



FACULTAD DE BIOLOGÍA

DEPARTAMENTO DE MICROBIOLOGÍA

**SISTEMAS DE SECRECIÓN TIPO  
VI (T6SS) EN *RHIZOBIALES*.  
ESTUDIO *IN SÍLICO* DEL  
CLÚSTER DE *S. FREDII* USDA257**

Trabajo Fin de Grado

**Ana Sánchez Reina**

Tutor: Francisco Javier López Baena

Directora: Patricia Bernal Guzmán

*Junio 2020*

# ÍNDICE

<b>1. RESUMEN</b> .....	1
<b>2. INTRODUCCIÓN</b> .....	1
2.1. Sistemas de secreción: clasificación y generalidades. ....	1
2.2. Sistema de secreción tipo VI. ....	4
2.2.1. Estructura y ensamblaje. ....	4
2.2.2. Tipos de efectores. ....	6
2.2.3. Componentes accesorios. ....	7
2.3. Bacterias relacionadas con plantas. ....	8
2.3.1. Microbiota existente en los diferentes órganos de una planta. ....	8
2.3.2. Interacción bacteria-bacteria. ....	9
2.3.3. Interacción planta-bacteria. ....	9
2.3.3.1. Simbiosis <i>Rhizobium</i> -leguminosa.....	10
<b>3. OBJETIVOS</b> .....	13
<b>4. MATERIALES Y MÉTODOS</b> .....	13
<b>5. RESULTADOS Y DISCUSIÓN</b> .....	14
5.1. Estudio filogenético de los T6SS en el orden <i>Rhizobiales</i> .....	15
5.2. Análisis bioinformático del T6SS de <i>S. fredii</i> USDA257.....	18
<b>6. CONCLUSIONES</b> .....	27
<b>7. BIBLIOGRAFÍA</b> .....	27
<b>8. ANEXOS</b>	

## 1. RESUMEN

El sistema de secreción tipo VI (T6SS) es una maquinaria bacteriana a través de la cual se secretan toxinas/efectores al interior de células eucariotas y/o procariotas. Este sistema se ha visto involucrado en patogénesis, pero su papel fundamental es la competición entre bacterias.

En este trabajo, se aborda el estudio de los T6SSs en el orden *Rhizobiales*, donde la función del T6SS es desconocida. En primer lugar, hemos llevado a cabo un análisis filogenético de los T6SSs en *Rhizobiales* para posteriormente centrarnos en la caracterización del clúster del T6SS de *Sinorhizobium fredii* USDA257, una cepa de especial interés en agricultura por su amplio rango de nodulación. Nuestro análisis filogenético muestra que los T6SS del orden *Rhizobiales* están distribuidos principalmente en dos grupos (grupo 3 y 5), de los cinco descritos en la literatura para los T6SSs. El análisis bioinformático del T6SS de nuestro organismo modelo *S. fredii* USDA257 muestra que el clúster génico contiene todos los genes necesarios para la biosíntesis de las proteínas estructurales necesarias para ensamblar un T6SS activo. Además, hemos encontrado genes accesorios que codifican proteínas reguladoras y estructurales que regulan el sistema y confieren nuevas características funcionales al mismo (TagF, TagM, TagL y TagY). El estudio *in silico* también nos ha permitido identificar genes (USDA\_c28430, USDA\_c28470 y USDA\_c28480) que codifican los posibles efectores Vgrg<sub>e</sub>, Tre2 y Tre1, respectivamente. Las predicciones funcionales de estos efectores indican que podrían estar relacionados con la degradación de la pared celular vegetal. En este caso, serían los primeros efectores de tipo VI descrito, involucrado en el establecimiento de la simbiosis *Rhizobium*-leguminosa y que actuarían de manera conjunta con el sistema de secreción tipo III para llevar a cabo este proceso.

## 2. INTRODUCCIÓN

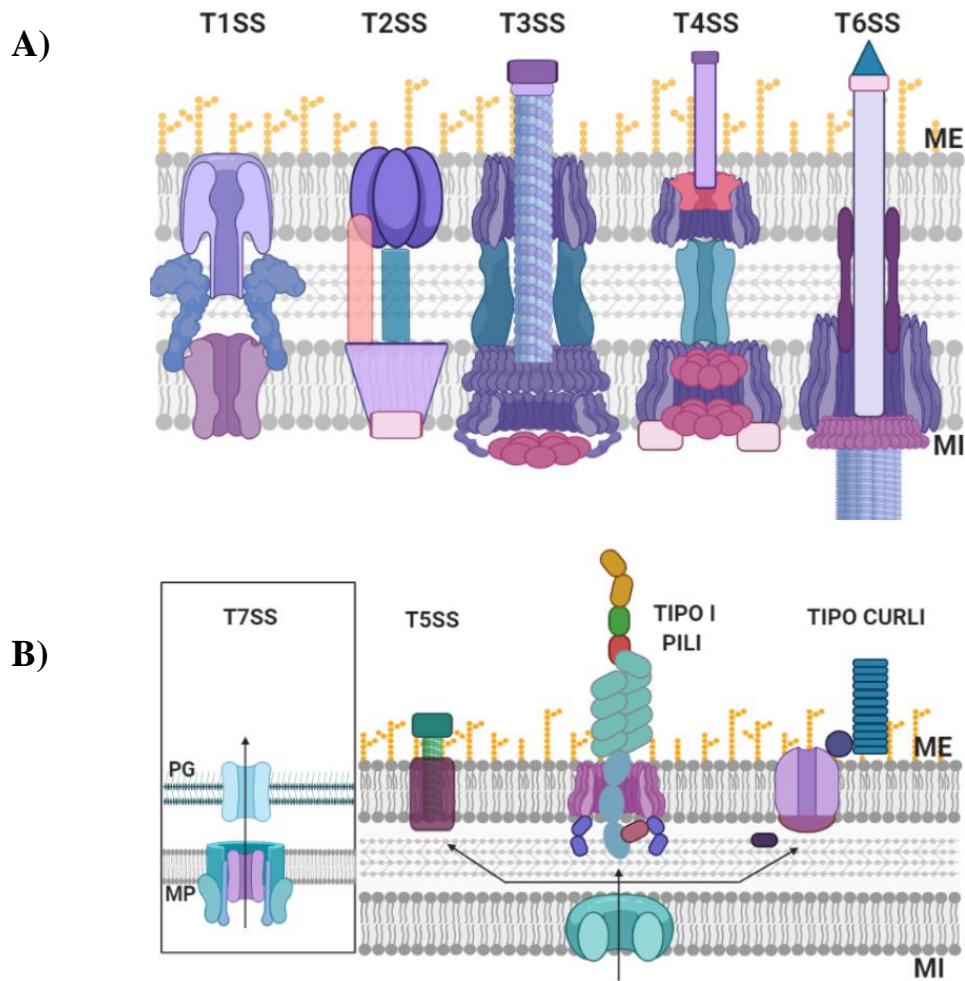
### 2.1. Sistemas de secreción: clasificación y generalidades.

Los sistemas de secreción bacterianos son maquinarias de naturaleza proteica, que llevan a cabo muchos procesos biológicos como la conjugación, la competición interbacteriana o la interacción con células hospedadoras<sup>1</sup>. Estas maquinarias se encargan de translocar proteínas desde el interior celular hacia el espacio extracelular, inyectarlas en otras células diana o depositarlas en su propia superficie. Actualmente, se distinguen hasta nueve tipos distintos de sistemas de secreción<sup>1</sup>. Los tipos I, II, III, V, VI, VIII y IX son

propios de bacterias Gram-negativas, el tipo IV se encuentra tanto en bacterias Gram-negativas como en Gram-positivas, mientras que el tipo VII es exclusivo de bacterias Gram-positivas y se ha estudiado principalmente en bacterias del género *Mycobacterium*<sup>2</sup> y *Staphylococcus*<sup>3</sup>.

En bacterias Gram-negativas, estos sistemas se han dividido en dos grandes categorías atendiendo a si la maquinaria atraviesa la doble membrana celular, interna y externa, o si la maquinaria atraviesa únicamente la membrana externa (**Fig. 1**). La primera categoría engloba cinco sistemas donde se incluyen los tipos I, II, III, IV y VI (**Fig. 1A**), mientras que en la segunda se encuentran el tipo V y otras maquinarias que actúan de apéndices en la superficie celular, conocidas como pili y curli (también denominado tipo VIII)<sup>1</sup> (**Fig. 1B**). Todos ellos desarrollan diversas funciones; así, el T1SS (**Fig. 1A**) se encarga fundamentalmente de la secreción de factores de virulencia al medio extracelular, como la hemolisina A (proteína formadora de poros)<sup>4</sup>. También secreta factores que ayudan a la bacteria en la adquisición de nutrientes, como el hemóforo HasA (proteína transportadora del cofactor hemo)<sup>5</sup> o promueven la adhesión celular, LapA en *Pseudomonas putida*<sup>6</sup>. El T2SS (**Fig. 1A**) presente en patógenos, secreta gran cantidad de enzimas hidrolíticas como las pseudolisinas secretadas por *P. aeruginosa*, que son fundamentales para su crecimiento y supervivencia<sup>7</sup>. A través del T2SS, también se secretan otros tipos de toxinas como la toxina del cólera en *Vibrio cholerae*<sup>8</sup>. El sistema de secreción tipo III o T3SS (**Fig. 1A**), se encuentra fundamentalmente en bacterias patógenas y es fundamental en los procesos de virulencia. Los efectores de tipo III en bacterias patogénicas de los géneros *Yersinia* y *Pseudomonas* manipulan al hospedador, modulando funciones celulares y pudiendo llegar a causar la muerte celular<sup>9</sup>. El T3SS también está presente en algunas bacterias no patógenas, como las pertenecientes al orden *Rhizobiales*, en las que está implicado en la simbiosis con leguminosas<sup>10</sup>. El T3SS presente en algunas estirpes de rizobios, desempeña un papel fundamental en la determinación del rango de hospedador, la supresión de las respuestas de defensa, la eficiencia de la simbiosis o la formación de nódulos en ausencia de factores de nodulación<sup>10</sup>. En cuanto al sistema de secreción de tipo IV o T4SS (**Fig. 1A**), es de especial interés su capacidad para mediar el paso de ADN hacia células eucariotas y procariontas. Ésta es una habilidad que únicamente posee este sistema y es responsable, por ejemplo, de la transmisión de plásmidos que contienen genes de resistencia a antibióticos, entre muchos otros<sup>1</sup>. Este sistema está muy bien estudiado en *Agrobacterium*

*tumefaciens*, patógeno de planta donde el T4SS transfiere ADN a las células de la planta huésped como primer paso de su mecanismo de virulencia<sup>11</sup>.



**Figura 1. Organización estructural de los sistemas de secreción presentes tanto en bacterias Gram-positivas como Gram-negativas. A)** Sistemas que atraviesan la doble membrana celular (T1SS, T2SS, T3SS, T4SS, T6SS). **B)** Sistemas que atraviesan únicamente la membrana externa (T5SS, Tipo I, Tipo Curli) o la membrana interna en el caso de Gram-positivas, como el T7SS presente en *Mycobacterium* y *Staphylococcus*.

El sistema de secreción de tipo V (**Fig. 1B**) está involucrado en la adhesión celular y la formación de biopelículas. En el caso de *Yersinia enterocolitica*<sup>12</sup>, la secreción de adhesinas a través del T5SS, como YadA, permite la invasión de tejidos para causar infecciones. Además, el T5SS está implicado en la secreción de toxinas que actúan como factores de virulencia, como es el caso de VacA de *Helicobacter pylori*<sup>13</sup> que promueve la apoptosis de la célula hospedadora mediante la formación de poros en la membrana plasmática de ésta. El tipo I o pili se caracteriza por poseer un apéndice (**Fig. 1B**), siendo su función principal el anclaje a células hospedadoras para la invasión de células eucariotas. El T8SS o curli (**Fig. 1B**), está implicado principalmente en la adhesión

celular<sup>14</sup>. Un ejemplo de ello es la cepa K-12 de *Escherichia coli*, la cual expresa este sistema de secreción para la formación de biopelículas<sup>14</sup>.

El único sistema exclusivo de Gram-positivas, el T7SS (**Fig. 1B**) está implicado en procesos de virulencia en *Mycobacterium*<sup>2</sup> y en competición entre bacterias en el caso de *Staphylococcus*<sup>3</sup>.

## 2.2. Sistema de secreción tipo VI.

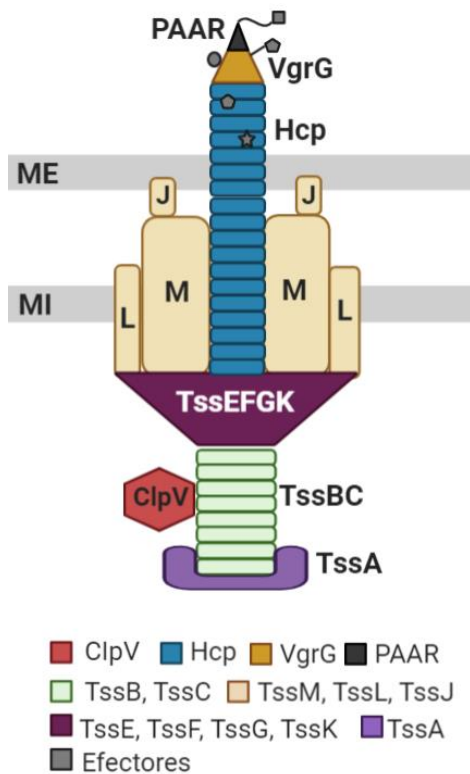
El sistema de secreción tipo VI (T6SS, del inglés Type Six Secretion System), en el que nos vamos a centrar en este trabajo, fue descubierto por primera vez en *Rhizobium leguminosarum*<sup>15</sup>, aunque hasta 2006 no se le asignó el término de T6SS cuando fue identificado y descrito en *V. cholerae*<sup>16</sup>. El T6SS se encuentra en el 25% de las bacterias Gram-negativas, principalmente en el filo Proteobacterias donde se incluyen las clases  $\alpha$ -,  $\beta$ -,  $\gamma$ -proteobacterias<sup>17</sup>.

Este sistema de secreción secreta efectores/toxinas principalmente al interior de células procariotas, por lo que juega un papel esencial en la competición entre bacterias por el mismo nicho ecológico. Un ejemplo de este proceso es el de *P. aeruginosa*, capaz de producir la muerte de *V. cholerae* mediante la inyección del efector *Tse1* en su periplasma<sup>18</sup>. Algunos efectores tienen como diana células eucariotas y en este caso el T6SS está implicado en patogénesis, como ocurre en la inflamación intestinal causada por *V. cholerae* gracias al dominio ACD de VgrG-1<sup>19</sup>.

### 2.2.1. Estructura y ensamblaje.

El T6SS es un complejo multiproteico formado por trece constituyentes principales. Los genes que codifican estas proteínas se encuentran agrupados en un mismo clúster génico<sup>20</sup> denominado *tss* (de sus siglas en inglés, Type Six Secretion)<sup>21</sup>. En el mismo clúster se pueden encontrar genes que codifican proteínas con funciones complementarias, también conocidas como componentes accesorios y denominados Tag (del inglés, Type VI Accessory Genes). Estos componentes no se encuentran en todos los sistemas de tipo VI, pero en aquellos que aparecen, llevan a cabo diferentes funciones como controlar el ensamblaje del sistema o la regulación del mismo<sup>22</sup>.

En cuanto a la estructura del sistema, podemos diferenciar tres partes principales: el complejo de membrana, la placa base y la cola. El complejo de membrana está formado por tres proteínas principales, TssJ, TssL y TssM y se encarga de anclar el sistema de secreción a la membrana<sup>23</sup> (**Fig. 2, beige**). Tras el ensamblaje de éste, se lleva a cabo el reclutamiento de TssA, que forma una estructura en forma de anillo que a su vez recluta a las proteínas que componen la placa base: TssK, TssE, TssF y TssG<sup>24</sup> (**Fig. 2, magenta**). Una vez formada, ésta establece la unión entre el complejo de membrana y la cola<sup>24</sup>. En el centro de la placa base se localiza una estructura en forma de aguja, constituida por un



**Figura 2. Representación esquemática de la estructura del T6SS.** Componentes de la placa base (TssEFGK) en color magenta, componentes del complejo de membrana (TssLMJ) en color beige y componentes de la vaina (TssBC) en color verde.

trímero de proteínas VgrG y la proteína PAAR en la parte superior, terminando en forma de punta<sup>25</sup> (**Fig. 2, negro**). A partir de la placa base y el anillo hexamérico formado por TssA situado en la parte distal del sistema, comienza la polimerización de la cola, hacia el citosol<sup>26</sup>. De esta manera, se sintetiza el tubo interno, formado por hexámeros en forma de anillo de la proteína Hcp (TssD) y rodeado por una vaina contráctil, compuesta por las proteínas TssB y TssC<sup>27</sup> (**Fig. 2**). Ambas se unen formando dímeros estables repetidos en conformación helicoidal, siendo suficiente seis dímeros TssBC para cubrir un anillo de Hcp<sup>28</sup>. La formación del T6SS finaliza cuando la parte distal de éste conecta con la membrana que se encuentra en el lado opuesto de la bacteria<sup>26</sup>.

Tras el ensamblaje del sistema completo, la secreción de los efectores tiene lugar una vez que se contrae la vaina. Después del estímulo necesario, como puede ser el contacto con la célula diana, se

inicia en la placa base la reorganización de TssBC, produciéndose una rotación del complejo. Esto hace que la vaina se acerque hacia la placa base y por tanto disminuya su longitud. De esta forma, se promueve la expulsión del tubo interno de la cola a través de la placa base y el complejo de membrana ya que ambos forman un canal en la membrana celular para permitir este paso<sup>29</sup>. La expulsión del tubo interno se da hacia el interior de la célula diana junto con la aguja, y los efectores que pueden ir conectados a ella o

localizarse en el interior del tubo, serán liberados al medio intracelular<sup>30</sup>. A continuación, la vaina, que queda en el citosol, se desensambla gracias a una AAA<sup>+</sup> ATPasa, conocida como ClpV (**Fig. 2, rojo**), la cual desempeña un papel crucial en el reciclaje del sistema para permitir su formación de nuevo<sup>31</sup>.

#### 2.2.2. Tipos de efectores.

Dentro o fuera del clúster de T6SS también podemos encontrar de forma habitual genes que codifican proteínas efectoras, que son moléculas secretadas por el mismo sistema de tipo VI y que por lo general inhiben el crecimiento o manipulan el funcionamiento de la célula diana<sup>32</sup>. Normalmente estos genes se encuentran aguas abajo de los genes que codifican las proteínas VgrG, Hcp y PAAR. Se diferencian dos grandes grupos de efectores basados en el modo de interacción con proteínas estructurales del sistema, los denominados “carga” y los “especializados”<sup>32</sup>. Los efectores tipo carga son aquellos que interactúan con las proteínas Hcp, VgrG o PAAR. Esta interacción proteína-efector puede ocurrir de forma directa o mediante adaptadores (Tap, EagR), proteínas independientes capaces de reconocer el sistema y acoplarse para unir el efector posteriormente<sup>32</sup>. Por otro lado, encontramos los efectores especializados, codificados en el mismo gen que codifica las proteínas Hcp, VgrG o PAAR y, por tanto, se encuentran fusionados a ellas en forma de dominio C-terminal<sup>33</sup>. En este caso, las proteínas estructurales se denominan “Hcp, VgrG o PAAR evolucionadas” o Hcp<sub>e</sub>, VgrG<sub>e</sub> o PAAR<sub>e</sub>, término que proviene del inglés “evolved” Hcp, VgrG o PAAR<sup>34</sup>.

Los efectores de tipo 6 pueden tener como diana tanto células eucariotas como procariontes. Dentro de las toxinas anti-bacterianas encontramos amidasas, peptidasas hidrolasas de peptidoglicano, fosfolipasas y nucleasas, las cuales van dirigidas hacia el periplasma, la membrana o el citoplasma de la bacteria con la que compiten<sup>33</sup>. Además, es esencial que la bacteria que sintetiza el T6SS produzca, a su vez, proteínas inmunitarias que inactiven las propias toxinas o aquellas que proceden de células hermanas evitando ser su diana<sup>33</sup>. Por otro lado, los efectores que tienen como diana células eucariotas están implicados en procesos de manipulación del hospedador. Intervienen en mecanismos de reordenación de los componentes del citoesqueleto, fagocitosis, estrés oxidativo, inflamación, proliferación celular o apoptosis<sup>33,35</sup>. Recientemente, también se han identificado efectores antifúngicos<sup>36</sup>. Éstos están involucrados en la inhibición de la ingesta de nutrientes, detienen el crecimiento celular y causan la muerte en hongos<sup>36</sup>.



### 2.2.3. Componentes accesorios.

En la mayoría de los sistemas de secreción de tipo VI podemos encontrar componentes accesorios que varían entre clústeres y que codifican proteínas que regulan la articulación/desarticulación y/o la actividad de estos sistemas<sup>37</sup>.

Algunos de estos componentes han sido ampliamente estudiados en *P. aeruginosa*, como es el caso de la serina/treonina kinasa PpkA<sup>38</sup>. PpkA es una proteína transmembrana que posee un dominio en el periplasma y otro en el citosol, siendo este último el que posee la actividad kinasa. Esta kinasa tras su dimerización y autofosforilación, es capaz de unir y fosforilar a la proteína citosólica Fha codificada en el mismo clúster<sup>38</sup>. Según otros autores, Fha fosforilada interacciona con algunos de los componentes estructurales del T6SS como TssL<sup>39</sup>. Esta interacción permite el reclutamiento de los demás componentes del sistema, promoviendo su ensamblaje y activación para llevar a cabo la secreción de toxinas/efectores. En cambio, se ha propuesto un modelo diferente para esta cascada regulatoria en el T6SS de *A. tumefaciens*<sup>39</sup>. En este caso, PpkA fosforila a TssL permitiendo su dimerización, para posteriormente reclutar a Fha. De igual forma, tras la unión de Fha a TssL se da el ensamblaje del sistema<sup>39</sup>. Por otro lado, encontramos la proteína con función inversa a PpkA, la fosfatasa PppA, la cual desfosforila a la proteína diana de PpkA, en el caso de *P. aeruginosa*, Fha, bloqueando el ensamblaje del sistema<sup>38</sup>. Esta inhibición también puede estar mediada por el represor TagF que se une directamente a Fha, bloqueando su interacción con TssL y por tanto inhibiendo el ensamblaje del T6SS; este proceso se ha estudiado en *A. tumefaciens* y *P. aeruginosa*<sup>39</sup>.

Otras proteínas accesorias que se pueden encontrar en *E. coli* y *P. putida* son TagL y TagP, respectivamente<sup>40,41</sup>. Estas proteínas poseen un dominio de unión a TssL similar al de TssM (IcmF), y además un dominio de unión a peptidoglicano (OmpA). Estos dominios le permiten establecer una interacción entre el complejo de membrana y el peptidoglicano<sup>41</sup>.

Diferentes estudios han mostrado que la estructura de los T6SS está muy conservada<sup>20</sup> a pesar de que sus componentes estructurales no tienen una homología muy alta. Se ha observado que las proteínas estructurales del T6SS están presentes en todos los grupos filogenéticos y son esenciales para la función del mismo, mientras que las accesorias son específicas de cada clúster y regulan y articulan el ensamblaje del

sistema<sup>17</sup>. Estas características han permitido clasificar a los clústeres del T6SS en cinco grupos filogenéticos<sup>17</sup>.

### 2.3. Bacterias relacionadas con plantas.

Desde la aparición de las primeras plantas en La Tierra, éstas han evolucionado de manera conjunta con las bacterias, conviviendo y compartiendo nichos ecológicos. Como resultado de esta co-evolución, han surgido dos tipos de interacción planta-bacteria, conocidas como patogénesis y simbiosis. En la primera de ellas, la patogénesis, la bacteria obtiene un beneficio y a la vez causa daños en la planta. En la segunda, la simbiosis, definida como la asociación física e íntima de dos o más organismos diferentes, se diferencian tres tipos distintos: el comensalismo, el mutualismo y el parasitismo<sup>42</sup>. De ellos, el más común es el mutualismo, que se define como una relación con beneficios mutuos entre ambos organismos<sup>42</sup>. En este caso, se encuentra muy bien estudiado el mecanismo de fijación de nitrógeno en la interacción simbiótica rizobio-leguminosa que se describirá en la sección 2.3.3.1.

Además de los diferentes tipos de interacciones planta-bacteria, también queremos destacar las interacciones que se desarrollan entre los microorganismos que habitan los distintos tejidos de la planta, como se describe a continuación.

#### 2.3.1. Microbiota existente en los diferentes órganos de una planta.

La microbiota de una planta es el conjunto de microorganismos, que establecen interacción con cualquier órgano de la misma, incluyendo hongos, bacterias y arqueas<sup>43</sup>. Los microorganismos pueden alojarse en cualquier tejido de la planta, desde las raíces hasta las hojas. Según el lugar donde se alojen reciben diferentes nombres, así se denominan epífitos a aquellos microorganismos localizados en las superficies de la planta como la filosfera (partes aéreas) o la rizosfera (zona del suelo inmediata a las raíces de la planta)<sup>44</sup>. Mientras que los microorganismos endófitos son los que se alojan en el interior de los tejidos de la misma (endosfera)<sup>44</sup>. Esta clasificación no es exacta, ya que la localización de los microorganismos puede ir variando a lo largo del ciclo de vida de la planta. De esta manera, existen organismos que en un principio residen en la filosfera y en un momento determinado pasan a la endosfera o viceversa<sup>44</sup>. Un gran número de bacterias que forman parte de la microbiota de la planta se alojan en la rizosfera, zona del suelo influenciada por las raíces de la planta y los compuestos que secreta, como son ácidos orgánicos, hormonas, aminoácidos, azúcares, vitaminas o nucleótidos. La rizosfera

representa uno de los ecosistemas más complejos, cuya composición y actividad microbiana es regulada por las secreciones de la propia planta<sup>43</sup>.

### 2.3.2. Interacción bacteria-bacteria.

La microbiota no interacciona únicamente con la planta hospedadora, sino que las bacterias y otros microorganismos que la componen también interaccionan entre ellos. Estas interacciones son fundamentales para la planta y se dan de forma mayoritaria en la rizosfera donde la población es más densa y las interacciones, tanto de cooperación como de competición, son más frecuentes<sup>45</sup>. Uno de los principales componentes de la microbiota de las plantas son las rizobacterias que promueven el crecimiento vegetal y son conocidas como PGPR (del inglés, Plant Growth Promoting Rhizobacteria). Este grupo de bacterias colabora en una gran variedad de procesos como son la fijación de nitrógeno o la solubilización de minerales que aumentan así su biodisponibilidad y mejoran la toma de nutrientes por parte de la planta<sup>44</sup>. Algunas de estas bacterias pueden proteger a la planta de enfermedades, y se denominan agentes de biocontrol<sup>46</sup>. Los agentes de biocontrol eliminan patógenos usando diferentes estrategias entre las que se incluye la producción de antibióticos, compuestos volátiles o incluso enzimas líticas por parte de las bacterias beneficiosas y el T6SS, que permite la secreción de agentes antimicrobianos directamente al interior de fitopatógenos<sup>45,43</sup>. Algunos géneros implicados en los distintos mecanismos de competición entre microorganismos, y por tanto de gran interés para la agricultura son *Pseudomonas*, *Nocardioides*, *Streptomyces* y *Terrabacter*. Un ejemplo bien estudiado es el caso de los cultivos de trigo, donde estos géneros bacterianos interaccionan con el hongo *Gaeumannomyces graminis* var. *tritici*<sup>47</sup> suprimiendo su patogenicidad. En el caso de cultivos de fresa, muy afectados por el hongo *Fusarium oxysporum*, se han identificado gran variedad de bacterias como *Actinobacteria*, *Proteobacteria*, *Acidobacteria*, *Gemmatimonadetes*, *Nitrospira* y *Chloroflexi* capaces de secretar efectores antifúngicos<sup>48</sup> y proteger de este modo las cosechas.

### 2.3.3. Interacción planta-bacteria.

Como se ha explicado anteriormente, las relaciones entre plantas y bacterias están muy bien estudiadas, distinguiéndose dos tipos: la patogénesis y la simbiosis. En el caso de la patogénesis, la gran mayoría de bacterias utilizan el T3SS para infectar plantas susceptibles<sup>9</sup>. Las bacterias fitopatógenas inyectan efectores a través del T3SS directamente al interior de la célula vegetal, donde alteran rutas de transducción de señales, provocando la supresión del sistema inmune, promoviendo la infección y

beneficiándose de los nutrientes de la planta<sup>49</sup>. En cambio, en algunas bacterias del orden *Rhizobiales* no patogénicas conocidas comúnmente como rizobios, el T3SS está implicado en la simbiosis fijadora de nitrógeno con las leguminosas<sup>50</sup>. En estas bacterias, el T3SS es inducible por flavonoides y está implicado en la determinación del rango de hospedador, la supresión de las respuestas de defensa, la eficiencia de la simbiosis y en la inducción de la formación de nódulos en ausencia de factores de nodulación<sup>50</sup>.

#### 2.3.3.1. Simbiosis *Rhizobium*-leguminosa.

La simbiosis *Rhizobium*-leguminosas es el proceso por el cual se fija el nitrógeno atmosférico (N<sub>2</sub>), que es usado por la planta, y ésta a cambio le aporta nutrientes a la bacteria para su crecimiento<sup>51</sup>. De esta manera, se establece un diálogo molecular complejo entre los rizobios y las leguminosas que permite la formación en las raíces de la planta de unas nuevas estructuras llamadas nódulos, donde existe un ambiente favorable para la diferenciación de la bacteria a bacteroide y posterior fijación de nitrógeno<sup>51</sup>. Este mecanismo es muy complejo y requiere grandes vías de señalización para que pueda desarrollarse correctamente. A grandes rasgos, se requieren dos fases principales: la organogénesis del nódulo y la infección bacteriana<sup>51</sup>.

El diálogo molecular comienza con la secreción de ciertos compuestos fenólicos, denominados flavonoides, por parte de la planta. Estos compuestos son captados por la bacteria y se unen al regulador transcripcional NodD, activando la transcripción de los genes *nod*, que codifican enzimas implicadas en la síntesis de moléculas conocidas como Factores nod (NF, del inglés Nod Factor) que inducen la formación de nódulos en las leguminosas<sup>51</sup>. Los NF son lipoquitoligosacáridos que pueden presentar grupos funcionales adicionales, adquiriendo estructuras diferentes para ser reconocidos correctamente por sus receptores. De esta manera, son específicos en la simbiosis y fundamentales en todas las etapas del proceso<sup>51</sup>. En el reconocimiento de NF por parte de la planta están implicados los receptores NFR (del inglés, Nod Factor Receptor) de tipo quinasa, que poseen dominios extracelulares de tipo LysM para unirse a los NF. Tras esta interacción, se unen a los NFR los receptores de simbiosis tipo quinasa (SYMKR, del inglés Symbiosis Receptor-like Kinase), los cuales están asociados a la enzima 3-hidroxi-3-metilglutaril-CoA reductasa (HMGR), encargada de la síntesis de mevalonato<sup>52</sup>. Este mevalonato puede actuar de segundo mensajero, interaccionando con los poros nucleares y con los canales de calcio que se hallan en la membrana nuclear, produciendo oscilaciones en las concentraciones de calcio dentro del núcleo. Estas oscilaciones serán

detectadas por varias enzimas, como la proteína quinasa calcio-calmodulina dependiente (CCaMK) que, tras su activación, media la fosforilación de otras proteínas como CYCLOPS y ellas, a su vez, activan otros genes de nodulación. Por tanto, se deduce que la activación de la vía CCaMK es esencial y el primer paso requerido para la organogénesis del nódulo<sup>51,53</sup>.

Este diálogo molecular viene acompañado, en una primera fase, de señalización hormonal en la planta, donde juegan un papel fundamental las citoquininas (hormona implicada en la diferenciación celular) y las auxinas (hormona implicada en el control del desarrollo). Ambas son esenciales y se combinan para conseguir un buen desarrollo y crecimiento de la planta. En términos generales, se requieren altas concentraciones de citoquinina y bajos niveles de auxina en las raíces para el desarrollo de los nódulos<sup>51</sup>.

En la segunda fase del proceso se produce la infección por los rizobios. Para ello es necesario que las bacterias atraviesen los distintos tejidos que componen la raíces, como la epidermis (parte más externa de la raíz que forma pelos) y el córtex (se encuentra bajo la epidermis y compone la mayor parte de la raíz). Así, podrán acceder al primordio nodular donde se está desarrollando el nódulo y alojarse allí<sup>51</sup>.

Aunque existen distintas vías de infección, el método más común es el mecanismo que utiliza tubos de infección o IT (del inglés, Infection Thread) que consiste en la formación de una estructura en forma de tubo en los pelos radicales de las raíces. Este mecanismo comienza cuando receptores específicos en los pelos radicales reconocen los NF secretados por las bacterias. Esto permite que se produzcan grietas entre las células de los pelos radicales y por tanto se adentren en ellas las bacterias, creando un bolsillo de infección, donde continúan dividiéndose y formando colonias conocidas como focos de infección. A partir de estos focos se comienza a formar el tubo de infección, proceso en el que están involucradas muchas vías de señalización, entre ellas el flujo de calcio y la reorganización de las estructuras celulares como son el núcleo y el citoesqueleto<sup>51</sup>. En primer lugar, el núcleo celular se desplaza hacia zonas cercanas al foco de infección, gracias a la reorganización del citoesqueleto que se sitúa a lo largo de toda la célula. Este cambio hace que se forme una estructura pre-IT que servirá de guía para el crecimiento del tubo de infección. De esta manera, los múltiples reordenamientos de microtúbulos, actina, flotilinas y remorinas, todos ellos componentes del citoesqueleto, permiten que se inicie el crecimiento del IT mediante una invaginación de la punta de los pelos radicales. Esta estructura crece hacia dentro en forma de tubo y se extiende por los espacios

intercelulares de los pelos radicales. Gracias a vesículas provenientes del aparato de Golgi que aportan componentes de la pared celular, ésta se va reorganizando y va cubriendo al tubo hasta llegar a las células del córtex, conocidas como células corticales. Una vez que el tubo entra en contacto con estas células carentes de pared celular, se detiene su formación, generándose IT epidérmicos. A continuación, se forman más tubos de infección desde el punto donde ha finalizado el anterior y se desarrollan por el córtex de la raíz hasta llegar a la zona donde se está formando el primordio nodular<sup>51</sup>.

Una vez finalizada la formación de nódulos y el proceso de infección, es necesario que los rizobios se acomoden en el interior de la célula vegetal. La entrada de la bacteria se produce gracias a que la membrana plasmática de la célula vegetal engloba a una o más bacterias hacia el interior formando un simbiosoma<sup>53</sup>. En este punto comienzan a entrar péptidos y moléculas señalizadoras en el simbiosoma obligando a las bacterias a adaptarse al medio y a especializarse, convirtiéndose en bacteroides. Estos bacteroides son capaces de fijar nitrógeno en un entorno carente de oxígeno en el interior celular gracias a los compuestos orgánicos que les aporta la planta, comportándose como un orgánulo más dentro de la célula. De esta forma las bacterias se vuelven dependientes del metabolismo de la planta, y ésta se asegura de que la bacteria le aportará beneficios<sup>51</sup>.

Un trabajo de Bladergroen y colaboradores (2003)<sup>15</sup> donde estudian el proceso de simbiosis entre *R. leguminosarum* RBL5523 y plantas de guisante, representa uno de los primeros estudios del T6SS. Se observó que esta cepa tenía la capacidad de formar nódulos en plantas de guisante a pesar de no ser su hospedador natural, pero estos nódulos eran ineficaces para la fijación de nitrógeno. Posteriormente, los autores procedieron a la selección de una cepa mutante obtenida mediante la inserción de un transposón y denominada RBL5787, capaz de fijar nitrógeno de manera eficaz<sup>15</sup>. Este transposón se localizó en el gen *tssK*, incluido dentro de un clúster que parecía codificar un sistema de secreción y al que los autores denominaron *imp* (del inglés, Impaired in Nodulation). Los autores analizaron las proteínas del sobrenadante de los cultivos bacterianos de la cepa silvestre y mutante y, tras observar las diferencias en el fenotipo de cada cepa llegaron a la conclusión de que estas proteínas estaban implicadas en la formación de nódulos y la fijación de nitrógeno. Estos resultados sugerían que el clúster *imp* codificaba un sistema de secreción cuyos efectores bloqueaban el proceso de infección en plantas de guisante<sup>15</sup>. Actualmente sabemos que el clúster *imp* es un sistema de secreción de tipo VI<sup>54</sup>.

El efecto contrario se ha descrito recientemente en un estudio realizado por Salinero-Lanzarote y sus colaboradores (2019) en *R. etli* Mim1. En este estudio se muestra la capacidad de esta cepa de formar nódulos de forma efectiva en plantas de frijol, mientras que mutantes en los genes que codifican los componentes estructurales del T6SS forman nódulos más pequeños<sup>55</sup>.

Estos dos trabajos muestran que el T6SS en los rizobios podría estar involucrado no sólo en la competición entre bacterias sino en simbiosis y, por tanto, el T6SS podría tener un papel complementario al ya conocido por el T3SS en los procesos de nodulación. El efecto del T6SS podría ser tanto positivo como negativo, dependiendo de si los efectores secretados son reconocidos por la planta hospedadora o no. De tal manera que, si la bacteria secreta efectores a una planta que es su hospedador, se produciría la simbiosis. Mientras que, si esta planta no es su hospedador, ésta reaccionará activando la respuesta inmune para evitar la proliferación bacteriana y por tanto el proceso de simbiosis.

En este estudio trabajamos con el organismo modelo *Sinorhizobium fredii* USDA257, cepa que posee dos T3SS y un T6SS. Esta cepa de crecimiento rápido se aisló por primera vez de plantas de soja<sup>56</sup> y posee un amplio rango de hospedador, pudiendo nodular una gran variedad de géneros de leguminosas. USDA257 es uno de los rizobios más versátiles junto a otro rizobio de su mismo género, *S. fredii* NGR234 que no posee el T6SS. El T3SS de USDA257 ha sido muy bien estudiado y tiene un papel muy importante en la simbiosis y en la determinación del rango de hospedador<sup>50</sup>. Por otro lado, no se conoce nada sobre el papel del T6SS en USDA257 y por ello estudiar este sistema de secreción en este microorganismo es el objetivo general de nuestro trabajo, para el que nos hemos planteado una serie de objetivos concretos.

### **3. OBJETIVOS**

Los objetivos de este proyecto son:

- Estudiar de la distribución filogenética de los T6SS en el orden *Rhizobiales*.
- Identificar y estudiar *in silico* el T6SS de *Sinorhizobium fredii* USDA257.

### **4. MATERIALES Y MÉTODOS**

En este estudio se ha seleccionado la proteína TssB para la construcción del árbol filogenético por ser una de las proteínas estructurales más conservadas del sistema. Para el estudio filogenético de los T6SS en el orden *Rhizobiales*, se han seleccionado un total de 153 cepas perteneciente a doce géneros: *Rhizobium*, *Mesorhizobium*, *Sinorhizobium*,

*Hoeflea*, *Agrobacterium*, *Methylobacterium*, *Microvirga*, *Methylocystis*, *Aminobacter*, *Methylosinus*, *Ensinfer* y *Phyllobacterium*. Mediante el uso de la herramienta BLASTp<sup>57</sup> se obtuvieron las secuencias correspondientes a la proteína TssB del T6SS, y se recopilaron en un archivo fasta (Anexo II). Para la construcción del árbol filogenético, se usó el programa MEGA7<sup>58</sup>. En primer lugar, las 161 secuencias TssB obtenidas de las 153 cepas que componen el estudio, se alinearon con el algoritmo ClustalW. El archivo obtenido se exportó en formato .meg, para la posterior construcción del árbol filogenético. Usando el algoritmo “Maximum Likelihood”<sup>59</sup> y basado en el modelo “JTT matrix” con un valor de 500 bootstrap. De esta manera, se construyó el árbol filogenético de mayor probabilidad en base logarítmica (-13307,37). El árbol inicial para la búsqueda heurística se obtuvo automáticamente aplicando los algoritmos “Neighbor-Join and BioNJ”. Para estimar las distancias de la matriz por parejas se usó el modelo JTT y luego se seleccionó la topología con el valor estadístico más alto. Hubo un total de 237 posiciones en la base de datos final.

El árbol filogenético se exportó en formato .nwk para ser analizado con la herramienta iTOL<sup>60</sup>. Este software se usó para personalizar el árbol, modificando el tamaño de fuente y el tamaño de las ramas, además de los colores asignados para los distintos grupos filogenéticos dentro del árbol.

La segunda parte del estudio constó de un análisis bioinformático del clúster de T6SS en *S. fredii* USDA257. Para su desarrollo se utilizaron distintos softwares, bases de datos y herramientas informáticas. Para el análisis de proteínas se usaron las herramientas BLASTp, SMART<sup>61</sup> y Pfam<sup>62</sup>. La herramienta “Protein Homology/analogy Recognition Engine” (Phyre<sup>2</sup>) fue usada para las predicciones estructurales basadas en homologías<sup>63</sup>. El peso molecular y punto isoeléctrico de cada proteína se calcularon con el software ExPaSy<sup>64</sup>. Mientras que PSORTb<sup>65</sup> se usó para la predicción de la localización celular de las proteínas. La predicción del número de dominios transmembrana fue determinada por TMHMM<sup>66</sup>. Para el alineamiento de las secuencias proteicas se usó el software Clustal Omega<sup>67</sup>. Por último, se utilizó SignalP<sup>68</sup> y PSORTb para la predicción de posibles péptidos señal. Para la creación de figuras se usaron los programas BioRender (<https://biorender.com/>) y Adobe Illustrator CC2019.

## 5. RESULTADOS Y DISCUSIÓN

El T6SS es una nanomáquina bacteriana que secreta toxinas/efectores al interior de células eucariotas y/o procariotas. Este sistema se encuentra tanto en bacterias patógenas



como no patógenas, principalmente en el filo Proteobacterias ( $\alpha$ -,  $\beta$ - y  $\gamma$ -proteobacterias). El último grupo es el mayoritario y mejor estudiado, en el que se encuentran las bacterias del orden *Pseudomonadales*, *Enterobacteriales* y *Xhantomonadales*. Sin embargo, los grupos  $\alpha$ - y  $\beta$ -proteobacterias son minoritarios<sup>54</sup>. Concretamente el grupo de  $\alpha$ -proteobacterias es el de interés en este trabajo, ya que en él se hallan las bacterias del orden *Rhizobiales*. Hasta la fecha se conoce muy poco sobre la función y distribución de los T6SSs en este orden por lo que en este estudio vamos a realizar un análisis filogenético de los T6SSs en *Rhizobiales*, seguido del estudio *in silico* de un clúster de T6SS, concretamente el presente en *S. fredii* USDA257, de interés por su potencial papel en el proceso de simbiosis.

### 5.1. Estudio filogenético de los T6SS en el orden *Rhizobiales*.

Los T6SS presentes en el filo Proteobacteria están distribuidos en cinco grupos filogenéticos denominados Grupo 1, 2, 3, 4 y 5<sup>17</sup>. En este trabajo nos centraremos en los T6SS presentes en el orden *Rhizobiales* para conocer su distribución. En el desarrollo de este estudio se han usado como cepas de referencias de cada grupo filogenético las pertenecientes al orden *Pseudomonadales*, con excepción del grupo 5 cuya cepa de referencia es *A. tumefaciens*.

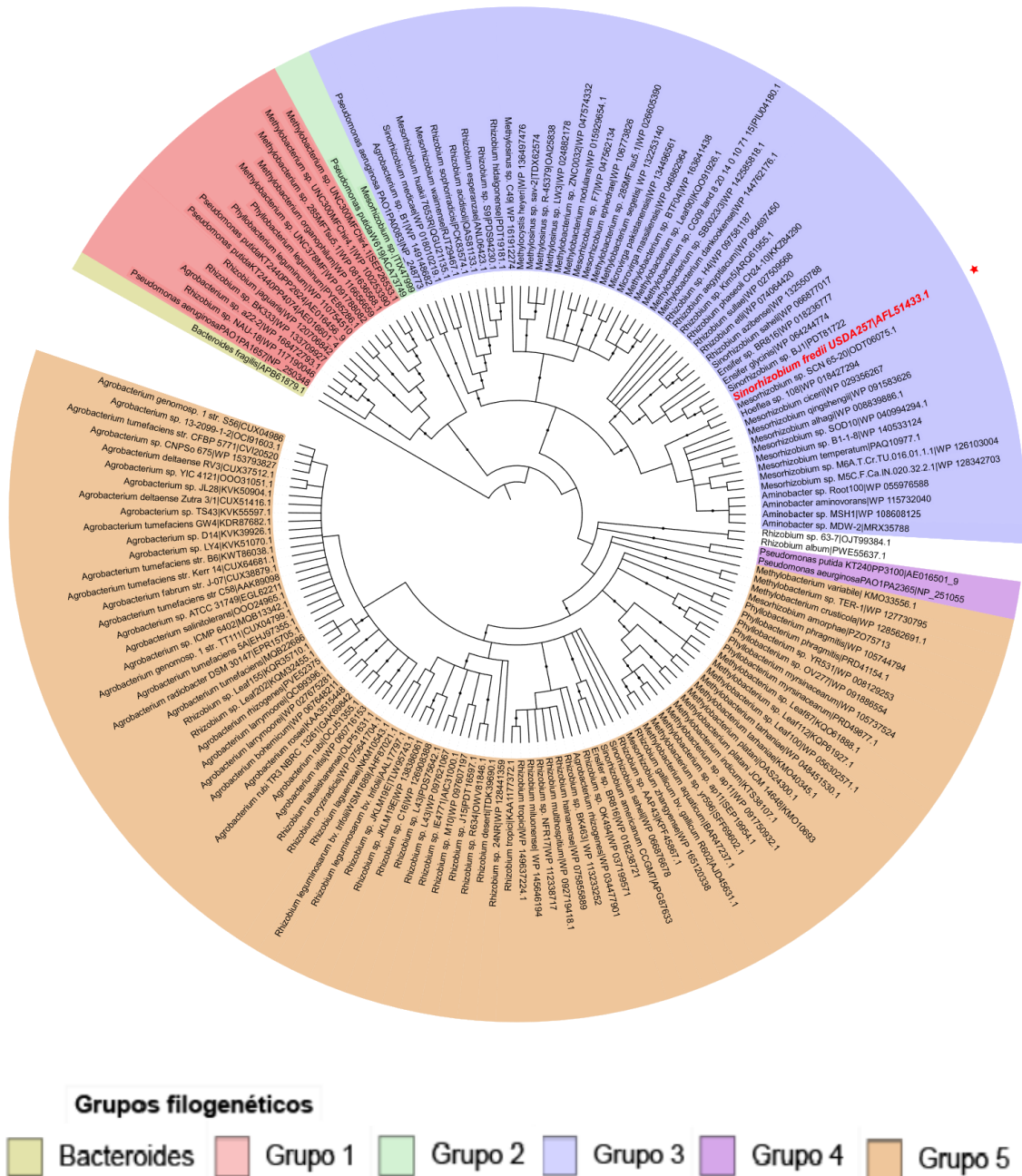
Como se ha explicado en el Apartado 2.2.1, la proteína TssB forma parte de la vaina contráctil del T6SS junto a TssC y está presente en todos los sistemas estudiados<sup>69</sup>. Al ser una de las proteínas más conservadas del sistema, se ha seleccionado para llevar a cabo nuestro estudio filogenético. En este análisis utilizamos 161 secuencias proteicas correspondientes a la proteína TssB del T6SS. Para ello se eligieron un número variable de cepas pertenecientes a distintos géneros del orden *Rhizobiales* (**Tabla 1**). El T6SS de *Bacteroides fragilis*, alejado evolutivamente de los T6SSs de Proteobacterias, se usó como raíz del árbol (**Fig. 3, verde grisáceo**).

**Tabla 1. Listado de los géneros bacterianos incluidos en el estudio filogenético, así como el número de cepas usadas de cada género.**

Género	Número de cepas utilizadas
<i>Microvirga</i>	2
<i>Methylobacterium</i>	29
<i>Bacteroides</i>	1
<i>Methylocystis</i>	1

<i>Methylosinus</i>	4
<i>Aminobacter</i>	4
<i>Mesorhizobium</i>	16
<i>Hoeflea</i>	1
<i>Phyllobacterium</i>	8
<i>Agrobacterium</i>	34
<i>Rhizobium</i>	44
<i>Sinorhizobium</i>	6
<i>Ensifer</i>	3

En la **Fig. 3**, donde se muestra el árbol filogenético de los T6SSs de *Rhizobiales*, podemos observar que la mayoría de éstos pertenecen a los grupos 3 (representado en color azul) con 53 miembros (34.6 %) y 5 (representado en color naranja) con 86 miembros (56.2 %). Por lo que un total de 139 clústeres (91 %) se encuentran distribuidos entre los grupos 3 y 5. Cabe destacar que nuestro organismo modelo *S. fredii* USDA257 (**Fig. 3, indicado con una estrella y en color rojo**) se halla en el grupo 3. Por otro lado, solo se encuentra un miembro del orden *Rhizobiales* en el grupo 2 (**Fig. 3, verde**) y ninguno en el grupo 4 (**Fig. 3, violeta**); el grupo 1 contiene 11 clústeres, es decir un 7.2 % del total (**Fig. 3, rojo**).

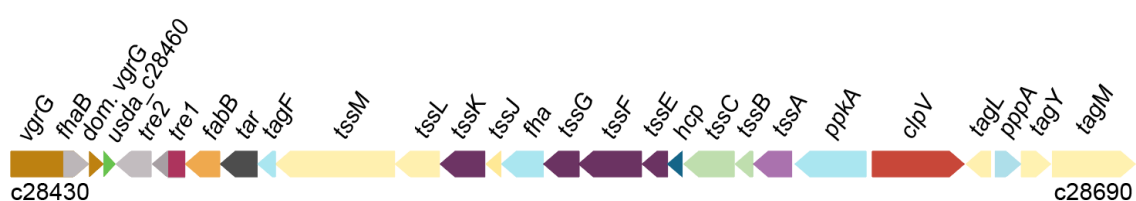


**Figura 3. Árbol filogenético de los T6SSs del orden *Rhizobiales*.** Las ramas que poseen círculos en negro representan un nivel de confianza  $\geq 0.75$ . Se distinguen 5 grandes grupos: grupo 1 (rojo), grupo 2 (verde), grupo 3 (azul), grupo 4 (violeta), grupo 5 (naranja). Bacteria del género *Bacteroides* (verde grisáceo) representa la raíz del árbol. La estrella en rojo corresponde a la cepa en estudio *S. fredii* USDA257.

Para ampliar nuestro conocimiento sobre el papel del T6SS de rizobios y su posible papel en los procesos de simbiosis con leguminosas o la competición interbacteriana, en este trabajo vamos a realizar un estudio *in silico* del T6SS en *S. fredii* USDA257.

## 5.2. Análisis bioinformático del T6SS de *S. fredii* USDA257.

La secuenciación del genoma de *S. fredii* USDA257 permitió identificar dos sistemas de secreción tipo III<sup>70</sup>. El T3SS está ampliamente estudiado en este organismo, jugando un papel esencial en la simbiosis con distintas leguminosas hospedadoras. Las proteínas secretadas a través de este sistema se conocen como Nop (del inglés, Nodulation Outer Protein)<sup>71</sup>. En cambio, el papel ejercido por el T6SS en este microorganismo es aún desconocido. En este trabajo fin de grado hemos identificado por primera vez un clúster que codifica un T6SS completo en este microorganismo (**Fig. 4**) y que pertenece al grupo filogenético 3 (**Fig. 3**).



**Figura 4. Organización del clúster del T6SS en *S. fredii* USDA257.** En color beige se muestran los genes que codifican los componentes del complejo de membrana (*tssM*, *L*, *J*), y genes que codifican proteínas accesorias (*tagL*, *tagY*, *tagM*). Mientras que en color magenta se hallan los genes que codifican los componentes de la placa base (*tssK*, *G*, *F*, *E*) y *tssA* (violeta). En verde claro se identifican los genes *tssB* y *tssC* (vaina contráctil), en azul oscuro el gen *hcp* que codifica la proteína que forma el tubo interno y en azul claro los genes que codifican proteínas reguladoras (*tagF*, *fha*, *ppkA* y *pppA*). En rojo se indica el gen *clpV* que codifica para la ATPasa. Indicado en color gris se encuentran los genes que codifican posibles efectores del sistema (*vgrG*, *tre2* y *tre1*), mientras que el gen que codifica un posible adaptador (*tar*) se indica en gris oscuro. El gen de función desconocida *usda\_c28460* se indica en color verde mientras que el gen *vgrG* y un fragmento del mismo están coloreados en marrón y el *paar* en color rosa.

A través de un detallado análisis bioinformático de este clúster se han identificado veintiséis genes, de los cuales catorce codifican proteínas estructurales que darán lugar al ensamblaje del sistema: complejo de membrana (TssJ, TssM, TssL) (**Fig. 4, beige**), TssA (**Fig. 4, violeta**), placa base (TssE, TssF, TssG, TssK) (**Fig. 4, magenta**), vaina contráctil (TssB, TssC) (**Fig. 4, verde**), tubo interno (Hcp) (**Fig. 4, azul oscuro**), aguja (VrgG (**Fig. 4, marrón**), PAAR (**Fig. 4, rosa**) y ATPasa (ClpV) (**Fig. 4, rojo**).

Muchos de los T6SSs descritos hasta la fecha poseen proteínas reguladoras. Un ejemplo bien estudiado es la pareja kinasa-fosfatasa PpkA-PppA, que media la activación del sistema a través de una cascada de fosforilación que termina con el ensamblaje de los componentes estructurales<sup>38,39</sup>. En el clúster en estudio hemos identificado dos genes que codifican proteínas con actividad kinasa y fosfatasa (PpkA y PppA, respectivamente), y las proteínas TagF, y Fha que interaccionan con el sistema y regulan su actividad (**Fig. 4, azul**). Según los resultados obtenidos, todas ellas son homólogas a proteínas reguladoras

de tipo VI en otros microorganismos (**Tabla 2, Fig. 4**), por lo que cabe esperar que desempeñen la misma función en este sistema y regulen el mismo a nivel post-traducciona

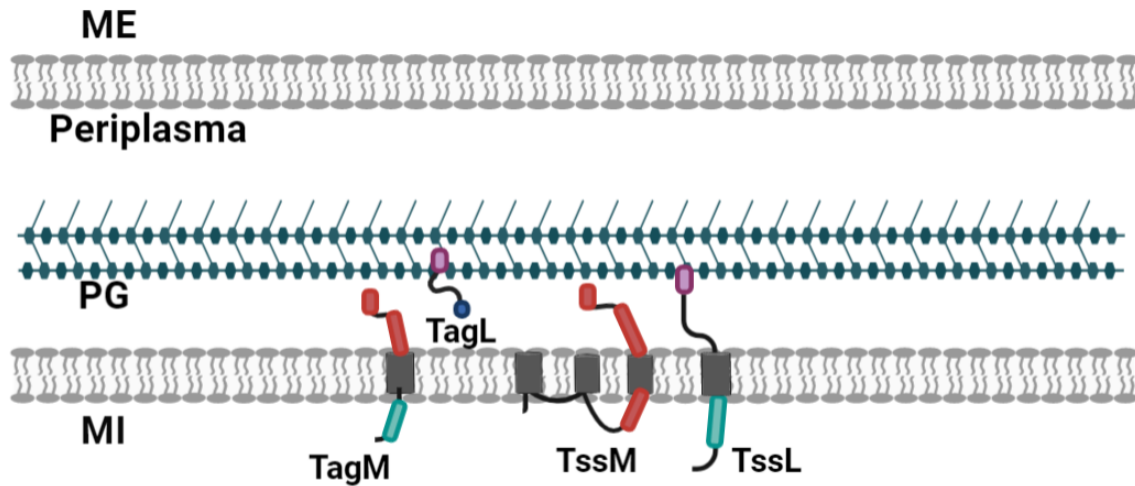
Dentro de los componentes que forman el complejo de membrana, la proteína estructural TssL posee un dominio citosólico (DotU), un dominio transmembrana, y otro periplásmico que en ocasiones puede tener un dominio de unión a peptidoglicano (OmpA) denominándose en este caso TssL especializada<sup>72</sup>. El dominio DotU de TssL (PF09850) es un dominio C-terminal (Ct) que establece una unión fuerte con el dominio citosólico IcmF de TssM y mantiene unidas estas dos proteínas del complejo de membrana<sup>72</sup>. Según el análisis realizado en nuestro estudio, observamos que la proteína del complejo de membrana TssL del T6SS de USDA257 (TssL<sub>U</sub>), posee un dominio DotU de unión a IcmF [intervalo 43-246 aminoácidos (aas)] (**Fig. 5**), por donde interacciona con TssM<sub>U</sub>. TssL<sub>U</sub> posee un dominio transmembrana y en el extremo Ct un dominio de unión a peptidoglicano OmpA (**Fig. 5**) por lo que es una TssL especializada. La predicción de localización celular indica que esta proteína está situada en la membrana interna (**Fig. 5, Tabla 2**). La otra proteína del complejo de membrana, TssM<sub>U</sub>, consta de tres dominios transmembrana (**Fig. 5**) y tres dominios conservados IcmF (COG3523) de unión a DotU (**Fig. 5**). Dos de ellos se encuentran en el dominio amino terminal (Nt) y el tercero en torno al intervalo 400-500 aas de un total de 1176 aas que componen la proteína. Según la predicción de localización, esta proteína se halla en la membrana interna (**Fig. 5, Tabla 2**).

En algunos T6SS también encontramos proteínas accesorias como TagL en *E. coli*<sup>40</sup>. TagL es una proteína estructural con homología a TssM y que al igual que ésta posee un dominio de unión a TssL (IcmF); TagL también contiene un dominio de unión a peptidoglicano (OmpA) que no se encuentra en TssM pero que se encuentra en las TssL especializadas<sup>72</sup>. Tras analizar la proteína TagL de USDA257 (TagL<sub>U</sub>), se observa cierta similitud con la proteína estructural TssL, ya que posee en el extremo Ct terminal el dominio OmpA pero al contrario que TagL de *E. coli*, no tiene dominio IcmF (**Fig. 5**). TagL<sub>U</sub> posee un péptido señal en el extremo Nt, que sitúa a la proteína en el periplasma, y que no se encuentra en la proteína de membrana interna TssL. Según las predicciones realizadas en este estudio, TagL<sub>U</sub> se sitúa en el periplasma (**Fig. 5, Tabla 2**).

En el clúster de USDA257, los dominios IcmF de unión a DotU (TssL) aparecen en la proteína TssM como se indicó anteriormente y también en una proteína específica

de este clúster, que hemos denominado TagM (**Fig. 4, Tabla 2**). La proteína TagM posee dos dominios IcmF (Ct) y un único dominio transmembrana a la altura del aminoácido 150 de la secuencia proteica según las predicciones de los programas TMHMM y Smart (**Fig. 5**). El dominio Nt de TagM es un dominio DotU (intervalo 5-125 aas, SMART), el cual se encuentra en las proteínas TssL del T6SS (**Fig. 5**). TagM presenta la misma probabilidad de localización en todos los compartimentos subcelulares según la herramienta PSORTb (**Tabla 2**), pero dada su topología podríamos esperar que esta proteína se situara en la membrana interna ya que posee un dominio transmembrana, un dominio DotU citosólico y no posee péptido señal para situarla en la membrana externa o en el periplasma (**Fig. 5**). Por tanto, TagM y TagL son proteínas nuevas del T6SS, las cuales poseen dominios comunes a otras proteínas estructurales que forman parte del complejo de membrana, como es el caso del dominio IcmF, de la proteína TssM y el dominio DotU de TssL para TagM y el dominio OmpA de TssL para TagL (**Fig. 5**). TagL y TagM parecen formar parte del complejo de membrana en USDA257 y podrían llevar a cabo funciones de anclaje a la membrana y al peptidoglicano como es el caso de TagL y TagP en *E. coli* y *P. putida*, respectivamente.

Para apoyar los resultados obtenidos, también se realizó un alineamiento de estas proteínas basado en el programa Clustal Omega. Tras realizar el alineamiento entre TssM y TagM (**Fig. 6A**), se observó que gran parte de la secuencia proteica de ambas proteínas presentaba homología (dominio IcmF), a excepción del dominio Nt de TagM, el cual correspondía al dominio DotU. En su lugar, en TssM encontramos dos dominios transmembrana y un dominio IcmF adicional. También se realizó un alineamiento entre TssL y TagM, donde se constató la alta homología existente a nivel del dominio DotU presente en ambas proteínas en sus extremos Nt (**Fig. 6C**). Además, el alineamiento de las proteínas TssL y TagL mostró la similitud entre ambas secuencias ya que poseen en común el dominio OmpA de unión a peptidoglicano (Ct). Este dominio se halla en el intervalo de aminoácidos de 327-429 en la proteína TssL, mientras que en TagL se encuentra en el intervalo 146-242 (**Fig. 6B**).

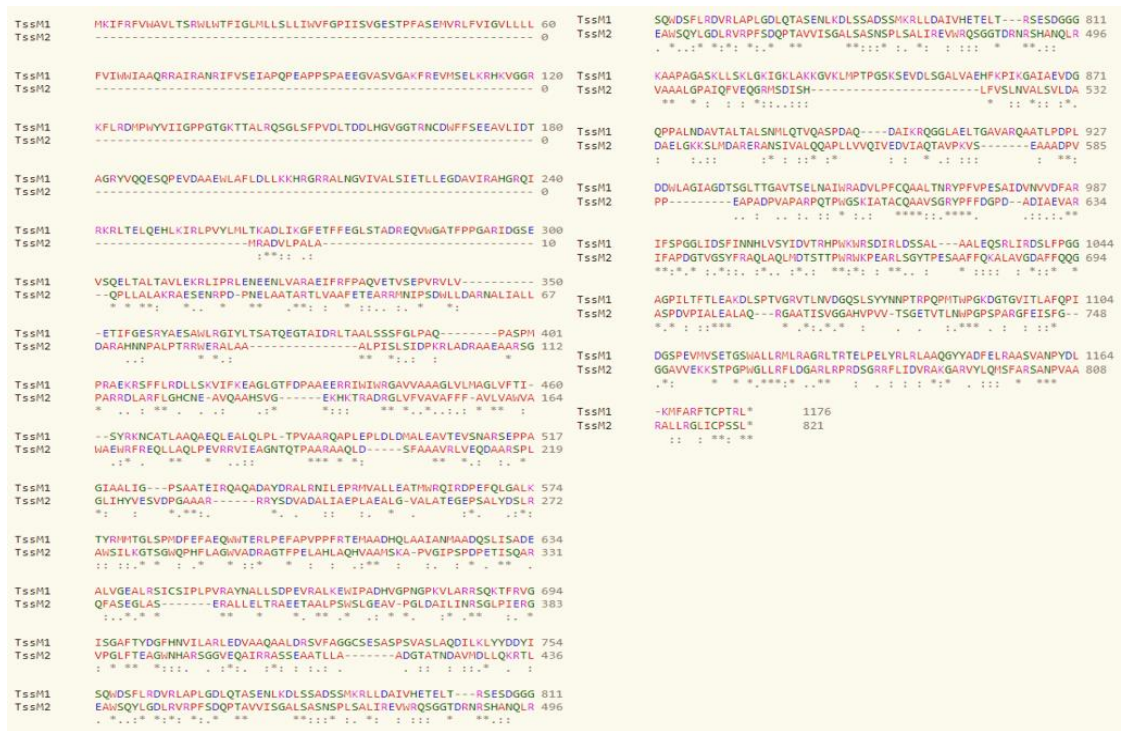


**Figura 5. Representación esquemática de las proteínas estructurales TssM<sub>U</sub>, TssL<sub>U</sub>, y la accesorias TagL y TagM y sus correspondientes dominios.** Dominio citosólico DotU (verde azulado), dominio transmembrana (gris), dominio estructural IcmF (rojo), dominio de unión a peptidoglicano OmpA (magenta) y péptido señal (azul). ME: membrana externa, MI: membrana interna, PG: peptidoglicano.

Como podemos observar, estas cuatro proteínas están interconectadas, ya que poseen dominios en común, e incluso algunas de ellas establecen uniones entre sí en la estructura del sistema. En el caso de TssL<sub>U</sub> y TagM, ambas poseen dominios DotU que interaccionan en el citosol con el dominio IcmF de TssM<sub>U</sub>. Mientras que TagL<sub>U</sub> no se conoce con qué proteína establece interacción, pero al poseer un dominio OmpA al igual que TssL<sub>U</sub>, se espera que ancle el sistema al peptidoglicano.



**A)**



**B)**



**C)**



**Figura 6. Alineamiento múltiple de las secuencias de las proteínas TssM<sub>U</sub>, TagM, TagL<sub>U</sub> y TssL<sub>U</sub>.** Representación del alineamiento de las secuencias proteicas de TssM<sub>U</sub> y TagM (A), de las secuencias de TssL<sub>U</sub> y TagL<sub>U</sub> (B) y de las secuencias de TssL<sub>U</sub> y TagM (C). TssM1 corresponde a TssM<sub>U</sub> y TssM2 corresponde a TagM.

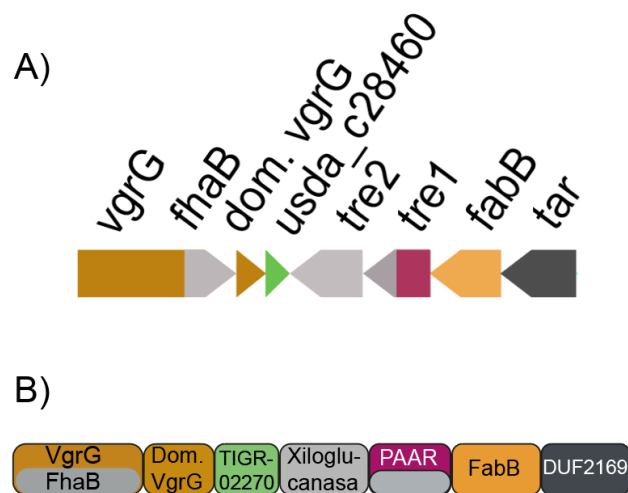


A continuación, entre los genes que codifican las proteínas TagL y TagM observamos un gen que codifica una proteína con función desconocida y que hemos denominado TagY (**Fig. 4, Tabla 2**). Nuestro estudio indica que esta proteína posee dos dominios, el primero situado en el Nt, perteneciente a la superfamilia SlyX (Pfam04102), mientras que en el Ct encontramos un dominio conocido como peptidasa M15\_4 (Pfam13539). El primero de ellos es de función desconocida y no hemos encontrado ninguna predicción funcional con las distintas herramientas utilizadas en este estudio. Según la base de datos Pfam, el dominio Pfam13539 corresponde a una D-alanil D-alanina carboxipeptidasa<sup>73</sup>. A pesar de que su localización celular no se conoce, las predicciones indican que posee un dominio transmembrana en el extremo Nt y un péptido señal, por lo que esta proteína parece encontrarse anclada a la membrana de la bacteria, con el extremo Ct hacia el periplasma o hacia el exterior celular. Hasta la fecha sólo se ha descrito un clúster de T6SS que codifica una peptidasa<sup>74</sup>, hallada en *Acinetobacter baumannii*. Esta enzima es una endopeptidasa que digiere el peptidoglicano de la pared bacteriana, concretamente ejerce su actividad sobre los enlaces L-ala-D-glu y se denomina TagX<sup>74</sup>. Esta enzima forma un poro en la membrana bacteriana que permite la correcta formación y actuación del T6SS<sup>74</sup>. De manera similar a TagX, TagY podría estar involucrada en la rotura local del peptidoglicano a través de sus enlaces D-alanina-D-alanina para la correcta formación del T6SS en *S. fredii* USDA257.

En los clústeres de los T6SS se pueden encontrar genes que codifican proteínas efectoras<sup>32</sup>. En el T6SS de USDA257, encontramos un gen que codifica una proteína VgrG evolucionada (VgrG<sub>e</sub>) (**Fig. 7A, marrón**). Una proteína VgrG<sub>e</sub> es la que posee en su dominio Ct un efector codificado en el propio gen<sup>34</sup>. El dominio Ct de VgrG<sub>e</sub> de USDA257 tiene homología con un dominio que se encuentra en proteínas pertenecientes a la superfamilia FhaB y que están implicadas en adhesión y en la utilización del grupo hemo, así como en el transporte y la secreción vesicular. En nuestro caso, la extensión de VgrG<sub>e</sub> es de 746 aas mientras que el dominio FhaB tiene una longitud de 225 aas. Los datos obtenidos nos indican que este dominio sólo corresponde con un fragmento de la proteína FhaB completa, la cual es precursora de filamentos de hemaglutinina (FHA), uno de los factores de virulencia de *Bordetella pertussis*<sup>75</sup>. FhaB ejerce de chaperona previniendo el plegamiento prematuro de FHA<sup>75</sup>. Este dominio Ct presenta además homología estructural con la proteína Cthe\_2159 procedente de *Hungateiclostridium*

*thermocellum* ATCC 27405 según nuestro análisis estructural realizado con el software Phyre<sup>2</sup>. Observamos un alineamiento del 85% entre el dominio Ct de VgrG<sub>e</sub> y la proteína Cthe\_2159 y una confianza del 86.5% (**Tabla 2**). Según un estudio realizado por Close y colaboradores (2014)<sup>76</sup>, Cthe\_2159 tiene características similares a las enzimas polisacárido liasas, entre ellas la capacidad de unirse a carbohidratos, concretamente celulosa y ácido poligalacturónico, aunque no se ha determinado la actividad enzimática de esta proteína<sup>76</sup>. La similitud del dominio Ct de VgrG<sub>e</sub> con la proteína Cthe\_2159, indica que éste podría estar implicado en la unión a polisacáridos, (celulosa y/o ácido poligalacturónico), ambos componentes de la pared celular vegetal. Estos resultados indican que este efector especializado podría estar involucrado en la adhesión y/o degradación de la pared celular vegetal, pasos imprescindibles en las etapas iniciales de la infección del proceso de nodulación.

En cuanto a los genes que se hallan aguas abajo de *vgrG<sub>e</sub>* en el clúster, los primeros suelen codificar posibles adaptadores en el caso de que los haya, y a continuación se encuentran los genes que codifican los efectores/toxinas del sistema. En nuestro organismo modelo podemos observar cómo la estructura de esta región del clúster está invertida, ya que los genes que codifican posibles efectores (*ter1* y *ter2*) se hallan aguas abajo de *vgrG* y aguas arriba del gen que codifica el adaptador (*tar*) (**Fig. 7A**).



**Figura 7. A)** Organización de los genes aguas abajo de *vgrG* en el clúster del T6SS de *S. fredii* USDA257. Gen *vgrG* y un fragmento del mismo (marrón) y los genes para posibles efectores (gris). *Usda\_c28460* (verde), *paar* (magenta), *fabB* (naranja) y el gen para un posible adaptador (gris oscuro). **B)** Proteínas codificadas por los genes representados en A) y sus correspondientes dominios: proteína VgrG y un fragmento de la misma (marrón), dominio con función efectora FhaB (gris), xiloglucanasa (gris), dominio TIGR02270 con función desconocida (verde), proteína PAAR<sub>e</sub> (magenta) con un dominio Ct de función desconocida (gris), proteína FabB (naranja) y adaptador (gris oscuro).

El gen USDA257\_c28500 codifica una proteína con un dominio perteneciente a la superfamilia DUF2169 (**Fig. 7B**). Esta proteína ha sido identificada en estudios anteriores como un adaptador del T6SS<sup>77,78</sup>. El dominio pertenece a una familia de chaperonas esenciales para el transporte de efectores tipo cargo, en este caso concreto, los dirigen hacia la proteína VgrG a la cual se unirán los efectores para su posterior secreción<sup>78</sup>. En nuestro estudio, hemos denominado a esta proteína Tar del inglés, *Type six adaptor rhizobium* (**Fig. 7A, Tabla 2**).

El gen USDA257\_c28490 codifica una proteína con un dominio perteneciente a la superfamilia PRK06147. Este dominio se encuentra en proteínas transportadoras de grupos acilo, y concretamente en una proteína denominada 3-oxoacil sintasa a la que hemos denominado FabB (**Fig. 7B, Tabla 2**). Recientemente se ha identificado una proteína con actividad tiolasa, también implicada en el metabolismo de lípidos, en un clúster VgrG huérfano de *P. aeruginosa*<sup>77</sup> con una estructura génica similar a la descrita en USDA257. Aún no se conoce la función que desempeñan estas proteínas en el sistema, pero sí parecen ser unas proteínas que se conservan en ciertos T6SS y están ligadas a la síntesis de VgrG, adaptadores tipo DUF2169 y efectores.

USDA257\_c28480 codifica una proteína que posee un dominio PAAR en el extremo Nt y es por tanto una PAAR<sub>e</sub> (proteína que posee en su extremo Ct un efector codificado por el propio gen) a la que hemos denominado Tre1 (*Type six rhizobium effector 1*) (**Fig. 7A**). El dominio Ct es un dominio efector con función desconocida (**Fig. 7B**), los estudios llevados a cabo en este trabajo no han podido revelar ninguna predicción sobre la función de este efector.

La proteína codificada por el gen USDA257\_c28470 ha sido denominada Tre2 (*Type six rhizobium effector 2*) por su posible papel como efector del sistema. La proteína presenta homología estructural con la enzima GH74 (glicósido hidrolasa) procedente de *Caldicellulosiruptor lactoaceticus 6A* (**Tabla 2**). El alineamiento cubre un 69% de la proteína del aminoácido 100 al 344, y tiene una confianza del 99.3%. GH74 posee actividad hidrolasa con un dominio de unión a xiloglucanos (componente principal de la hemicelulosa) (**Fig. 7B**), capaz de hidrolizar la pared vegetal<sup>79</sup>. Estos resultados indican que Tre2 podría estar implicada en la degradación de los componentes de la pared vegetal. Los efectores que llevan a cabo la degradación de la pared celular vegetal durante la etapa de infección del proceso de nodulación para la creación del tubo de infección en el

proceso de simbiosis, son normalmente secretados por el T3SS<sup>80</sup>. Este estudio presenta los primeros datos que indican que estos efectores pueden ser secretados también a través del T6SS (**Fig. 7A, Tabla 2**).

Por último, el análisis de la secuencia proteica codificada por el gen USDA257\_c28460 muestra un dominio de la familia TIGR02270 de función desconocida (**Fig. 7B**). Se requiere una investigación experimental para conocer en profundidad el papel que desempeñan las proteínas accesorias y los posibles efectores del sistema.

**Tabla 2. Listado de genes que forman el clúster de T6SS de *S. fredii* USDA257.**

Nombre del locus	Nombre de la proteína	Dominios conservados COG/pfam/TIGR (Abreviatura)/ Phyre <sup>2</sup> (P)	Predicción de localización celular
USDA257_C28440	VgrGe	COG3501 (VgrG) + COG3210 (FhaB)/ pfam06715 (Gp5_C)	C
USDA257_C28450	Dominio VgrG	COG3501 (VgrG)/ TIGR01646 (vrg_GE)	D
USDA257_C28460		TIGR02270	C
USDA257_C28470	Tre2	P: GH74 (100-344 a.a) C: 99.3%	C
USDA257_C28480	Tre1	Pfam13665 (DUF4150)	C
USDA257_C28490	FabB	COG0304 (FabB)	MI
USDA257_C28500	Tar	DUF2169	C
USDA257_C28510	TagF	COG3913 (SciT)	D
USDA257_C28520	TssM	COG3523 (IcmF)	MI
USDA257_C28530	TssL	COG3455 (DotU) + pfam00691 (OmpA)	ME
USDA257_C28540	TssK	COG3522	C
USDA257_C28550	TssJ	COG3521	MI
USDA257_C28560	Fha	COG3456	C
USDA257_C28570	TssG	COG3520	C
USDA257_C28580	TssF	COG3519	C
USDA257_C28590	TssE	COG3518	C
USDA257_C28600	Hcp	COG3157 (Hcp)	C
USDA257_C28610	TssC	COG3517	C
USDA257_C28620	TssB	COG3516	C
USDA257_C28630	TssA	COG3515 + TIGR03363	C
USDA257_C28640	PpkA	COG0515 (SPS1)	P y PC (2ª localización)
USDA257_C28650	ClpV	COG0542 (ClpA) + pfam07724 (AAA_2)	C
USDA257_C28660	TagL	COG2885 (OmpA)	ME

USDA257_C28670	PppA	COG0631 (PTC) + TIGR02865 (spore_II_E)	C
USDA257_C28680	TagY	Pfam04102 (SlyX) + pfam13539 (peptidasa_M15_4)	D
USDA257_C28690	TagM	COG3523 (IcmF) + COG3455 (DotU)	D

En la tabla se indica el nombre de cada locus, el nombre de la proteína codificada por ese locus, la predicción de los dominios que posee cada proteína, así como la predicción de la localización celular de cada una. ME: membrana externa, MI: membrana interna, C: citoplasmática, P: periplasma, PC: pared celular, D: desconocida, aa: aminoácido. En el Anexo I se puede encontrar una versión extendida de esta tabla.

## 6. CONCLUSIONES

En este trabajo se ha llevado a cabo un estudio filogenético de los T6SSs en rizobios, donde se ha determinado que éstos se encuentran principalmente en los grupos filogenéticos 3 y 5. En ambos grupos encontramos especies bacterianas relacionadas con plantas, sin embargo, la composición de cada uno es muy diferente. Mientras que el grupo 3 está formado por especies patógenas y algunas simbióticas, el grupo 5 incluye mayoritariamente especies simbióticas y por tanto beneficiosas para las plantas. Posteriormente, hemos llevado a cabo un análisis detallado del clúster del T6SS de *S. fredii* USDA257 (grupo filogenético 3) el cual codifica las proteínas estructurales conservadas (TssABCDEFGHIJKLM) así como proteínas accesorias conocidas (TagF, PpkA, PppA, Fha) y proteínas propias que no figuran en otros T6SS, como el caso de TagM, TagL y TagY. Los resultados de este estudio indican que los efectores identificados en el T6SS de USDA257 (VgrG<sub>e</sub>, Tre1 y Tre2) podrían estar involucrados en la degradación de celulosa y otros componentes de la pared celular vegetal y por tanto implicados en el proceso de simbiosis ayudando a la bacteria en la formación del tubo de infección. De esta forma, el T6SS de USDA257 cooperaría con el T3SS presente en este microorganismo en el establecimiento del proceso de simbiosis con leguminosas.

## 7. BIBLIOGRAFÍA

1. Costa, T. R. D. *et al.* Secretion systems in Gram-negative bacteria: Structural and mechanistic insights. *Nat. Rev. Microbiol.* **13**, 343–359 (2015).
2. Houben, E. N. G., Korotkov, K. V. & Bitter, W. Take five - Type VII secretion systems of Mycobacteria. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843**, 1707–1716 (2014).
3. Cao, Z., Casabona, M. G., Kneuper, H., Chalmers, J. D. & Palmer, T. The type VII secretion system of *Staphylococcus aureus* secretes a nuclease toxin that targets competitor bacteria. *Nat.*

- Microbiol.* **2**, 16183 (2016).
4. Lenders, M. H. H., Beer, T., Smits, S. H. J. & Schmitt, L. In vivo quantification of the secretion rates of the hemolysin A Type I secretion system. *Sci. Rep.* **6**, 1–10 (2016).
  5. Masi, M. & Wandersman, C. Multiple signals direct the assembly and function of a type I secretion system. *J. Bacteriol.* **192**, 3861–3869 (2010).
  6. Espinosa-Urgel, M., Salido, A. & Ramos, J. L. Genetic analysis of functions involved in adhesion of *Pseudomonas putida* to seeds. *J. Bacteriol.* **182**, 2363–2369 (2000).
  7. Yang, J. *et al.* Mechanistic Insights into Elastin Degradation by Pseudolysin, the Major Virulence Factor of the Opportunistic Pathogen *Pseudomonas aeruginosa*. *Sci. Rep.* **5**, 1–7 (2015).
  8. Reichow, S. L., Korotkov, K. V., Hol, W. G. J. & Gonen, T. Structure of the cholera toxin secretion channel in its closed state. *Nat. Struct. Mol. Biol.* **17**, 1226–1232 (2010).
  9. Deslandes, L. & Rivas, S. Catch me if you can: Bacterial effectors and plant targets. *Trends Plant Sci.* **17**, 644–655 (2012).
  10. López-Baena, F. J., Ruiz-Sainz, J. E., Rodríguez-Carvajal, M. A. & Vinardell, J. M. Bacterial molecular signals in the *Sinorhizobium fredii*-soybean symbiosis. *Int. J. Mol. Sci.* **17**, 755 (2016).
  11. Aguilar, J., Zupan, J., Cameron, T. A. & Zambryski, P. C. *Agrobacterium* type IV secretion system and its substrates form helical arrays around the circumference of virulence-induced cells. *PLoS* **107**, 1–6 (2010).
  12. Leo, J. C., Grin, I. & Linke, D. Type V secretion: Mechanism(S) of autotransport through the bacterial outer membrane. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 1088–1101 (2012).
  13. Henderson, I. A. N. R. & Nataro, J. P. MINIREVIEW Virulence Functions of Autotransporter Proteins. *Society* **69**, 1231–1243 (2001).
  14. Kikuchi, T., Mizunoe, Y., Takade, A., Naito, S. & Yoshida, S. I. Curli fibers are required for development of biofilm architecture in *Escherichia coli* K-12 and enhance bacterial adherence to human uroepithelial cells. *Microbiol. Immunol.* **49**, 875–884 (2005).
  15. Bladergroen, M. R., Badelt, K. & Spaink, H. P. Infection-blocking genes of a symbiotic *Rhizobium leguminosarum* strain that are involved in temperature-dependent protein secretion. *Mol. Plant-Microbe Interact.* **16**, 53–64 (2003).
  16. Pukatzki, S. *et al.* Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 1528–1533 (2006).
  17. Boyer, F., Fichant, G., Berthod, J., Vandenbrouck, Y. & Attree, I. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: What can be learned from available microbial genomic resources? *BMC Genomics* **10**, 104 (2009).
  18. Basler, M., Ho, B. T. & Mekalanos, J. J. Tit-for-tat: Type VI secretion system counterattack during bacterial cell-cell interactions. *Cell* **152**, 884–894 (2013).
  19. Ma, A. T. & Mekalanos, J. J. In vivo actin cross-linking induced by *Vibrio cholerae* type VI secretion system is associated with intestinal inflammation. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 4365–4370 (2010).
  20. Cascales, E. The type VI secretion toolkit. *EMBO Rep.* **9**, 735–741 (2008).

21. Shalom, G., Shaw, J. G. & Thomas, M. S. In vivo expression technology identifies a type VI secretion system locus in *Burkholderia pseudomallei* that is induced upon invasion of macrophages. *Microbiology* **153**, 2689–2699 (2007).
22. Silverman, J. M., Brunet, Y. R., Cascales, E. & Mougous, J. D. Structure and Regulation of the Type VI Secretion System. *Annu. Rev. Microbiol.* **66**, 453–472 (2012).
23. Felisberto-Rodrigues, C. *et al.* Towards a structural comprehension of bacterial type vi secretion systems: Characterization of the TssJ-TssM complex of an *Escherichia coli* pathovar. *PLoS Pathog.* **7**, 1–11 (2011).
24. Cherrak, Y. *et al.* Biogenesis and structure of a type VI secretion baseplate. *Nat. Microbiol.* **3**, 1404–1416 (2018).
25. Leiman, P. G. Secretion System Spike. *Nature* **500**, 350–353 (2014).
26. Zoued, A. *et al.* Priming and polymerization of a bacterial contractile tail structure. *Nature* **531**, 59–63 (2016).
27. Bönemann, G., Pietrosiuk, A. & Mogk, A. Tubules and donuts: A type VI secretion story: MicroReview. *Mol. Microbiol.* **76**, 815–821 (2010).
28. Brackmann, M., Wang, J. & Basler, M. Type VI secretion system sheath inter-subunit interactions modulate its contraction. *EMBO Rep.* **19**, 225–233 (2018).
29. Christie, P. J. The Rich Tapestry of Bacterial Protein Translocation Systems. *Protein J.* **38**, 389–408 (2019).
30. Cianfanelli, F. R., Monlezun, L. & Coulthurst, S. J. Aim, Load, Fire: The Type VI Secretion System, a Bacterial Nanoweapon. *Trends Microbiol.* **24**, 51–62 (2016).
31. Kapitein, N. *et al.* ClpV recycles VipA/VipB tubules and prevents non-productive tubule formation to ensure efficient type VI protein secretion. *Mol. Microbiol.* **87**, 1013–1028 (2013).
32. Durand, E., Cambillau, C., Cascales, E. & Journet, L. VgrG, Tae, Tle, and beyond: The versatile arsenal of Type VI secretion effectors. *Trends Microbiol.* **22**, 498–507 (2014).
33. Alcoforado Diniz, J., Liu, Y. C. & Coulthurst, S. J. Molecular weaponry: Diverse effectors delivered by the Type VI secretion system. *Cell. Microbiol.* **17**, 1742–1751 (2015).
34. Pukatzki, S., McAuley, S. B. & Miyata, S. T. The type VI secretion system: translocation of effectors and effector-domains. *Curr. Opin. Microbiol.* **12**, 11–17 (2009).
35. Hachani, A., Wood, T. E. & Filloux, A. Type VI secretion and anti-host effectors. *Curr. Opin. Microbiol.* **29**, 81–93 (2016).
36. Trunk, K. *et al.* The type VI secretion system deploys antifungal effectors against microbial competitors. *Nat. Microbiol.* **3**, 920–931 (2018).
37. Wang, J., Brodmann, M. & Basler, M. Assembly and Subcellular Localization of Bacterial Type VI Secretion Systems. *Annu. Rev. Microbiol.* **73**, 621–638 (2019).
38. Mougous, J. D., Gifford, C. A., Ramsdell, T. L. & Mekalanos, J. J. Threonine phosphorylation post-translationally regulates protein secretion in *Pseudomonas aeruginosa*. *Nat. Cell Biol.* **9**, 797–803 (2007).
39. Lin, J. S. *et al.* TagF-mediated repression of bacterial type VI secretion systems involves a direct interaction with the cytoplasmic protein Fha. *J. Biol. Chem.* **293**, 8829–8842 (2018).

40. Santin, Y. G. *et al.* Role and recruitment of the TAGL peptidoglycan-binding protein during type VI secretion system biogenesis. *J. Bacteriol.* **201**, 1–12 (2019).
41. Aschtgen, M. S., Thomas, M. S. & Cascales, E. Anchoring the type VI secretion system to the peptidoglycan TssL, TagL, TagP,... what else? *Virulence* **1**, 535–540 (2010).
42. Leung, T. L. F. & Poulin, R. Parasitism, commensalism, and mutualism: Exploring the many shades of symbioses. *Vie Milieu* **58**, 107–115 (2008).
43. Compant, S., Samad, A., Faist, H. & Sessitsch, A. A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. *J. Adv. Res.* **19**, 29–37 (2019).
44. Rout, M. E. The Plant Microbiome. *Adv. Bot. Res.* **69**, 279–309 (2014).
45. Levy, A., Conway, J. M., Dangl, J. L. & Woyke, T. Elucidating Bacterial Gene Functions in the Plant Microbiome. *Cell Host Microbe* **24**, 475–485 (2018).
46. Bloemberg, G. V. & Lugtenberg, B. J. J. Molecular basis of plant growth promotion and biocontrol by rhizobacteria. *Curr. Opin. Plant Biol.* **4**, 343–350 (2001).
47. Chng, S. *et al.* Take-all decline in New Zealand wheat soils and the microorganisms associated with the potential mechanisms of disease suppression. *Plant Soil* **397**, 239–259 (2015).
48. Cha, J. Y. *et al.* Microbial and biochemical basis of a Fusarium wilt-suppressive soil. *ISME J.* **10**, 119–129 (2016).
49. Chang, J. H., Desveaux, D. & Creason, A. L. The ABCs and 123s of Bacterial Secretion Systems in Plant Pathogenesis. *Annu. Rev. Phytopathol.* **52**, 317–345 (2014).
50. Staehelin, C. & Krishnan, H. B. Review Article: Nodulation outer proteins: Double-edged swords of symbiotic Rhizobia. *Biochem. J.* **470**, 263–274 (2015).
51. Oldroyd, G. E. D., Murray, J. D., Poole, P. S. & Downie, J. A. The Rules of Engagement in the Legume-Rhizobial Symbiosis. *Annu. Rev. Genet.* **45**, 119–144 (2011).
52. Zipfel, C. & Oldroyd, G. E. D. Plant signalling in symbiosis and immunity. *Nature* **543**, 328–336 (2017).
53. Long, S. R. SnapShot: Signaling in Symbiosis. *Cell* **167**, 582–582 (2016).
54. Bernal, P. & Filloux, A. Minireview Type VI secretion systems in plant-associated bacteria. *Environmental Microbiology* **20**, 1–15 (2018).
55. Salinero-Lanzarote, A. *et al.* The type VI secretion system of *Rhizobium etli* Mim1 has a positive effect in symbiosis. *FEMS Microbiology Ecology* **95**, 5 (2019).
56. Keyser, H. H., Bohlool, B. B., Hu, T. S. & Weber, D. F. Fast-Growing Rhizobia Isolated from Root Nodules of Soybean. *Science* **215**, 1631–1632 (1982).
57. Boratyn, G. M. *et al.* BLAST: a more efficient report with usability improvements. *Nucleic Acids Res.* **41**, 29–33 (2013).
58. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
59. Le, S. Q., Lartillot, N. & Gascuel, O. Phylogenetic mixture models for proteins. *Philos. Trans. R. Soc. B Biol. Sci.* **363**, 3965–3976 (2008).
60. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).



61. Letunic, I., Doerks, T. & Bork, P. SMART: Recent updates, new developments and status in 2015. *Nucleic Acids Res.* **43**, D257–D260 (2015).
62. Finn, R. D. *et al.* The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **44**, D279–D285 (2016).
63. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
64. Artimo, P. *et al.* ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* **40**, 597–603 (2012).
65. Yu, N. Y. *et al.* PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26**, 1608–1615 (2010).
66. Krogh, A., Larsson, B., Von Heijne, G. & Sonnhammer, E. L. L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
67. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
68. Almagro Armenteros, J. J. *et al.* SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423 (2019).
69. Cascales, E. & Cambillau, C. Structural biology of type VI secretion systems. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 1102–1111 (2012).
70. Schuldes, J. *et al.* Complete genome sequence of the broad-host-range strain *Sinorhizobium fredii* USDA257. *J. Bacteriol.* **194**, 4483–4483 (2012).
71. Marie, C. *et al.* Characterization of Nops, nodulation outer proteins, secreted via the type III secretion system of NGR234. *Mol. Plant-Microbe Interact.* **16**, 743–751 (2003).
72. Durand, E. *et al.* Structural characterization and oligomerization of the TssL protein, a component shared by bacterial type VI and type IVb secretion systems. *J. Biol. Chem.* **287**, 14157–14168 (2012).
73. Evers, S. & Courvalin, P. Regulation of VanB-type vancomycin resistance gene expression by the VanSB-VanRB two-component regulatory system in *Enterococcus faecalis* V583. *J. Bacteriol.* **178**, 1302–1309 (1996).
74. Weber, B. S. *et al.* Genetic dissection of the type VI secretion system in *Acinetobacter* and identification of a novel peptidoglycan hydrolase, TagX, required for its biogenesis. *MBio* **7**, 1–17 (2016).
75. Coutte, L., Antoine, R., Drobecq, H., Loch, C. & Jacob-Dubuisson, F. Subtilisin-like autotransporter serves as maturation protease in a bacterial secretion pathway. *EMBO J.* **20**, 5040–5048 (2001).
76. Close, D. W., D’angelo, S. & Bradbury, A. R. M. A new family of  $\beta$ -helix proteins with similarities to the polysaccharide lyases. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **70**, 2583–2592 (2014).
77. Pissaridou, P., Allsopp, L. P., Wettstadt, S., Howard, S. A. & Mavridou, D. A. I. The

- Pseudomonas aeruginosa* T6SS-VgrG1b spike is topped by a PAAR protein eliciting DNA damage to bacterial competitors. *PNAS* **115**, 12519-12524 (2018).
78. Lien, Y. W. & Lai, E. M. Type VI secretion effectors: Methodologies and biology. *Front. Cell. Infect. Microbiol.* **7**, 1–11 (2017).
79. Arnal, G. *et al.* Substrate specificity, regiospecificity, and processivity in glycoside hydrolase family 74. *J. Biol. Chem.* **294**, 13233–13247 (2019).
80. Nelson, M. S. & Sadowsky, M. J. Secretion systems and signal exchange between nitrogen-fixing rhizobia and legumes. *Front. Plant Sci.* **6**, 1–11 (2015).