够识别更短的或者 (DNA Box 的 3' 端) 不同的目标 DNA 序列<sup>[31]</sup>。另外, 在 UPT Box 的 5' 末端, 对应于第一个预测的重复碱基的前一个位点上含有一个保守的 T<sup>[31]</sup>。后来的研究发现保守的 T 碱基对于诱导活性至关重要<sup>[27]</sup>。Boch 等通过分析多个无毒蛋白重复类型以及受其诱导的基因启动子区的 UPT Box 的对应关系, 构建了不同重复类型特异性识别目标 DNA 碱基的分子密码模型<sup>[31]</sup>。与此同时, Moscou 和 Bogdanove 利用计算机扫描比对 TAL 效应子中可变重复类型与目标启动子中的 DNA 序列, 统计分析得到了类似的分子密码模型 (图 1B)<sup>[32]</sup>。

为了验证病原菌无毒蛋白与靶基因 DNA 特异性识别的分子密码模型, Boch 等分析受 AvrXa27、PthXo1、PthXo6 及 PthXo7 诱导的基因及其等位基因的启动子区, 发现受诱导的基因中均含有相匹配

的 UPT Box, 而不受诱导的等位基因找不到相应的 UPT Box<sup>[31]</sup>。他们将预测的 UPT Box 构建到最小 Bs4(-55~+25) 启动子的上游, 驱动 *GUS* 基因, 与含有相对应 35S 启动子驱动的 TAL 效应子载体的农杆菌共注射本茗烟, 证明预测的 UPT Box 受相应的 TAL 效应子诱导。他们还根据已知的密码模型, 成功预测了 Hax2 的靶基因 PAPI(At1G56650)。此外, Boch 等构建了 7 个人造的 TAL 效应子 (ArtX), 将它们构建成 N 端融合了 *GFP* 报告基因的人造效应子, 并检验其诱导含有预测的 UPT box 的 Bs4 启动子驱动报告基因表达的能力。这 7 个人造效应子都只激发含相应的 UPT Box 的启动子引起 *GUS* 表达活性, 且人造的 ArtX 和 GFP 融合蛋白能够在植物细胞中表达<sup>[31]</sup>。

### 4 TAL 效应子与寄主靶基因识别密码的应用前景

#### 4.1 TALN 的产生及应用前景

TAL 效应子与寄主靶基因的互作是一种新型的蛋白-DNA 结合方式, 即 2 个特异的氨基酸组合对应 1 个特异的碱基。根据 TAL 效应子与寄主靶基因相互识别的分子密码, 科学家已将 TAL 效应子与核酸内切酶 *Fok I* 融合构建了新型的位点特异性识别的工具酶 TALNs (TAL effector nucleases), 并通过体外实验, 对识别的特异性进行了验证<sup>[33-34]</sup>。Li 等构建了 AvrXa7 和 PthXo1 与核酸内切酶 *Fok I* 融合的酶, 通过体外的 DNA 酶切实验以及酵母内的重组实验均证明, TALN 能够特异性地识别其目标 DNA 序列, 并使 *Fok I* 在特定的位点进行酶切<sup>[33]</sup>。Christian 等先是利用 AvrBs3 及 PthXo1 和核酸内切酶 *Fok I* 融合的酶, 发现了不同的融合酶对于酶切位点的活性受识别位点的距离的影响, 他们还证明使用 2 个不同的融合酶能够分别识别特异的位点并共同完成酶切。同时, 为了验证该分子密码, 他们找到了基因组上的序列, 并按照频率最大的密码 (NI 对应 A,

HD 对应 C, NN 对应 G 以及 NG 对应 T) 人工设计了相对应的 TAL 效应子, 通过酵母内的实验发现许多人工设计的融合酶具有很强的活性 (图 2A)<sup>[34]</sup>。

这些实验证明了利用分子密码设计相应的 TALN 完成真核生物基因组的重组和突变成为可能。例如可以设计能与特定 DNA 序列结合的人造效应子, 以完成靶基因的定点突变和重组。另外, 这种转录因子模式可以控制目的基因转录水平上的表达, 如利用特异识别启动子的人造 TAL 效应子作为基因表达的分子开关, 在转录水平上进行基因的表达调控。当然, 这种机制不仅仅用于植物, 也可能应用于人类重要疾病比如癌症等的基因治疗。我们预计, TALN 将在未来的基因治疗中受到重视并发挥重要作用。

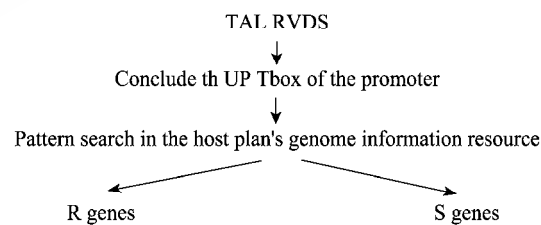
#### 4.2 利用分子密码发掘更多的抗病基因和感病基因

根据分子识别密码, 可以利用 TAL 效应子的可变区序列预测并寻找抗病基因或感病基因 (图 2B)。比如水稻基因组序列已经完成, 可以利用 TAL 效应子可变区序列, 扫描基因组序列, 找到受其诱导的启动子, 进而研究相应基因的功能。这可能为寻找抗病基因开辟一条新的途径。如 Boch 等根据 TAL 效应子 HAX2 的序列, 推测出其对应启动子的序列, 并在拟南芥基因组中找到了靶基因 MYB 转录因子类的 PAP1 (At1G56650)<sup>[31]</sup>。现已发现多个毒性因子 TAL 效应子, 这样可以根据其可变区 (RVD) 的变化, 推测出与感病相关的基因, 研究植物感病的分子机制, 并利用 RNAi 的办法抑制感病基因从而达到提高抗病性的目的。

##### A Construct the hybrid TAL nucleases with designed endonuclease site



##### B Identify the resistance related genes



##### C Create broader spectrum resistance gene



图 2 TAL effector 与靶基因识别的分子密码的应用前景

Fig. 2 Applications and future prospects of molecular recognition code between TAL effectors and host genes. (A) Construction of hybrid TAL nucleases (TALN). This chimeric protein can bind the target nucleotides and cut at the relative position using the *Fok I* digestion function. This diagram takes the *AvrXa7* and *PthXo1* used in Li's paper<sup>[33]</sup> as an example. (B) Identify the resistance related genes. The known effectors can be used to identify their target genes through pattern search in the host plant's information resource. In this way, more and more resistance or susceptible genes can be identified in the future. (C) Create broader spectrum resistance gene. Multiple UPT-Boxes can be combined in one promoter to drive a R gene like *Rxo1* or *Xa27*. This new engineered R gene complex can be recognized and induced by different TAL effectors, resulted in resistance to more pathogenic bacteriums.

### 4.3 利用分子密码开发广谱抗病基因启动子

Römer 等利用基因工程的方法将不同的 UPT Box 聚合到一个启动子上, 获得受多个 TAL 效应子诱导的新启动子<sup>[27]</sup>。这种新型的启动子使利用基因工程扩展抗病基因的抗谱成为可能。例如, 在水稻白叶枯病抗病育种中, 选育和种植抗病品种是防治该病的最经济、有效和环保的途径。目前国内外报道的白叶枯病抗病基因已经有 36 个之多<sup>[35]</sup>, 但是由于水稻白叶枯病原菌致病性的变异和新毒性菌群的不断出现, 一些抗病品种的抗性逐渐丧失<sup>[36]</sup>。根据“基因对基因”理论, 抗病基因是和相应的无毒基因互作后起作用的, 介导的抗性是小种特异性的, 向植物中导入某一抗病基因后, 它可以对携带特定无毒基因的病原菌产生抗性。目前克隆得到的抗白叶枯病基因的无毒基因很多, 可以根据分子密码设计受多种无毒基因诱导的 UPT Box<sup>[27]</sup>。将新设计的含多个 UPT Box 的启动子驱动一个抗病基因, 使该抗病基因受多个无毒基因产物的诱导, 进而扩大抗病基因的抗谱 (图 2C)。另外, 目前在细菌性条斑病细菌 (*Xooc*) 中也克隆到一些 TAL 效应子类基因<sup>[37]</sup>。但是抗细菌性条斑病的主效基因报道仍然很少<sup>[38]</sup>, 目前比较有效的为非寄主抗性基因 *Rxo1*<sup>[39]</sup>, 而组成型表达的抗白叶枯病基因 *Xa27* 也可以抗一些细菌性条斑病小种<sup>[40]</sup>。可以利用 TAL 效应子与寄主靶基因 DNA 识别的分子密码设计受多个细菌性条斑病细菌 TAL 效应子诱导的启动子, 驱动抗病基因如 *Rxo1* 或 *Xa27*, 进而得到抗细菌性条斑病的抗病基因。目前我们拟利用已经克隆得到的 *avrXa23* 以及已经发表的无毒基因, 根据分子密码推测的 UPT Box 组合成一个受多个 TAL 效应子诱导的抗病基因启动子, 用于驱动抗病基因, 期望能够获得更大的抗谱。

### 4.4 问题及展望

TAL 效应子与其靶基因互作的研究取得了很大进展, 但是仍然有许多问题需要进一步揭示。例如 TAL 效应子到底如何与 DNA 碱基相结合, 包括结

合的方式及分子机制, 以及为什么发生特异性的结合, 这些都需要完成 TAL 效应子及 TAL-DNA 复合物的晶体结构解析才能准确地回答。TAL 效应子与寄主靶基因相互识别密码的广泛性还需要新的实验证明。对于分子密码的应用研究也只是刚刚开始, TALN 在基因组靶向修饰中的使用条件以及在各个物种中的通用性仍然需要进一步研究。受多种 TAL 效应子诱导的启动子已经构建成功, 但是构建的新型启动子能否获得更大的抗谱仍然需要实践证明。尽管如此, TAL 效应子与寄主靶基因识别的分子密码作为一种新型的识别机制, 具有很强的专一性, 利用该分子密码设计新的基因工程酶, 寻找更多的抗病基因, 构建更多抗谱的抗病基因将会有巨大的应用前景。继续研究黄单胞杆菌 TAL 效应子与其对应的寄主靶基因相互识别的分子机制, 并将其相互识别的分子密码应用于生物医学工程及农业抗病育种技术中去, 将为人类战胜疾病及获得更多的粮食提供更大的支持。

## REFERENCES

- [1] Boch J, Bonas U. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol*, 2010, 48: 419–436.
- [2] Bonas U, Stall RE, Stskawicz BJ. Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *campestris*. *Mol Gen Genet*, 1989, 218(1): 127–136.
- [3] Herbers K, Conrads-Strauch J, Bonas U. Race-specificity of plant resistance to bacterial spot disease determined by repetitive motifs in a bacterial avirulence protein. *Nature*, 1992, 356(6365): 172–174.
- [4] Salanoubat M, Genin S, Artiguenave F, et al. Genome sequence of the plant pathogen *Ralstonia solanacearum*. *Nature*, 2002, 415(6871): 497–502.
- [5] Flor H H. Host parasite interactions in flax rust—its genetics and other implications. *Phytopathology*, 1955, 45: 680–685.
- [6] Li P, Long JY, Zhang Y, et al. Avirulence genes from *Xanthomonas oryzae* pv. *oryzae*. *J Nanning Agric Univ*,



- 2004, 27(3): 119–124.
- 李平, 龙菊英, 张燕, 等. 水稻黄单胞细菌的无毒基因. 南京农业大学学报, 2004, 27(3): 119–124.
- [7] Li YR, Zou LF, Wu XM, et al. Current Progress in the Research On *avrBs3/PthA* Family Genes of *Xanthomonas oryzae*. *Sci Agric Sin*, 2007, 40(10): 2193–2199.
- 李玉蓉, 邹丽芳, 武晓敏, 等. 水稻黄单胞菌 *avrBs3/PthA* 家族基因研究进展. 中国农业科学, 2007, 40(10): 2193–2199.
- [8] da Silva ACR, Ferro JA, Reinach FC, et al. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature*, 2002, 417(6887): 459–463.
- [9] Lee BM, Park YJ, Park DS, et al. The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Res*, 2005, 33(2): 577–586.
- [10] Qian W, Jia Y, Ren SX, et al. Comparative and functional genomic analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv. *campestris*. *Genome Res*, 2005, 15(6): 757–767.
- [11] Salzberg SL, Sommer DD, Schatz MC, et al. Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99A. *BMC Genomics*, 2008, 9(1): 204–219.
- [12] Thieme F, Koebnik R., Bekel T, et al. Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria* revealed by the complete genome sequence. *J Bacteriol*, 2005, 187(21): 7254–7266.
- [13] Wang CL, Xu AB, Gao Y, et al. Generation and characterization of Tn5-tagged *Xanthomonas oryzae* pv. *oryzae* mutants that overcome *Xa23*-mediated resistance to bacterial blight of rice. *Eur J Plant Pathol*, 2009, 123(3): 343–351.
- [14] White FF, Potnis N, Jones JB, et al. The type III effectors of *Xanthomonas*. *Mol Plant Pathol*, 2009, 10(6): 749–766.
- [15] Schornack S, Meyer A, Römer P, et al. Gene -for-gene-mediated recognition of nuclear-targeted AvrBs3-like bacterial effector proteins. *J Plant Physiol*, 2006, 163: 256–272.
- [16] Kay S, Hahn S, Marois E, et al. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science*, 2007, 318(5850): 648–651.
- [17] Schornack S, Peter K, Bonas U, et al. Expression levels of *avrBs3*-like genes affect recognition specificity in tomato *Bs4* but not in pepper *Bs3* mediated perception. *Mol Plant Microbe Interact*, 2005, 18(11): 1215–1225.
- [18] Gu KY, Yang B, Tian DS, et al. R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature*, 2005, 435(7045): 1122–1125.
- [19] Yang B, Sugio A, White FF. *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. *Proc Natl Acad Sci USA*, 2006, 103(27): 10503–10508.
- [20] Chu ZH, Yuan M, Yao JL, et al. Promoter mutations of an essential gene for pollen development result in disease resistance in rice. *Gene Dev*, 2006, 20(10): 1250–1255.
- [21] Hopkins CM, White FF, Choi SH, et al. Identification of a family of avirulence genes from *Xanthomonas oryzae* pv. *oryzae*. *Mol Plant Microbe Interact*, 1992, 5: 451–459.
- [22] Römer P, Strauss T, Hahn S, et al. Recognition of AvrBs3-like proteins is mediated by specific binding to promoters of matching pepper *Bs3* alleles. *Plant Physiol*, 2009, 150(4): 1697–1712.
- [23] Römer P, Recht S, Lahaye T. A single plant resistance gene promoter engineered to recognize multiple TAL effectors from disparate pathogens. *Proc Natl Acad Sci USA*, 2009, 106(48): 20526–20531.
- [24] Yuan M, Chu ZH, Li XH, et al. Pathogen-induced expressional loss of function is the key factor in race-specific bacterial resistance conferred by a recessive *R* gene *xa13* in rice. *Plant Cell Physiol*, 2009, 50(5): 947–955.
- [25] Sugio A, Yang B, Zhu T, et al. Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes *OsTFIIAγ1* and *OsTFX1* during bacterial blight of rice. *Proc Natl Acad Sci USA*, 2007, 104(25): 10720–10725.
- [26] Antony G, Zhou JH, Huang S, et al. Rice *xa13* recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene *OsIIN3*. *Plant Cell*, 2010, 22(11): 3864–3876.
- [27] Römer P, Recht S, Strauß T, et al. Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. *New Phytol*, 2010, 187(4): 1048–1057.
- [28] Römer P, Hahn S, Jordan T, et al. Plant pathogen recognition mediated by promoter activation of the pepper *Bs3* resistance gene. *Science*, 2007, 318(5850): 645–648.

- [29] Marois E, Van den Ackerveken G, Bonas U. The *Xanthomonas* type III effector protein AvrBs3 modulates plant gene expression and induces cell hypertrophy in the susceptible host. *Mol Plant Microbe Interact*, 2002, 15(7): 637–646.
- [30] Kay S, Hahn S, Marois E, et al. Detailed analysis of the DNA recognition motifs of the *Xanthomonas* type III effectors AvrBs3 and AvrBs3Drep16. *Plant J*, 2009, 59(6): 859–871.
- [31] Boch J, Scholze H, Schornack S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, 2009, 326(5959): 1509–1512.
- [32] Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science*, 2009, 326(5959): 1501.
- [33] Li T, Huang S, Jiang ZW, et al. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and *Fok I* DNA-cleavage domain. *Nucleic Acids Res*, 2010, 39(1): 359–372.
- [34] Christian M, Cermak T, Doyle EL, et al. TAL effector nucleases create targeted DNA double-strand breaks. *Genetics*, 2010, 110: 120717.
- [35] Miao LL, Wang CL, Zheng CK, et al. Molecular mapping of a new gene for resistance to rice bacterial blight. *Sci Agric Sin*, 2010, 43(15): 3051–3058.  
苗丽丽, 王春连, 郑崇珂, 等. 水稻抗白叶枯病新基因的初步定位. *中国农业科学*, 2010, 43(5): 3051–3058.
- [36] Zeng LX, Zhu XY, Yang JY, et al. A new pathotype of *Xanthomonas oryzae* pv. *oryzae* was found and tested for pathogenicity in Guangdong. *Guangdong Agric Sci*, 2005, 2: 58–59.  
曾列先, 朱小源, 杨建源, 等. 广东水稻白叶枯病菌新致病型的发现及致病性测定. *广东农业科学*, 2005, 2: 58–59.
- [37] Zou LF, Chen GY, Wu XM, et al. Cloning and analysis of diverse members of avrBs3/PthA Family of *Xanthomonas oryzae* pv. *oryzicola*. *Sci Agric Sin*, 2005, 38(5): 929–935.  
邹丽芳, 陈功友, 武晓敏, 等. 中国水稻条斑病细菌 avrBs3/PthA 家族基因的克隆和序列分析. *中国农业科学*, 2005, 38(5): 929–935.
- [38] Zeng JM, Lin WX. Research progress on rice bacterial leaf streak and its resistance. *Mol Plant Breed*, 2003, 1(2): 257–263.  
曾建敏, 林文雄. 水稻细菌性条斑病及其抗性研究进展. *分子植物育种*, 2003, 1(2): 257–263.
- [39] Zhao BY, Lin XH, Poland J, et al. A maize resistance gene functions against bacterial streak disease in rice. *Proc Natl Acad Sci USA*, 2005, 102(43): 15383–15388.
- [40] Tian DS, Yin ZC. Constitutive heterologous expression of *avrXa27* in rice containing the R gene *Xa27* confers enhanced resistance to compatible *Xanthomonas oryzae* strains. *Mol Plant Pathol*, 2009, 10(1): 29–39.