



## 副溶血弧菌 TF2 中整合接合元件 ICEVpaTF2 的生物信息学特征和剪切、环化活性

李颖颖<sup>1,2</sup>, 谢媚<sup>1,2</sup>, 程子骞<sup>3</sup>, 李卓波<sup>1,2</sup>, 罗鹏<sup>1\*</sup>

<sup>1</sup> 中国科学院南海海洋研究所, 中国科学院热带海洋生物资源与生态重点实验室, 广东省应用海洋生物学重点实验室, 广东 广州 510301

<sup>2</sup> 中国科学院大学, 北京 101408

<sup>3</sup> 河北农业大学海洋学院, 河北 沧州 060010

**摘要:**【目的】从副溶血弧菌 TF2 基因组框架序列中拼接获得整合性接合元件(ICE) ICEVpaTF2 的全序列, 分析其基因组学特征; 研究 ICEVpaTF2 是否具有从基因组中剪切、环化活性以及剪切、环化时 *attB* 和 *attP* 位点如何形成。【方法】对副溶血弧菌 TF2 基因组框架序列进行 RAST 注释, 发现其基因组中可能存在一个完整的 ICE 元件, 命名为 ICEVpaTF2, 经过人工拼接和 PCR 扩增测序验证, 得到完整的 ICEVpaTF2 序列, 并对 ICEVpaTF2 再次进行 RAST 注释和 ICE 元件部分特征分析。通过 PCR 检测, 探索 ICEVpaTF2 是否能够从基因组剪切并环化。通过 *attL*、*attR*、*attB*、*attP* 位点序列比较, 探索 *attB* 和 *attP* 重组特征。【结果】ICEVpaTF2 全长 83588 bp, 包含 SXT/R391 家族 ICE 元件的 52 个保守核心基因, 它们与 ICE 切除、整合、自我转移和调节机制相关。ICEVpaTF2 也包含 5 个外源基因插入的热点区(HS)、2 个可变区(VR)以及 3 个非典型插入位点。HS 和 VR 包含了大量可变基因, 它们负责编码限制性修饰系统、DNA 修复系统或毒素-抗毒素系统等, 赋予宿主广泛适应性功能, ICEVpaTF2 也包含独特未知功能基因。通过对 *int*、*xis* 二个核心保守基因种系发生分析, 发现 ICEVpaTF2 的 *int*、*xis* 分别属于以 R391 和 SXT 为代表的亚群。ICEVpaTF2 在 *attL* 和 *attR* 处发生剪切并完成环化, 新形成杂合重组的 *attB* 和 *attP* 位点。【结论】ICEVpaTF2 属于 SXT/R391 家族, 是一个具备自我剪切和环化能力的完整 ICE 元件, 剪切后新形成的 *attB* 和 *attP* 位点由 *attL*、*attR* 杂合重组形成。

**关键词:** 副溶血弧菌, SXT/R391 元件, ICEVpaTF2, 生物信息学特征, 剪切和环化

基金项目: 国家自然科学基金(31370149); 广东省渔业科技攻关项目(A201701B03)

\*通信作者。E-mail: luopengli@163.com

收稿日期: 2019-07-23; 修回日期: 2019-08-23; 网络出版日期: 2019-10-17

整合接合元件(integrating conjugative elements, ICEs)是一类可自我转移的遗传元件,它们既能像前噬菌体一样特异性整合入宿主染色体的特定位点,又能像接合质粒一样通过宿主与受体菌接合形成的通道实现细菌间的接合转移<sup>[1]</sup>。ICE水平转移对细菌的广泛环境适应性以及细菌的遗传多样性和基因组进化均产生重要影响<sup>[2-4]</sup>。ICE包含一系列保守基因组成的核心区以及位于核心区间的插入热点区(HS)和可变区(VR),位于HS和VR的基因赋予ICE独特的功能<sup>[5]</sup>。

SXT/R391家族是迄今为止发现最早、数量最多、研究最为深入的一类ICE<sup>[6-7]</sup>。SXT/R391家族编码高度保守的整合酶Int,Int是一种酪氨酸重组酶,介导SXT/R391元件的attP位点与宿主prfC基因(编码肽链释放因子3的保守染色体基因)内近5'端的位点(attB)处发生特异性整合,被认为是SXT/R391元件的一个决定性特征,整合后的线性化SXT/R391两端分别形成一个新的高度相似的attL和attR位点<sup>[2,6]</sup>。在重组相关因子Xis的辅助下,Int识别整合SXT/R391元件的最左端的attL位点和最右端的attR位点,并介导重组,使整个元件从染色体上切除,切除的元件环化形成新的attP位点,同时染色体上剪切处也重组形成新的attB位点<sup>[5-7]</sup>。被切除的环化SXT/R391是共轭转移的主要底物,当受体菌具有与其attP位点相似的attB'位点,SXT/R391就可能自主转移整合到受体菌相应的attB'位点,从而实现SXT/R391水平转移<sup>[3,5,7]</sup>。

SXT/R391家族的典型代表SXT元件最初发现于O139型霍乱弧菌中<sup>[8]</sup>。后来发现1972年分离于南非的雷氏普罗威登斯菌(*Providencia rettgeri*)携带一种与STX密切相关的可自主转移

的遗传元件,命名为R391<sup>[9]</sup>。此后,SXT/R391元件被发现广泛存在于亚洲和非洲地区临床和环境分离的霍乱弧菌和其他弧菌属细菌中<sup>[10]</sup>。此外,SXT/R391元件也存在于美人鱼发光杆菌(*Photobacterium damsela*)、腐败斯瓦尼菌(*Shewanella putrefaciens*)和奇异变形杆菌(*Proteus mirabilis*)的分离株中<sup>[10]</sup>。在人工实验条件下,发现原存在于弧菌中的SXT/R391元件能以极高的频率转移至大肠杆菌中<sup>[11-12]</sup>,表明SXT/R391元件具备在遗传关系遥远的细菌群体间转移的巨大潜力,因此具有进一步扩大分布群体范围的可能。

副溶血弧菌(*Vibrio parahaemolyticus*)普遍存在于海洋和河口环境,是三大对人类健康具有重大威胁的弧菌之一<sup>[13]</sup>,副溶血弧菌同时也是海水养殖动物的重要致病菌,可造成巨大的经济损失<sup>[14-15]</sup>。迄今为止,未有在副溶血弧菌中发现SXT/R391元件的报道,本研究从一株对虾致病菌——副溶血弧菌TF2的框架基因组序列中获得了一个完整的ICE元件ICEVpaTF2,从比较基因组学角度对此元件的基因组学特征进行了分析,对其剪切、环境化活性进行了检测,此外分析了att位点的重组特征。

## 1 材料和方法

### 1.1 菌株来源和全基因组测序

副溶血弧菌TF2分离于广东湛江养殖的患病凡纳滨对虾肝胰腺,并保存于本实验室。首先对副溶血弧菌TF2提取的基因组进行机械剪切,剪切后的DNA进行末端补平,再通过3'端加碱基A,使得DNA片段能与3'端带有T碱基的接头连接,用电泳回收目的片段(400-500 bp)连接产物,再使用PCR扩增两端带有接头的DNA片段,最后用

检验合格的文库进行 Illumina Miseq(美吉生物, 上海)上机测序。

## 1.2 副溶血弧菌 TF2 的基因组注释和 ICE 拼接

采用线上注释工具 Rapid Annotations using Subsystem Technology(RAST)<sup>[16]</sup>对 91 个框架序列进行注释并进行注释结果分析, 找出包含 SXT/R391 元件的片段的 Scaffold11 和 Scaffold7 序列。以 Scaffold11 末端序列(300 bp)进行 BLAST 搜索发现高度相似的 DNA 序列(GenBank No: MH160822), 以 MH160822 与 Scaffold11 末端序列高度匹配部分的侧翼序列(30 bp)为质询序列在 Scaffold 系列中搜索到包含 Scaffold7 侧翼未知序列的 Scaffold50。进一步对 Scaffold11 (5'末端包含 *attL* 序列)、Scaffold50、Scaffold7 (3'末端包含 *attR* 序列)进行了人工拼接, 得到一个完整的 ICE 序列, 该 ICE 命名为 ICEVpaTF2。设计跨 3 个 Scaffold 连接区的引物(表 1)。ICEVpaTF2-F2/R2 和 ICEVpaTF2-F3/R3 二对引物分别锚定 3 个 Scaffold 连接处的上下游, 进行 PCR 扩增和测序以验证拼接结果。获得完整 ICEVpaTF2 后再次以 RAST 对其基因注释。

## 1.3 ICEVpaTF2 的基因组学分析

以 SXT 序列(GenBank No: AY055428)以及

ICEVpaE0601 序列(GenBank No: KT072768)为参考, 对 ICEVpaTF2 中的保守基因进行识别, 对 ICEVpaTF2 中部分假想基因进行了保守结构域查找 (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>), 以补充完善基因的功能。通过 Blast 搜索与 ICEVpaTF2 全长最为相似的 ICE 序列。以 MEGA 6.0 对 ICE 整合和剪切的关键保守基因 *int* 和 *xis* 进行系统进化树构建和种系发生分析, 高度相似的序列来源于 BLAST 搜索并随机选取。

## 1.4 ICEVpaTF2 的剪切和环化活性检测

在 ICEVpaTF2 元件的起点(*attL*)两端和终点(*attR*)两端设计引物 *attL*-F1、*attL*-R1、*attL*-F2、*attL*-R2, 引物间进行不同的组合(表 1), 进行 PCR 扩增, 借以检测 ICEVpaTF2 是否存在整合、剪切环化的形式。

## 2 结果和分析

### 2.1 ICE 的组装

以副溶血弧菌 TF2 基因组 DNA 为模板, 以 ICEVpaTF2-F2/R2、ICEVpaTF2-F3/R3 引物对分别进行 PCR 扩增均得到预期大小的 PCR 片段, PCR 产物递交测序, 测序结果表明对 Scaffold11、

表 1. 引物和引物扩增的 DNA 区域

Table 1. Primers used for the amplification of different DNA regions in this study

Primer name	Nucleotide sequence (5'→3')	PCR amplification for
ICEVpaTF2-F2	AAGTTGAGTAACGGCAAGAGC	Gap between Scaffold 11 and Scaffold 50
ICEVpaTF2-R2	TATCTAGGATGACGGGCTTTG	
ICEVpaTF2-F3	CAACAAGAAGCTGCCAACACC	Gap between Scaffold 50 and Scaffold 7
ICEVpaTF2-R3	CGATGGTATCAATGTGGTGC	
<i>attL</i> -F1	AGCGCGTGGCTAGAAGTAAG	<i>attL</i> , <i>attR</i> , <i>attB</i> , and <i>attP</i> sites when using different primer combinations
<i>attL</i> -R1	ACACCGCCAGACTCGATTC	
<i>attR</i> -F4	ACAATACCCTGCAATACCGATC	
<i>attR</i> -R4	TACTTCCATTGGATCACGAACG	

Scaffold50、Scaffold7 的拼接完全正确，因此得到一个完整的 ICEVpaTF2 序列(GenBank No: MN201567)。

## 2.2 ICEVpaTF2 的一般结构

经生物信息学预测和分析，ICEVpaTF2 全长为 83588 bp，包含 78 个蛋白编码基因，由 52 个保守核心基因形成的骨架和 26 个可变基因组成。可变基因所在区域总长约 34.5 kb，占 ICE 元件全长的 41.3%。ICE 主要存在 5 个外源 DNA 频繁插入的区域，称为热点区域 HS1-HS5<sup>[17]</sup>；此外，在一些 ICE 中包含另外的可变区域，命名为 VR I-VR IV<sup>[17]</sup>。ICEVpaTF2 中可变基因主要分布在 5 个 HS、可变区 VR I 和 VR III 之间(图 1)，VR II 和 VR IV 没有基因插入。此外，有 3 个单基因分别出现在 HS 和 VR 以外的插入位点(图 1)。最大的可变区域插

入片段位于 HS5 区，长度达到 15.5 kb，该区包括 7 个编码基因。最小的可变区域 HS4 仅有一个基因插入。

## 2.3 ICEVpaTF2 核心保守基因

与其他 SXT/R391 元件相似，ICEVpaTF2 也注释到 3 个保守功能模块 C1-C3(图 1)，其中 C1 内的保守基因为 *int* 和 *xis*，负责 ICE 整合和切除<sup>[17]</sup>；C2 内的保守基因主要是一系列 *tra* 基因，主要负责单链 DNA 运输、DNA 加工以及 DNA 输出的交配孔道形成等<sup>[18-19]</sup>；C3 区的保守基因 *setCD* 和 *setR* 负责整个 ICE 整合、切除和转移活性的调控<sup>[20]</sup>。C1-C3 保守基因的排列顺序与 SXT/R391 家族元件完全相同。在预测到的 ICEVpaTF2 的 52 个保守核心基因中，有 32 个已经获知功能注释(图 1 中标注基因名称)，但是仍然

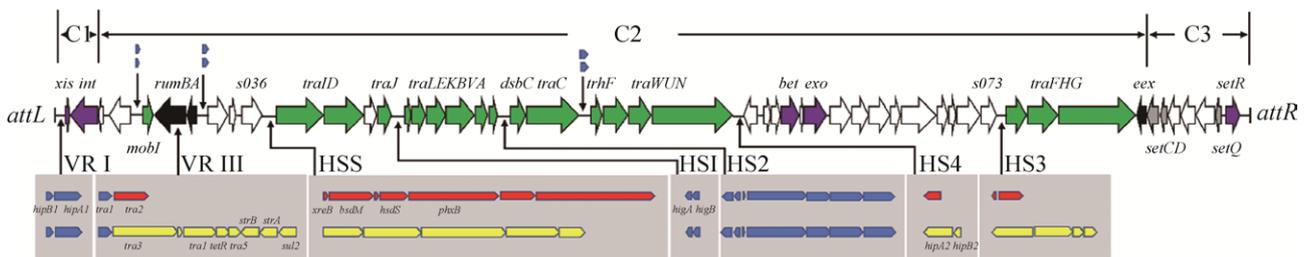


图 1. SXT/R391 家族 52 个核心基因示意图以及副溶血弧菌 TF2 的 ICEVpaTF2 与 *V. owensii* V180403 的 ICEVowV18 的结构比较

Figure 1. Schematic representation of 52 conserved core genes of SXT/R391 family and structural comparison between ICEVpaTF2 of *V. parahaemolyticus* TF2 and ICEVowV18 of *V. owensii* V180403. The thick arrows represent 52 core genes of SXT/R391 elements. The length of the arrows corresponds to the length of core genes of ICEVpaTF2. The short vertical lines at both sides of SXT/R391 elements represent the start (*attL*) and the end (*attR*) of them. The white thick arrows represent conserved genes with unknown functions. The purple thick arrows represent the genes closely related to prophage that are responsible for the integration, the excision, and the recombination of SXT/R391 ICEs. The green thick arrows represent the genes closely related to conjugative plasmids that are responsible for conjugation transfer of SXT/R391 ICEs. The upward vertical arrows indicate the insertion positions of the variable regions. The downward vertical arrows indicate three unusual insertion sites for variable gene. The red arrows indicate the specific genes in ICEVpaTF2 when compared to ICEVowV18, the yellow arrows indicate the specific genes in ICEVowV18, and the blue arrows indicate common genes of two ICEs in variable regions.

有 20 个保守基因功能未知(图 1 中白色粗箭头, 未标注基因名称)。

分别以 C1 保守模块内 *int* 和 *xis* 两个基因预测的核苷酸序列构建系统进化树(图 2-A 和图 2-B), 结果发现二者具有明显不同的拓扑学构型。基于 *int* 的种系进化树中, 来源于不同 ICE 的 *int* 基因明显分为 2 个分支, ICEVpaTF2 的 *int* 基因处在以 R391 为代表的分支上; 基于 *xis* 的种系进化树中, 来源于不同 ICE 的 *xis* 基因也明显分为 2 个分支, 但 ICEVpaTF2 的 *xis* 基因处在以 SXT 为代表的分支上。在采用相同比例的标尺(Bar)上, *xis* 分枝的平均长度大于 *int* 分枝的平均长度, 表明同处 C1 保守模块的 *int*、*xis* 虽然功能密切相关, 位置紧邻, 但可能具有明显不同的进化速率。

## 2.4 ICEVpaTF2 可变基因

对 ICEVpaTF2 的可变基因区域进行生物信息学分析, 发现 HS1 包含 *higA* 和 *higB* 基因, 它们

编码一个毒素-抗毒素系统(TAS)。此外, ICEVpaTF2 可变区域 VR I 插入了 *hipA* 和 *hipB* 基因, 它们也编码 1 个 TAS。TAS 广泛存在于 ICE 中, 推测它们通过杀死失去该元件的细胞来维持 ICE 在宿主基因组的存在<sup>[21]</sup>。HS2 区包含 7 个功能不完全明确基因 *orf16-orf22* (表 2), 推测它们与乙酰基转移、磷脂合成和能量供应有关, 可能参与调节生物膜的形成、细菌运动和生物体的毒力<sup>[22]</sup>。HS3 由 2 个假想基因组成, HS4 由一个假想基因构成, 这些假想基因在 GenBank 中缺少同源基因, 因此功能无法预测。HS5 包含 7 个基因, 编码多个功能各异的蛋白, 包括: XRE 家族转录调节因子 XreB、I 型限制修饰系统中的甲基转移酶 M 亚基 HsdM、S 亚基 HsdS、蛋白激酶 PknB, 以及 3 个功能未知的假想蛋白。VR III 插入了 2 个不同的 Tn3 家族转座酶基因 *tra1*、*tra2*, 破坏了 *rumB* 基因。*rumB* 基因常因外源携带抗生素抗性

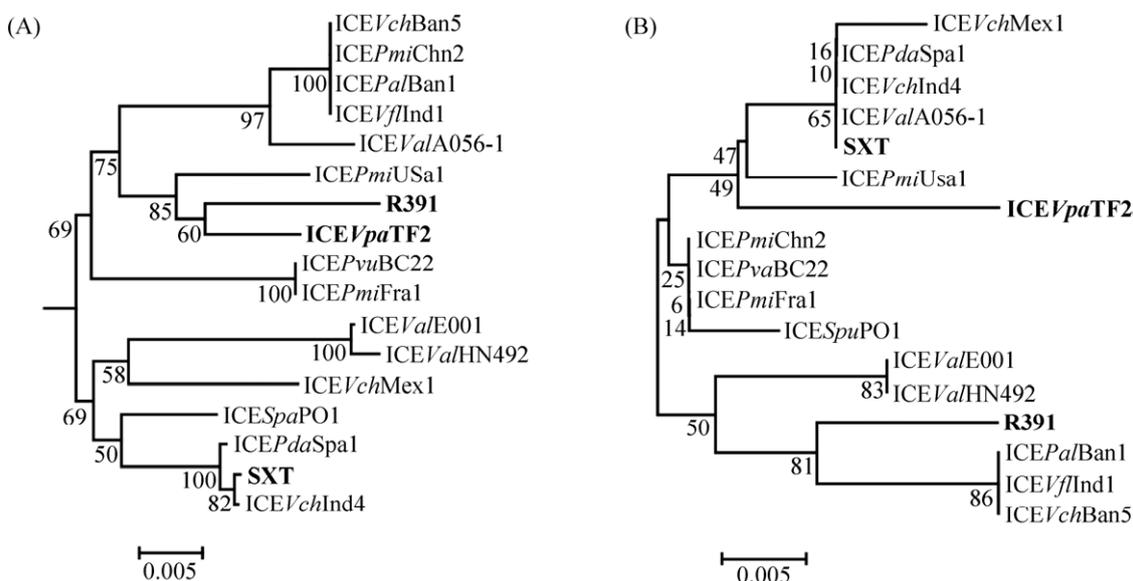


图 2. ICEVpaTF2 和随机选取的 SXT/R391 元件 *int* 基因(A)和 *xis* 基因(B)系统发育树

Figure 2. Phylogenetic trees of *int* (A) and *xis* (B) genes from ICEVpaTF2 and some randomly selected SXT/R391 ICEs. The trees were constructed using the neighbor-joining method. Bootstrap values were obtained after 1000 repetitions.

表 2. ICEVpaTF2 可变区域中的基因及功能预测

Table 2. Genes and functions prediction of variable genes in ICEVpaTF2

ORF	Gene	Functions of genes	Start	Stop	Length/aa
<i>orf1</i>	<i>hipB1</i>	XRE family transcriptional regulator	809	1129	321
<i>orf2</i>	<i>hipA2</i>	Type II toxin-antitoxin system toxin	1122	2336	1215
<i>orf3</i>	\	Hypothetical protein	5349	5516	168
<i>orf4</i>	<i>tra1</i>	Tn3 family transposase	7628	8221	594
<i>orf5</i>	<i>tra2</i>	Tn3 family transposase	8335	9948	1614
<i>orf6</i>	\	Hypothetical protein	11019	11273	255
<i>orf7</i>	<i>xreB</i>	XRE family transcriptional regulator	14338	14565	228
<i>orf8</i>	<i>hsdM</i>	Type I RM system methyltransferase subunit M	14565	16619	2055
<i>orf9</i>	\	Hypothetical protein	16616	16828	213
<i>orf10</i>	<i>hsdS</i>	Type I RM system methyltransferase subunit S	16825	18240	1416
<i>orf11</i>	<i>pknB</i>	Protein kinase	18250	22383	4134
<i>orf12</i>	\	Hypothetical protein	22380	24026	1647
<i>orf13</i>	\	Hypothetical protein	24023	29497	5475
<i>orf14</i>	<i>higA</i>	Helix-turn-helix domain-containing antitoxin	34898	35176	279
<i>orf15</i>	<i>higB</i>	Type II toxin-antitoxin system toxin	35173	35496	324
<i>orf16</i>	\	GNAT family N-acetyltransferase	39824	40330	507
<i>orf17</i>	\	DUF1778 domain-containing protein	40321	40587	267
<i>orf18</i>	\	Hypothetical protein	41147	41266	120
<i>orf19</i>	\	AAA family ATPase	41403	44132	2730
<i>orf20</i>	\	Hypothetical protein	44129	45238	1110
<i>orf21</i>	\	Phospholipid synthase family protein	45251	46783	1533
<i>orf22</i>	\	FIS family transcriptional regulator	46786	48231	1446
<i>orf23</i>	\	Conjugative transfer protein 345	51460	51807	348
<i>orf24</i>	\	Hypothetical protein	58299	59072	774
<i>orf25</i>	\	Hypothetical protein	71317	71517	201
<i>orf26</i>	\	Hypothetical protein	71579	72679	1101

基因的转座子插入而打断成 2 个基因 *rumB'* 和 *rum'B*<sup>[23-24]</sup>, 但是在 ICEVpaTF2 的 *rumB'* 和 *rum'B* 之间未发现抗性基因, 推测转座子的插入与插入位点基因的失活或者激活有关<sup>[24]</sup>。此外, 在 ICEVpaTF2 中还注释到位于 3 个非常规插入位点的单基因 *orf3*、*orf6*、*orf23*, 编码的假想蛋白长度不超过 116 个氨基酸, 其功能大部分完全未知, 小的假想基因大量存在于细菌基因组中, 其生物学意义不明。

## 2.5 ICEVpaTF2 与 *Vibrio owensii* V180403 的 ICE 基因组比较

以获得的完整 ICEVpaTF2 元件为质询序列, 通过 NCBI 的 Blast 搜索, 发现 GenBank 中与

ICEVpaTF2 最为相似的 ICE 来自于弧菌 *V. owensii* V180403 (GenBank No: CP033144), 命名为 ICEVowV18。在 78% 的覆盖度范围内 ICEVpaTF2 与 ICEVowV18 整体具有高达 98% 的相似性。ICEVpaTF2 与 ICEVowV18 的 HS1、HS2、VR I 区的可变基因显示高度的同源相似性, 此外 3 个非常规插入位点的 3 个单独基因也同样匹配到高度的同源相似性(图 1, 蓝色箭头)。ICEVpaTF2 与 ICEVowV18 具有高度的整体序列相似性, 由高度相似的结构和基因组成, 表明 2 个 ICE 很可能来自于共同的祖先。但二者也存在一些明显区别: ICEVowV18 中注释到三套 TAS; ICEVowV18 的 VR III 区是最大的可变区域, 主要包含 2 个 Tn3 家

族转座酶基因 *tra1* 和 *tra3*, 2 个 IS91 家族转座酶基因 *tra4* 和 *tra5*, 1 个四环素抗性基因 *tetR*, 2 个链霉素抗性基因 *strA*、*strB*, 和 1 个磺胺类抗生素抗性基因 *sul2*。目前 GenBank 中已经发布 1327 株副溶血弧菌全基因组序列(包括基因组草图), 以 ICEVpaTF2 为质询序列, 在测序菌株中仅发现 5 株副溶血弧菌存在与 ICEVpaTF2 相似的未鉴定 ICE, 5 株菌中有 2 株分离自对虾, 1 株来源于环境, 另外 2 株来源未知。其中与 ICEVpaTF2 最相似的一个未鉴定 ICE 位于副溶血弧菌菌株 Vb0624 中(CP041202.1), 它与 ICEVpaTF2 相比, 在 62% 的覆盖范围内具有 97% 的一致度, 5 个未鉴定副溶血弧菌株的 ICE 彼此间并不相同。

### 2.6 ICEVpaTF2 的剪切、环化活性

PCR 检测结果表明, 采用 *attL*-F1/*attL*-R1、*attR*-F4/*attR*-R4、*attR*-F4/*attL*-R1、*attL*-F1/*attR*-R4 引物组合均能扩增出预期大小的条带(图 3), 4 条引物的锚定位点见图 4。这一结果表明, 副溶血弧菌 TF2

细胞群体中, 部分细胞中的 ICEVpaTF2 处于整合于宿主基因组的状态, 部分细胞中 ICEVpaTF2 处于剪切、环化状态(图 4)。结合 ICEVpaTF2 具有完整的结构和功能基因, 表明 ICEVpaTF2 是一个有剪切、环化活性的完整 ICE, 因此也具有潜在的转移活性。

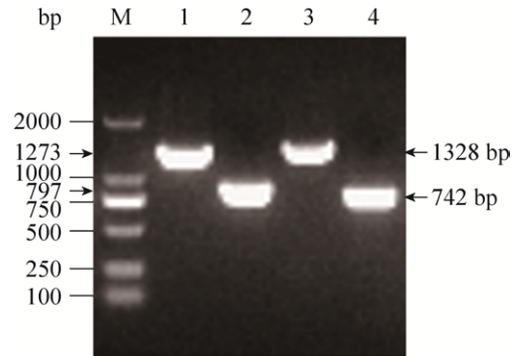


图 3. 锚定 *attL* 和 *attR* 位点上下游区域的引物不同组合 PCR 扩增结果

Figure 3. The PCR results based on different combinations of primers anchoring upstream and downstream regions of *attL* and *attR* sites. M: DNA marker DL2000; 1: *attL*-F1/*attL*-R1; 2: *attR*-F4/*attR*-R4; 3: *attR*-F4/*attL*-R1; 4: *attL*-F1/*attR*-R4.

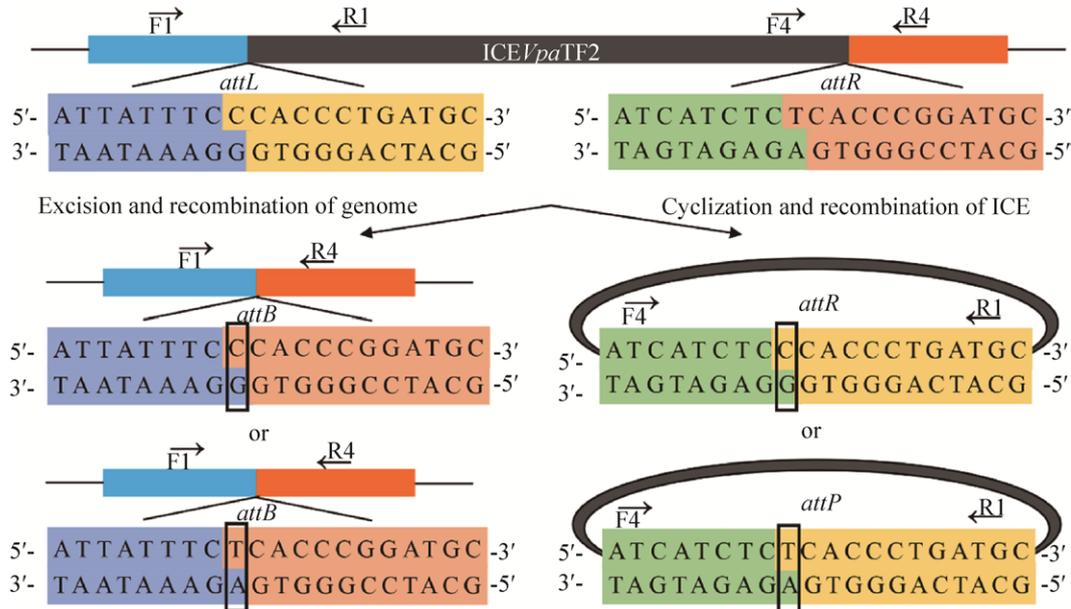


图 4. 推测的 ICEVpaTF2 剪切、环化时 *att* 位点形成示意图

Figure 4. Speculated schematic diagram of the *att*-sites forming during the excision/cyclization of ICEVpaTF2

进一步比较了 ICEVpaTF2 处于整合状态的 *attL* 和 *attR* 序列,发现 *attL* 和 *attR* 长度为 20 bp,二者并非完全相同,二者在第 3、6、9、15 位点存在碱基差异。ICEVpaTF2 发生剪切后,染色体 *attB* 位点呈现 *attL* 和 *attR* 的杂合重组,即 *attB* 位点由 *attL* 的 5'端和 *attR* 的 3'端组成(图 4),通过测序峰图发现,*attB* 第 9 个碱基具有 2 个同等大小的峰,因此第 9 碱基不确定,可能为 T 或 C(图 5-A)。同样环化后的 ICEVpaTF2 产生的 *attP* 位点也呈现 *attL* 和 *attR* 的杂合重组,即 *attP* 位点由 *attR* 的 5'端和 *attL* 的 3'端组成(图 4),通过测序峰图发现,*attP* 第 9 碱基同样具有 2 个同等大小的峰,因此第 9 碱基不确定,可能为 C 或 T(图 5)。

### 3 讨论

此前尽管在弧菌中发现了大量 SXT/R391 元件存在,但是并未有在副溶血弧菌中发现 SXT/R391 元件的明确报道,我们以 ICEVpaTF2 为质询序列,通过 Blast 搜索发现有 5 个副溶血弧菌菌株存在与 ICEVpaTF2 相似的未鉴定 ICE 序

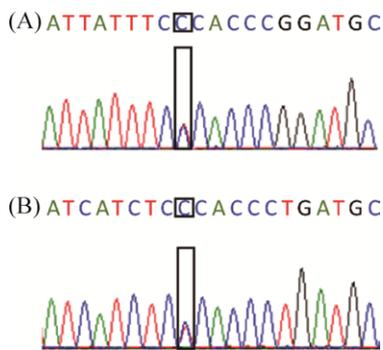


图 5. *attB* 位点测序峰图(A)和 *attP* 位点测序峰图(B)  
Figure 5. Sequencing peaks of *attB* (A) and *attP* (B). Four colors of peaks represent four different bases. Both *attB* and *attP* have two overlapping peaks at ninth base sites, which represent base T or C

列,因此表明本研究中发现的副溶血弧菌菌株 TF2 存在 ICE 元件并非是孤立事件,这 5 个 GenBank 中发现的未鉴定的 ICE 元件彼此并非不相同,表明副溶血弧菌也存在多样化的 ICE 元件。ICEVpaTF2 内大量基因编码功能各异蛋白,包括限制性修饰系统、DNA 修复系统、毒素-抗毒素系统等,它们赋予宿主对环境多样化的适应功能,同时对副溶血弧菌的遗传多样性产生重要影响。大多数 ICE 元件可变基因编码对多种抗生素的耐药性基因<sup>[11,25]</sup>,虽然 ICEVpaTF2 中未发现抗性基因,但是根据其结构预测,至少在 *rumB* 基因内部具有捕获和容纳携带抗性基因转座子的潜力。

SXT/R391 元件集中出现在非洲和亚洲沿海地区,且来源多为海洋细菌<sup>[10]</sup>,表明海洋环境很可能是 SXT/R391 主要的贮藏库,并且经由海洋环境株向临床株扩散<sup>[10,23,26]</sup>。对 *int* 的研究表明,一些诱发细菌 DNA 损伤的因素,如抗生素、丝裂霉素 C、紫外线等可以使 SXT/R391 的转移频率增加 10–100 倍以上<sup>[1,11,25,27]</sup>,因此有理由相信某些化学污染物或抗生素的滥用极有可能加剧 SXT/R391 的扩散。

多数情况下处于 C1 模板的 *int* 和 *xis* 紧密相连,功能上密切相关,*int* 负责 SXT/R391 的整合和剪切,*xis* 负责剪切<sup>[1]</sup>。一般而言,同一个 ICE 元件内的 *int* 如果归属 SXT 亚群,其 *xis* 也会归属 SXT 亚群,反之二者同属于 R391 亚群(图 2),但本文中的 ICEVpaTF2 中的 *int* 与 *xis* 分属于 SXT 和 R391 亚群(图 2),这一现象揭示同一模块内功能密切相关的基因也可能发生同源重组,因此很可能在 ICEVpaTF2 的进化史中,它可能曾与 R391 源的 ICE 处于同一个祖先细菌细胞中,导致了因同源重组而产生的杂合。2001 年, Hochhut 等<sup>[28]</sup>发现 R391 与 SXT 可以同时转移到大肠杆菌中形

成串联 R391、SXT 排列, 并且二者之间因为高度的同源, 发生了重组, 产生了杂合体。因此, 这一研究加深我们的推测。基于 *int* 和 *xis* 种系进化分析还发现相比 *int* 基因, *xis* 具有更快的进化速率, 这说明二者可能承受着不同的进化压力, 侧面反映出 *int* 相比 *xis* 具有更为重要的功能。

SXT/R391 元件中, HS1-HS5 以及可变区 VR I-VR IV 是最为常见的容纳外源 DNA 片段区域, 本研究中 ICEVpaTF2 和最为相似的 ICEVowV18 均产生 3 个独立的非典型小插入位点。此前我们对溶藻弧菌多个 SXT/R391 元件的研究中发现在 ICEVpaTF2 和 ICEVowV18 同样的位置具有非典型小插入位点<sup>[23]</sup>, 因此这些发现丰富了我们对 SXT/R391 元件遗传结构的认识。

在正常非压力诱导环境下, 大多数细胞中 SXT/R391 处于整合状态, 只有极少部分细胞中 ICE 处于剪切状态<sup>[6,29]</sup>, 本文中我们发现副溶血弧菌 ICEVpaTF2 的 SXT/R391 元件在细菌细胞群体中以整合和剪切环化两种形式存在, 再一次证实了这一发现。此前关于 SXT/R391 元件遗传结构的研究中, 并未对 *att* 位点序列以及其整合、剪切中的分配进行过分析。本研究发现 ICEVpaTF2 中 *attL* 的 5'端和 *attR* 的 3'端重组成新的 *attB* 位点, *attR* 的 5'端和 *attL* 的 3'端重组成剪切后环化 ICE 的 *attP* 位点。然而, 并不知道具体在 *attL* 和 *attR* 的哪一个位点开始剪切, 测序峰图发现 *attB* 或 *attP* 的正链第 9 位碱基具有 T 和 C 二种可能, 且两个峰几乎是重合的(图 5), 表明该位点出现碱基 C 或 T 的机率可能相同。因此我们推测 *attL* 和 *attR* 很可能在正链的第 8 和第 9 碱基间以及负链的第 9 和第 10 碱基间产生缺口, 在位点特异性整合酶 Int 的作用下可能发生链交换式重组(图 4), 由于 *attL*

和 *attR* 正链第 9 位碱基分别为 C 和 T, 导致重组时新产生的 *attB* 或 *attP* 的二条链在第 9 位碱基不配对, 必然产生修复以配对, 理论上这种错配修复随机, 结果 *attB* 或 *attP* 正链该位点具有 C 和 T 可能性的机会是均等的, 导致测序时该位点产生 C 和 T 近乎相同的峰值。这是第一次探讨 SXT/R391 元件从基因组剪切时 *attB* 和 *attP* 的重组产生的机制。

## 参考文献

- [1] Garriss G, Poulin-Laprade D, Burrus V. DNA-damaging agents induce the RecA-independent homologous recombination functions of integrating conjugative elements of the SXT/R391 family. *Journal of Bacteriology*, 2013, 195(9): 1991-2003.
- [2] Weiss E, Spicher C, Haas R, Fischer W. Excision and transfer of an integrating and conjugative element in a bacterial species with high recombination efficiency. *Scientific Reports*, 2019, 9(1): 8915.
- [3] Di Noto GP, Iriarte A, Ramírez MS, Centrón D, Quiroga C. ICE SXT vs. ICESh95: Co-existence of integrative and conjugative elements and competition for a new host. *Scientific Reports*, 2019, 9(1): 8045.
- [4] Gonzalez M, Huston D, McLenigan MP, McDonald JP, Garcia AM, Borden KS, Woodgate R. SetR<sub>ICE391</sub>, a negative transcriptional regulator of the integrating conjugative element 391 mutagenic response. *DNA Repair*, 2019, 73: 99-109.
- [5] Delavat F, Miyazaki R, Carraro N, Pradervand N, van der Meer JR. The hidden life of integrative and conjugative elements. *FEMS Microbiology Reviews*, 2017, 41(4): 512-537.
- [6] Garriss G, Burrus V. Integrating conjugative elements of the SXT/R391 family//Roberts AP, Mullany P. Bacterial Integrative Mobile Genetic Elements. Austin: Landes Bioscience, 2013: 217-234.
- [7] Bi DX, Xu Z, Harrison EM, Tai C, Wei YQ, He XY, Jia SR, Deng ZX, Rajakumar K, Ou HY. Iceberg: a web-based resource for integrative and conjugative elements found in bacteria. *Nucleic Acids Research*, 2012, 40(D1): D621-D626.
- [8] Waldor MK, Tschäpe H, Mekalanos JJ. A new type of conjugative transposon encodes resistance to

- sulfamethoxazole, trimethoprim, and streptomycin in *Vibrio cholerae* O139. *Journal of Bacteriology*, 1996, 178(14): 4157–4165.
- [9] Coetzee JN, Datta N, Hedges RW. R factors from *Proteus rettgeri*. *Microbiology*, 1972, 72(3): 543–552.
- [10] Luo P, He XY, Hu CQ. Bacterial SXT/R391 family from integrating conjugative elements—a review. *Acta Microbiologica Sinica*, 2014, 54(5): 471–479. (in Chinese) 罗鹏, 何香燕, 胡超群. 细菌整合性接合元件 SXT/R391 研究进展. *微生物学报*, 2014, 54(5): 471–479.
- [11] Hochhut B, Lotfi Y, Mazel D, Faruque SM, Woodgate R, Waldor MK. Molecular analysis of antibiotic resistance gene clusters in *Vibrio cholerae* O139 and O1 SXT constins. *Antimicrobial Agents and Chemotherapy*, 2001, 45(11): 2991–3000.
- [12] Rodríguez-Blanco A, Lemos ML, Osorio CR. Integrating conjugative elements as vectors of antibiotic, mercury, and quaternary ammonium compound resistance in marine aquaculture environments. *Antimicrob Agents Chemother*, 2012, 56(5): 2619–2626.
- [13] Daniels NA, Shafaie A. A review of pathogenic *Vibrio* infections for clinicians. *Infections in Medicine*, 2000, 17(10): 665–685.
- [14] Makino K, Oshima K, Kurokawa K, Yokoyama K, Uda T, Tagomori K, Iijima Y, Najima M, Nakano M, Yamashita A, Kubota Y, Kimura S, Yasunaga T, Honda T, Shinagawa H, Hattori M, Iida T. Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of *V. cholerae*. *Lancet*, 2003, 361(9359): 743–749.
- [15] Lee CT, Chen IT, Yang YT, Ko TP, Huang YT, Huang JY, Huang MF, Lin SJ, Chen CY, Lin SS, Lightner DV, Wang HC, Wang AHJ, Wang HC, Hor LI, Lo CF. The opportunistic marine pathogen *Vibrio parahaemolyticus* becomes virulent by acquiring a plasmid that expresses a deadly toxin. *Proceedings of the National Academy of Sciences of the United States of America*, 2015, 112(34): 10798–10803.
- [16] Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. The RAST server: rapid annotations using subsystems technology. *BMC Genomics*, 2008, 9: 75.
- [17] Wozniak RAF, Fouts DE, Spagnoletti M, Colombo MM, Ceccarelli D, Garriss G, Déry C, Burrus V, Waldor MK. Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs. *PLoS Genetics*, 2009, 5(12): e1000786.
- [18] Burrus V, Waldor MK. Control of SXT integration and excision. *Journal of Bacteriology*, 2003, 185(17): 5045–5054.
- [19] Ceccarelli D, Daccord A, René M, Burrus V. Identification of the origin of transfer (*oriT*) and a new gene required for mobilization of the SXT/R391 family of integrating conjugative elements. *Journal of Bacteriology*, 2008, 190(15): 5328–5338.
- [20] Beaber JW, Burrus V, Hochhut B, Waldor MK. Comparison of SXT and R391, two conjugative integrating elements: Definition of a genetic backbone for the mobilization of resistance determinants. *Cellular and Molecular Life Sciences*, 2002, 59(12): 2065–2070.
- [21] Wozniak RAF, Waldor MK. A toxin-antitoxin system promotes the maintenance of an integrative conjugative element. *PLoS Genetics*, 2009, 5(3): e1000439.
- [22] Jenal U, Malone J. Mechanisms of cyclic-di-gmp signaling in bacteria. *Annual Review of Genetics*, 2006, 40: 385–407.
- [23] Luo P, He XY, Wang YH, Liu QT, Hu CQ. Comparative genomic analysis of six new-found integrative conjugative elements (ICEs) in *Vibrio alginolyticus*. *BMC Microbiology*, 2016, 16: 79.
- [24] Saier Jr MH, Kukita C, Zhang ZG. Transposon-mediated directed mutation in bacteria and eukaryotes. *Frontiers in Bioscience-Landmark*, 2017, 22: 1458–1468.
- [25] Beaber JW, Hochhut B, Waldor MK. SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature*, 2004, 427(6969): 72–74.
- [26] Tian YS, Luo P, Liu QT, Hu CQ. Progress on bacterial recombination systems and their application. *Microbiology China*, 2017, 44(2): 473–482. (in Chinese) 田雨顺, 罗鹏, 刘秋婷, 胡超群. 细菌重组系统及其应用研究进展. *微生物学通报*, 2017, 44(2): 473–482.
- [27] He XY, Luo P, Gao Y, Hu CQ, Wei L, Liu QT. Effects of antibiotics on the transfer frequency of SXT/R391 element of *Vibrio alginolyticus*. *Acta Microbiologica Sinica*, 2016, 56(4): 643–650. (in Chinese) 何香燕, 罗鹏, 高沿, 胡超群, 韦露, 刘秋婷. 抗生素对溶藻弧菌 SXT/R391 元件转移频率的影响. *微生物学报*, 2016, 56(4): 643–650.
- [28] Hochhut B, Beaber JW, Woodgate R, Waldor MK. Formation of chromosomal tandem arrays of the SXT element and R391, two conjugative chromosomally integrating elements that share an attachment site. *Journal of Bacteriology*, 2001, 183(4): 1124–1132.
- [29] Carraro N, Poulin D, Burrus V. Replication and active partition of integrative and conjugative elements (ICEs) of the SXT/R391 family: the line between ICEs and conjugative plasmids is getting thinner. *PLoS Genetics*, 2015, 11(6): e1005298.

# Bioinformatics characteristics and excision/cyclization activities of an ICE element, ICEVpaTF2, in *Vibrio parahaemolyticus* TF2

Yingying Li<sup>1,2</sup>, Mei Xie<sup>1,2</sup>, Ziqian Cheng<sup>3</sup>, Zhuobo Li<sup>1,2</sup>, Peng Luo<sup>1\*</sup>

<sup>1</sup> CAS Key Laboratory of Tropical Marine Bio-resources and Ecology (LMB), Guangdong Provincial Key Laboratory of Applied Marine Biology (LAMB), South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, Guangdong Province, China

<sup>2</sup> University of Chinese Academy of Sciences, Beijing 101408, China

<sup>3</sup> Marine College, Agricultural University of Hebei, Cangzhou 060010, Hebei Province, China

**Abstract:** [Objective] To obtain complete sequence of an integrating conjugative element (ICE), ICEVpaTF2, from *Vibrio parahaemolyticus* TF2, and analyze its genomic characteristics; to investigate if ICEVpaTF2 has excision/cyclization activities from host genome and how *att* sites recombine during excision/cyclization. [Methods] The genomic frame sequences of *V. parahaemolyticus* TF2 was first annotated by RAST, which led to the finding of a potential and complete ICE element, named ICEVpaTF2. After manual splicing, PCR amplification followed by sequencing verification, a complete sequence of ICEVpaTF2 was obtained, and then it was annotated and characterized again. PCR detection and subsequent sequencing were performed to explore whether ICEVpaTF2 could excise from its host genome and cyclize itself and to characterize the forming of *attB* and *attP* by comparing the sequences of *attL*, *attR*, *attB* and *attP*. [Results] ICEVpaTF2, with a total length of 83588 bp, contains 52 conserved core genes of SXT/R391 ICE elements, which are related to excision, integration, self-transfer, and regulatory mechanisms. ICEVpaTF2 also contains five hot spots (HS), two variable regions (VR), and three atypical insertion sites accommodating variable genes. HS and VR contain a large number of variable genes, encoding for restrictive modification system, DNA repair system, toxin-antitoxin system, and etc., which confer the host extensive adaptive profits to the bacterial host. ICEVpaTF2 also contains unique genes with unknown functions. Through phylogenetic analysis of the two core conserved genes, *int* and *xis*, it was found that *int* and *xis* of ICEVpaTF2 belong to subgroups represented by that of R391 and SXT, respectively. ICEVpaTF2 excises at chromosomal *attL* and *attR* sites and cyclizes resulting in the forming of new heterozygous and recombined *attB* and *attP* sites. [Conclusion] ICEVpaTF2 belongs to the SXT/R391 family and it is a complete ICE element harboring self-excision and cyclization activities. New *attB* and *attP* sites are generated by hybrid recombination between primary *attL* and *attR* sites.

**Keywords:** *Vibrio parahaemolyticus*, SXT/R391 elements, ICEVpaTF2, bioinformatics characteristics, excision/cyclization

(本文责编: 张晓丽)

Supported by the National Natural Science Foundation of China (31370149) and by the Program of Fishery Problem Tackling of Guangdong Province (A201701B03)

\*Corresponding author. E-mail: luopengli@163.com

Received: 23 July 2019; Revised: 23 August 2019; Published online: 17 October 2019