



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

**FACULTAD DE QUÍMICA**

**ANÁLISIS MOLECULAR DE LA REGIÓN PROMOTORA  
DEL GEN *NASTEP* DE *NICOTIANA ALATA***

**T E S I S**  
QUE PARA OBTENER EL TÍTULO DE :  
**QUÍMICA FARMACÉUTICA BIÓLOGA**  
**P R E S E N T A :**  
**CAROLINA SOTO PALMA**



**MÉXICO, D.F.**

**2011**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## **JURADO ASIGNADO:**

<b>PRESIDENTE:</b>	<b>Prof. J. Eleazar Martínez Barajas</b>
<b>VOCAL:</b>	<b>Prof. Felipe Cruz García</b>
<b>SECRETARIO:</b>	<b>Prof. Francisco Javier Plasencia de la Parra</b>
<b>1er. SUPLENTE:</b>	<b>Prof. Martha Patricia Coello Coutiño</b>
<b>2° SUPLENTE:</b>	<b>Prof. Tzvetanka Dimitrova Dinkova</b>

El presente trabajo se desarrolló en el Laboratorio 104 del Departamento de Bioquímica del Conjunto E de la Facultad de Química de la Universidad Nacional Autónoma de México, bajo dirección del Dr. Felipe Cruz García y con apoyos de los proyectos:

- **CONACYT.** Número de proyecto: 81968
- **DGAPA.** Número de proyecto: IN205009
- **PAIP.** Número de proyecto: 629015

**ASESOR:** **DR. FELIPE CRUZ GARCÍA**

**SUSTENTANTE :** **CAROLINA SOTO PALMA**

## **Agradecimientos**

A mis padres, mis primeros y constantes maestros, quienes además me brindaron la oportunidad de cumplir mi sueño de licenciarme en la UNAM. Agradezco a mi madre por ser quien me inculcó los valores de la constancia y disciplina en el estudio, y a mi padre por mostrarme que cuando algo se entiende es fácil de explicar (aún cuando se trate de termodinámica y mecánica cuántica). Sepan los amo y admiro.

Al Dr. Felipe Cruz, que mediante la pasión y belleza de sus clases, redescubrió en mi el gusto e interés por la genética, y en general, por la ciencia. Y quien además, como mi tutor, me otorgó siempre su confianza, apoyo y conocimiento.

A mi tía Ana por brindarme un hogar, re-educar mis gustos y apoyarme como madre y amiga durante toda esta etapa.

A mis dos hermanas, Itzel y Mari, por ser el gran motivo e inspiración de mi vida. Las amo.

A Josué y a Pati, por ser indispensables en mi primera etapa de libertad substancial. A Bernardo, por los momentos vividos. A Rous, por haber sido la que inició la conversación entre las dos, por su solidaridad en los momentos difíciles, y por la construcción continua de recuerdos memorables. Y a Gabriel, mi gran amigo, por que desde el día que lo reencontré ha estado ahí.

A la Facultad de Ciencias, en especial a la carrera de Biología, que me abrió sus puertas y me estimuló, mediante la fotografía y la evolución, a mirar desde un ángulo distinto.

Gracias Leti por tu alegría y calidez. Gracias Nestor por ser la persona que eres, comprometido y genuino. Gracias Neto por tu amistad. Convirtieron del espacio y cotidianidad compartida, algo verdaderamente único.

A Clau por sus enseñanzas, y a Flor, Lili y Yu por brindarme una sonrisa. A mis compañeros del laboratorio 103 y en especial a Carlos, por siempre tener la disposición de discutir conmigo los resultados y proponer nuevas ideas al proyecto.

A ER, por que eres tú, la imagen visible aunque cerréis mis ojos, y la canción que oigo, aunque tapéis mis oídos. Y por que tu amor es un fragmento del corazón de la Vida.

A la Cineteca Nacional, simplemente por que en ella no pasa el tiempo!

**Agradecemos a la Facultad de Química de la Universidad  
Nacional Autónoma de México, por el apoyo otorgado para la  
realización de esta tesis a través del Subprograma 127  
Formación Básica en Investigación.**

# ÍNDICE

RESUMEN .....	1
<b>1. INTRODUCCIÓN</b>	
1.1 La regulación de la expresión genética en las plantas .....	2
1.2 El sistema de incompatibilidad como un mecanismo en el cual, la regulación de la expresión genética juega un papel importante .....	2
1.2.1 Función y expresión de las determinantes, femenina y masculina, del locus <i>S</i> .....	3
1.2.2 Genes modificadores .....	4
1.2.3 <i>NaStEP</i> como un gen modificador en <i>N. alata</i> y cuyo patrón de expresión es único .....	6
1.3 La expresión genética tiene diferentes niveles de regulación .....	8
1.4 Estructura y características de un promotor .....	8
1.4.1 El promotor mínimo ó <i>core promoter</i> .....	11
1.5 Clonación de la secuencia promotora de un gen .....	12
1.6 El uso de herramientas bioinformáticas en el análisis de una secuencia promotora .....	12
1.7 Elementos reguladores en <i>cis</i> .....	13
<b>2. MARCO TEÓRICO</b> .....	15
<b>3. HIPÓTESIS</b> .....	15
<b>4. OBJETIVOS</b> .....	16
<b>5. MATERIAL Y MÉTODOS</b> .....	17

<b>6. RESULTADOS</b> .....	30
<b>6.1 Clonación, secuenciación y análisis de los fragmentos de ADN genómico adyacentes al extremo 5' de la secuencia del gen <i>NaStEP</i></b> .....	31
<b>6.2 La secuencia 1.4 Kb río arriba de la región codificadora del gen <i>NaStEP</i>, posee varios elementos de una región promotora</b> .....	41
<b>6.3 Elementos en <i>cis</i> involucrados en la regulación de la expresión específica del estigma y el estilo en el promotor de <i>NaStEP</i></b> .....	48
<b>6.4 <i>NaStEP</i> fusionado a su promotor potencial: construcción y clonación</b> .....	50
<b>6.5 Generación de un vector binario de expresión para evaluar la función como una región promotora, de la secuencia 1.4 Kb río arriba de la región codificadora del gen <i>NaStEP</i></b> .....	53
<b>7. DISCUSIÓN</b> .....	57
<b>8. PERSPECTIVAS</b> .....	64
<b>9. CONCLUSIONES</b> .....	67
<b>10. BIBLIOGRAFÍA</b> .....	68
<b>11. Apéndice</b> .....	74

## RESUMEN

La autoincompatibilidad (AI) es el mecanismo más eficaz usado por las plantas con flores para evitar la endogamia y promover así la generación y mantenimiento de la diversidad genética. La especificidad en el rechazo del polen en las especies de *Nicotiana* está controlada principalmente por un solo locus multialélico conocido como el locus S; sin embargo, también se requiere de la presencia de los productos de otros genes no codificados en este locus, a los cuales se les ha dado el nombre de genes modificadores (GM). En nuestro laboratorio, se descubrió un GM esencial en la vía genética del rechazo del polen, al cual se le nombró como *NaStEP* (*N. alata Stigma Expressed Protein*). *NaStEP* codifica para una proteína específica del estigma la cual es detectable únicamente en especies AI y no en las de tipo autocompatible. Además, la expresión de *NaStEP* alcanza un primer pico hacia la madurez del estigma con un segundo aumento después del evento de la polinización. Debido a que *NaStEP* despliega un patrón específico de expresión en el estigma de *Nicotiana*, se sugiere fuertemente la presencia de un mecanismo fino de regulación, el cual pudiera estar actuando a través de los elementos presentes a nivel de su secuencia promotora. De esta manera, mediante una técnica de *DNA walking*, se clonó un segmento de ADN río arriba de la región codificadora del gen *NaStEP*, el cual posee una longitud de 1.4 Kb; y al cual se clonó en un vector binario de expresión con la finalidad de evaluar su función como una región promotora. Asimismo, el ensayo bioinformático de esta región, mostró que posee varios elementos potenciales reguladores en *cis*, los cuales incluyen elementos del promotor primordial y otros sitios de unión a factores de transcripción; de los cuales destacan por su relevancia en la expresión de este gen, elementos potenciales en *cis* implicados en la expresión específica de estigma y estilo.

# 1. INTRODUCCIÓN

## 1.1 La regulación de la expresión genética en las plantas

Debido a su naturaleza sésil y a su continuo desarrollo, las plantas deben constantemente regular y modular su expresión genética; así las células son capaces de percibir y responder correctamente a su ambiente, de participar en la homeostasis y de colaborar en el desarrollo adecuado del organismo. Además, se ha visto que las modificaciones en las redes de regulación genética provocan divergencia en la expresión genética lo que correlaciona con la capacidad de los organismos para responder a un determinado ambiente; sugiriendo fuertemente que la evolución puede estar principalmente conducida por cambios en el complejo mecanismo de regulación genética y no únicamente por los polimorfismos presentes en las secuencias codificadoras de las proteínas como se creía anteriormente (Wray, 2003; Carroll, 2005; Wagner and Lynch, 2010). Todo esto nos lleva a pensar en la importancia de los mecanismos de regulación genética en las plantas.

## 1.2 El sistema de incompatibilidad como un mecanismo en el cual, la regulación de la expresión genética juega un papel importante

Los órganos reproductivos femenino y masculino en una flor bisexual se encuentran muy próximos, por lo que su cercanía favorece que el polen llegue al estigma de la misma flor que lo produjo, aumentando la probabilidad de la autofecundación. Si esto ocurre, la variabilidad genética de las siguientes generaciones se verá reducida. Es por ello que las plantas con flores hermafroditas desarrollaron varios mecanismos genéticos para reducir la frecuencia de la autofecundación (Jiménez-Durán y Cruz-García, 2011)

El mecanismo más eficaz usado por las plantas con flores para evitar la endogamia y promover la generación y mantenimiento de la diversidad genética entre las especies y con ello el potencial para adaptarse a las nuevas condiciones ambientales es la autoincompatibilidad (AI) (de Nettancourt, 1977).

En el sistema de incompatibilidad homomórfico existe un solo morfo floral en la población, y la compatibilidad en muchas especies está generalmente controlada por un solo locus multialélico conocido como el locus S, el cual determina la especificidad del reconocimiento o rechazo del polen. Esto implica que los productos del locus S expresados en el polen y en el pistilo interactúan para determinar si un polen dado es autocompatible (AC) ó autoincompatible (AI) (de Nettancourt, 1997; Franklin-Tong *et al.*, 2003).

Debido a que el sistema de incompatibilidad juega un papel durante el evento de la reproducción de las angiospermas, la expresión de los genes involucrados en esta vía debe contar con una regulación altamente específica capaz de delimitar la localización de los productos de estos genes en los diferentes tejidos del polen y del pistilo.

### **1.2.1 Función y expresión de las determinantes, femenina y masculina, del locus S**

En las familias Solanaceae (*Nicotiana*, *Petunia*, *Solanum*), Rosaceae (*Prunus*) y Scrophulariaceae (*Antirrhinum*) los productos génicos del locus S en el estilo son proteínas con actividad de ribonucleasa, conocidas como S-RNasas (McClure *et al.*, 1989; McCubbin and Kao, 1996; Kao and Tsukamoto, 2004; de Nettancourt, 2001; McClure, 2004; McClure and Franklin-Tong, 2006).

Las S-RNasas son proteínas básicas, con un punto isoeléctrico en el rango de 7.6 a 9.5, y con un peso molecular de 28 a 35 kDa dependiendo de su grado de glicosilación (Woodward *et al.*, 1989). Estas proteínas con actividad de ribonucleasas, se secretan hacia la matriz extracelular del tejido de transmisión del estilo, sitio por el cual crece y se dirige el tubo polínico al ovario (Anderson *et al.*, 1986; Cornish *et al.*, 1987; McClure *et al.*, 1993; Broocthaerts *et al.*, 1995; Sassa *et al.*, 1996; Xue *et al.*, 1996). En este sitio, las S-RNasas acceden al citoplasma del tubo polínico donde causan degradación del RNA de los tubos polínicos incompatibles (McClure *et al.*, 1990; Gray *et al.*, 1991). Esta actividad citotóxica afecta el crecimiento de los tubos polínicos incompatibles interfiriendo con el programa de expresión

genética del polen (McClure and Franklin-Tong, 2006). Cuando las S-RNAs pierden su actividad, no se observa degradación del RNA y las plantas no presentan el fenotipo de rechazo del polen (Huang *et al.*, 1994).

El producto génico del locus *S* en el polen se conoce como *SLF* (*S-locus F-box gene*) en Solanáceas ó *SFB* (*S-haplotype-specific F-box protein*) en Rosaceae, y es miembro de una familia de proteínas que contiene una caja F hacia su extremo amino terminal. La expresión de *SLF* es haplotipos *S*-específica, y se restringe al polen (Ikeda *et al.*, 2004; Yamame y Tao, 2009). Los miembros de la familia de proteínas con caja F se asocian con la vía de degradación de proteínas por el proteosoma 26S (Sullivan *et al.*, 2003). La proteína *SLF* podría funcionar como una enzima ligasa tipo E3 (Sijacic *et al.*, 2004). La degradación de proteínas vía el proteosoma 26S involucra el marcaje por poliubiquitinación de las proteínas que se van a degradar. En este proceso las enzimas ligasas E3 representan un paso crucial en el control de la poliubiquitinación de la proteína blanco, ya que reconocen la señal de ubiquitinación en la proteína seleccionada y coordinan la transferencia de la ubiquitina (proteína de 76 kD) desde la ligasa E2 a un residuo de lisina de la proteína blanco (Sullivan *et al.*, 2003).

### **1.2.2 Genes modificadores (GM)**

Aunque se sabe que la especificidad del rechazo del polen depende de la interacción específica del haplotipo, entre la S-RNasa y *SLF*, también se requiere de los productos de otros genes no codificados en el locus *S*, a los cuales se les ha dado el nombre de genes modificadores (GM). Estos genes, al igual que los genes codificados por el locus *S*, poseen un patrón de expresión el cual es específico de polen y de pistilo.

Los GM se clasifican en tres grupos (McClure *et al.*, 2000). El Grupo I incluye genes cuyos productos regulan la expresión de las determinantes de la especificidad (*SLF* y S-RNasa). El Grupo II comprende GM que regulan la actividad biológica de las determinantes de especificidad y, por tanto, su función se restringe a la incompatibilidad. De este grupo sólo tres GM han sido identificados con función confirmada en el rechazo del polen, *HT-B*,

*NaStEP* y *120K* (McClure *et al.*, 1999; Hancock *et al.*, 2005; Jiménez-Durán *et al.*, no publicado).

HT-B posee una localización específica de estilos maduros de plantas AI. La proteína madura tiene una masa de 8.2 kDa, con un dominio constituido por 20 residuos de asparagina y ácido aspártico (dominio ND) hacia el extremo carboxilo. Su papel en la AI fue confirmado silenciando su gen por antisentido y RNA de interferencia (RNAi) en plantas transgénicas de *Nicotiana* y *Solanum*, ya que las plantas transgénicas pasaron de ser AI a AC (McClure *et al.*, 1999).

La proteína 120K es una arabinogalactoproteína específica de estilos de *N. alata*. Además, posee un peso molecular de 120 kDa, por ello su nombre. 120K forma complejos *in vitro* con las S-Rnasas (Cruz-García *et al.*, 2005) e ingresa al tubo polínico (Lind *et al.*, 1994) sin importar si es compatible o incompatible (Goldraij *et al.*, 2006). En cruzas compatibles y en etapas tempranas de la polinización de cruzas incompatibles, 120K se localiza delimitando a la vacuola donde se encuentran secuestradas las S-RNasas en el tubo polínico. Sin embargo, en etapas tardías de una polinización incompatible (36 h), 120K ya no se detecta en los tubos polínicos (Goldraij *et al.*, 2006). Estas evidencias hacen suponer que 120K se encuentra asociada a sistemas membranosos en el tubo polínico y que su participación en el rechazo del polen podría ser importante en la liberación de las S-RNasas de su compartimiento.

En el Grupo III se encuentran los genes que funcionan en el rechazo del polen y en otros procesos de la interacción polen-pistilo. En este grupo podrían estar *TTS*, *PELP III*, *Nap11*, *NaTrxh*, y *SBP* (O'Brien *et al.*, 2004; Juárez-Díaz *et al.*, 2006).

### 1.2.3 *NaStEP* como un gen modificador en *N. alata* y cuyo patrón de expresión es único

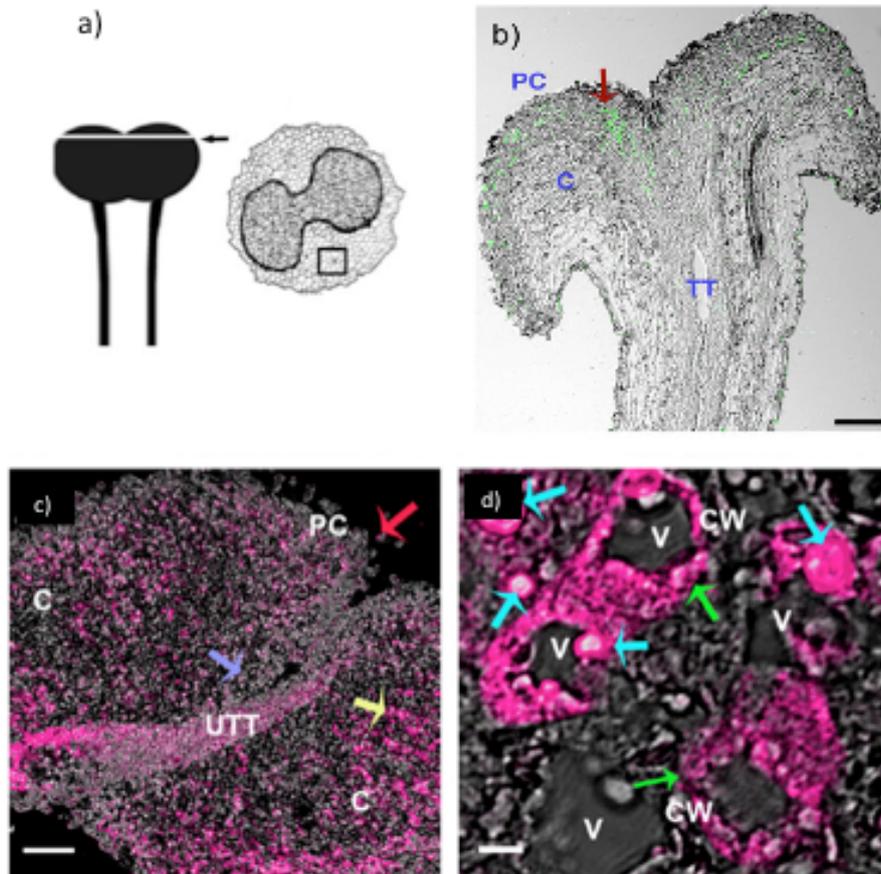
*NaStEP* es una proteína con una masa teórica de 24.37 KDa cuyo punto isoeléctrico estimado es de 4.68. Esta proteína posee homología con inhibidores de proteasas tipo Kunitz, no obstante, al momento no se ha podido comprobar que posee esta actividad.

*NaStEP* es un gen con expresión exclusiva de especies autoincompatibles como *N. bonariensis*, *N. forgetiana* y cultivares AI de *N. alata*. Además, este gen se expresa exclusivamente en los estigmas maduros de éstas especies, y su transcrito se detecta desde etapas tempranas del desarrollo del estigma, alcanzando sus máximos niveles hacia la madurez del pistilo (Busot *et al.*, 2008).

Los estudios inmunocitoquímicos indican que *NaStEP* se almacena en las vacuolas de las células estigmáticas de estilos sin polinizar. No obstante, después de la polinización con el polen propio y con aquel proveniente de *N. tabacum*, *N. alata* y *N. plumbaginifolia*, *NaStEP* se detecta en el exudado estigmático, lo que indica que la polinización cambia su localización celular, liberándose al exudado estigmático después de que ésta ocurre. Además, su síntesis y transferencia al exudado se ven estimulados fuertemente por polinizaciones incompatibles, como la autopolinización ó polinización con polen de especies autocompatibles como *N. tabacum* ó *N. plumbaginifolia*. La síntesis total de *NaStEP* sólo se incrementa 1.7 veces después de una polinización compatible, mientras que aumenta 3.7 veces después de una polinización incompatible. Lo mismo sucede con la deposición de *NaStEP* hacia el exudado estigmático, donde se observa que se estimula 20 veces en una polinización incompatible y 12 veces después en una polinización compatible (Busot *et al.*, 2008). Una vez en el exudado estigmático, *NaStEP* se incorpora al los tubos polínicos donde se almacena en sus vacuolas (Jiménez-Durán *et al.*, en preparación).

Experimentos de pérdida de función por RNAi de *NaStEP* en plantas transgénicas de *Nicotiana* permitieron determinar la función de este gen en la respuesta de rechazo del polen. Cuando este gen fue silenciado en plantas transgénicas de *Nicotiana*, éstas perdieron la capacidad de reconocer y

rechazar el polen con el mismo haplotipo S que el de la planta receptora, perdiendo así la autoincompatibilidad (Jiménez-Durán *et al.*, en preparación).



**Fig. 1 Análisis inmunohistoquímicos de NaStEP en tejidos estigmáticos y estilares de *N. alata* BT.** a) Diagramas de el estigma más el estilo de *N. Alata* y sección transversal de un estigma. La flecha y el cuadrado muestran las zonas donde las imágenes fueron tomadas. b) Estigma y estilo tratados con anticuerpo anti-NaStEP (verde). c) Secciones transversales de estigma mostrando a NaStEP (señal magenta) en las papilas estigmáticas (flecha roja), las células secretoras (flecha amarilla), células basales (flecha azul) y en una porción de la parte superior del tejido de transmisión. d) Acercamiento de las células estigmáticas secretoras donde se muestra que NaStEP está en pequeños cuerpos (flechas azules) y en la proximidad de la pared celular (flechas verdes). CW, pared celular; V, vacuolas; TT, tejido de transmisión; UTT, parte superior del tejido de transmisión; C, cortex; PC, células papilares. . Las barras corresponden a 50 mm en (b), 20 mm en (c), y 5 mm en (d). La figura fue adaptada a partir de la publicada por Busot *et al.* (2008).

### 1.3 La expresión genética tiene diferentes niveles de regulación

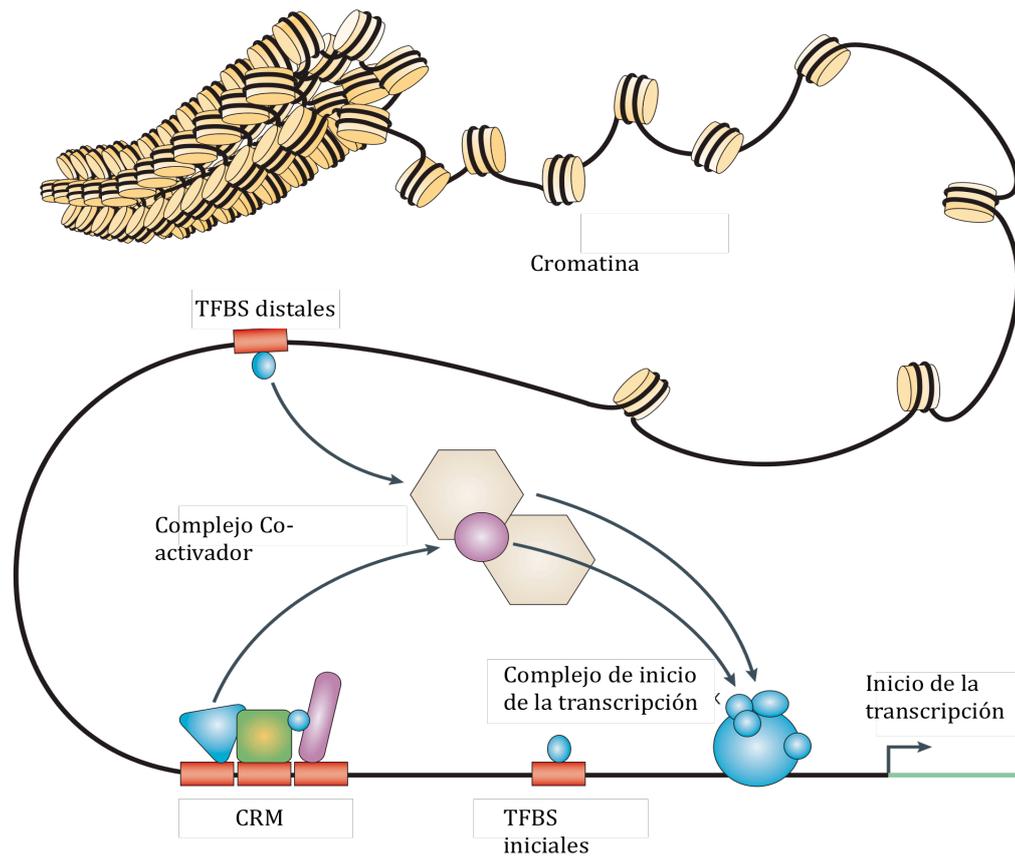
En el control de la expresión genética interviene más de un evento, los cuales van desde el inicio de la transcripción hasta las modificaciones post-traduccionales de la proteína. Al tipo de regulación de la expresión genética que actúa sobre el inicio de la transcripción, se le conoce como regulación en *cis*. A todos los otros pasos reguladores de la expresión se les define como ARLs (por sus siglas en inglés, *alternative regulatory levels*) (Alonso and Wilkins, 2005). Los ARLs incluyen el papel de los RNAs pequeños, por ejemplo microRNAs, los cuales son capaces de modificar los patrones de expresión genéticos mediante su actuación como factores en *trans*, es decir algunos RNAs tienen un papel importante y decisivo como reguladores de la expresión genética, jugando un papel que se creía exclusivo de las proteínas. (Zhao y Srivastava, 2007; Filipowicz *et al.*, 2008; Hobert *et al.*, 2008; Rhoades *et al.*, 2002; Allen *et al.*, 2005). Sin embargo, el punto clave en la regulación de la expresión genética se encuentra en el inicio de la transcripción, evento sobre el cual actúa la regulación en *cis*.

La regulación en *cis* es el mecanismo a través del cual se regula el inicio de la transcripción. Ésta se define como el control de la transcripción basada en la interacción entre secuencias cortas de ADN situadas en el mismo cromosoma que el gen regulado, y las proteínas de actuación en *trans*, llamadas factores de transcripción (FTs) .

El inicio de la transcripción de los genes que codifican proteínas en los eucariontes es precedido por múltiples eventos. Estos son la descondensación del locus, la reordenación del nucleosoma y la modificación de histonas, la unión de los factores de transcripción y coactivadores a intensificadores ó *enhancers* y promotores, y finalmente el reclutamiento de la maquinaria basal de la transcripción a la secuencia del promotor mínimo (Fig. 2) (Smale y Kadonaga, 2003).

### 1.4 Estructura y características de un promotor

Los elementos reguladores en *cis* son secuencias funcionales de ADN que, por su naturaleza modular y combinatoria, controlan de manera precisa el

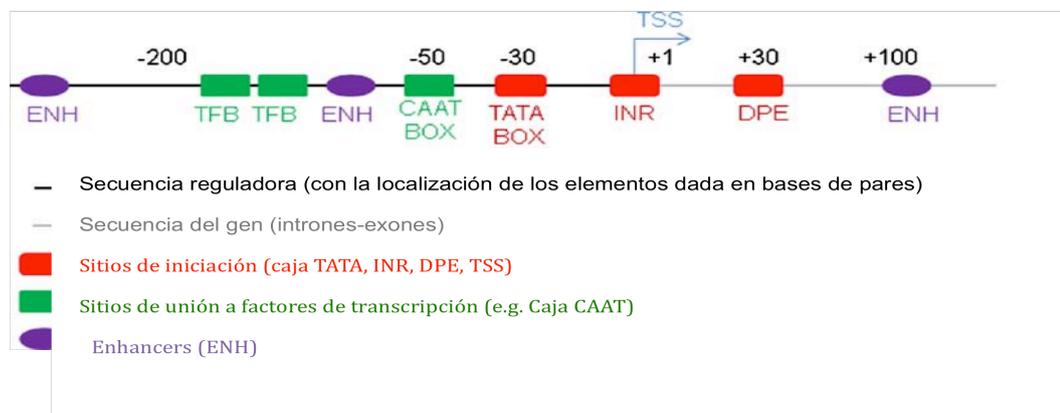


**Figura 2. Componentes de la regulación transcripcional.** Los factores de transcripción (TFs) se unen a sitios específicos (TFBS) los cuales son proximales ó distales de acuerdo al sitio de inicio de la transcripción. Conjuntos de TFs pueden operar en un módulo funcional cis-regulador (CRMs) para alcanzar propiedades reguladoras específicas. La interacción entre la TFs unidos y los cofactores estabiliza la maquinaria de la iniciación de la transcripción lo que permite la expresión genética. La regulación que es conferida por la unión de TFs a secuencias específicas es altamente dependiente de la estructura tri-dimensional de la cromatina. La figura fue adaptada a partir de la publicada por Wasserman y Sandelin (2004).

patrón de expresión espacial y temporal de decenas de miles de genes expresados en las células de eucariontes. Los elementos en *cis* suelen agruparse en diferentes módulos a los cuales se les conoce como CRMs (por sus siglas en inglés, *cis regulatory modules*). Estos CRMs integran señales combinadas para múltiples factores de transcripción. Es decir, este tipo de agrupación de elementos reguladores en *cis* es la que da resultado a patrones de expresión altamente específicos y ligeramente diferenciados entre sí (Priest *et al.*, 2009).

La región que engloba el total de los elementos reguladores en *cis* está compuesta del promotor primordial ó *core promoter*, y de los reguladores próximos y remotos. Cercano al sitio de inicio de la transcripción, se ubica el promotor primordial ó *core promoter*. Ríó arriba ó ríó abajo se ubican los reguladores próximos. Los reguladores remotos también llamados *enhancers*, pueden finalmente encontrarse dentro de los intrones del gen, ó bien, a una distancia de varios kilobases tanto ríó arriba como ríó abajo (Figura 3) (Wasylyk, 1988; Fassler y Gussin, 1996).

Los módulos de *enhancers* pueden variar su tamaño e ir de 50 pb hasta 1.5 K pb (Blackwood and Kadonaga, 1998). Además, pueden aumentar los niveles de transcripción de un gen mediante la unión de proteínas activadoras y no por la actuación de manera directa sobre los motivos a los cuales se unen los factores de transcripción.



**Fig. 3 Estructura general de una secuencia *cis*-reguladora.** La secuencia *cis*-reguladora esta dividida en el *core promoter*, el cual generalmente se localiza entre 0 y -200pb, en esta región se encuentran los sitios de unión de la ADN polimerasa (caja TATA, elemento iniciador, INR, *Downstream promoter element*, DPE), sitios de unión a factores de transcripción (TFB) y enhancers (ENH), situados en el mismo cromosoma pero posiblemente lejos ríó arriba ó ríó abajo de la secuencia promotora. Sitio de inicio de la transcripción +1 (TSS). La figura fue adaptada a partir de la publicada por Vedel y Scotti (2011).

Los promotores se caracterizan por un mayor grado de modularidad que las proteínas (los motivos son más cortos que los dominios de éstas, y más numerosos por promotor que los dominios por proteína). Además, el orden en el cual los motivos de secuencias reguladoras se encuentran parece irrelevante sobre la función de un promotor. Como consecuencia, es más fácil

la obtención de nuevos arreglos funcionales en las secuencias reguladoras que en las regiones que se transcriben, y por lo tanto las secuencias reguladoras suelen estar menos conservadas que las regiones codificadoras, manteniendo aún su funcionalidad (Vedel y Scotti, 2011).

La segunda característica que define a un promotor, es la redundancia que existe a nivel de los motivos de secuencias reguladoras. Como consecuencia de este hecho, puede esperarse que los eventos de recombinación y mutación cuando ocurren a nivel de la secuencia de un promotor, provoquen un menor número de efectos deletéreos sobre la función de un gen, y al mismo tiempo, igual que la modularidad, una mayor variabilidad, que cuando éstos caen sobre la región codificadora de la proteína (Vedel y Scotti, 2011).

#### **1.4.1 El promotor mínimo ó *core promoter***

La secuencia del promotor mínimo ó *core promoter*, de aproximadamente 200 pb se ubica cercana al sitio de inicio de la transcripción, en el extremo 5' de la secuencia codificadora del gen. La mayoría de los elementos que pertenecen al *core promoter* parecen interaccionar directamente con los componentes de la maquinaria basal de la transcripción.

A pesar de que inicialmente se pensó como invariable a la composición de los sitios de unión en el promotor primordial para la RNA polimerasa II se ha encontrado que por el contrario, existe en él una alta diversidad tanto a nivel estructural como funcional (Smale, 1994). Además, se ha demostrado que la diversidad en el *core promoter* contribuye de manera importante a la regulación combinatoria de la expresión genética (Smale, 2001; Butler y Kadonaga, 2002).

A través del estudio computacional realizado por Ohler *et al* (2002) se encontró como elementos predominantes en el *core promoter*, a la caja TATA, Inr y DPE, cuyas localizaciones relativas al sitio de inicio de la transcripción se encuentran restringidas. Sin embargo, es altamente probable que existan elementos adicionales al *core promoter*, los cuales se encuentren a distancias no conservadas del sitio de inicio de la transcripción.

## **1.5 Clonación de la secuencia promotora de un gen**

A pesar de la complejidad y diversidad de los motivos en los promotores, existe más de un camino en la identificación de la secuencia de un promotor. Por ejemplo, tal como sucede con organismos modelos como *A. thaliana*, *Zea mays* y *Populus trichocarpa*, el genoma ha sido enteramente secuenciado y por lo tanto, la secuencia reguladora de un gen puede ser clonada posicionalmente mediante el análisis de la secuencia río arriba de su región codificadora. Sin embargo cuando el genoma de la especie a la cual pertenece el gen no ha sido codificado en su totalidad, como sucede con el genoma de *N. alata*, especie del cual fue clonado el gen *NaStEP*, un segundo tipo de aproximación es necesario en la búsqueda de la secuencia reguladora (Vedel y Scotti, 2011).

Basándose en el conocimiento de la secuencia codificadora del gen, el aislamiento de la secuencia de un región promotora puede lograrse a través de técnicas como el *Tail PCR* ó *Gene Walking*. Esto tipo de métodos hacen uso de la extensión de los *primers*, diseñados sobre la secuencia conocida, para obtener a la secuencia genómica adyacente, la cual era desconocida mediante la reacción en cadena de la polimerasa. Este segundo tipo de acercamiento es más laborioso, pero posee la ventaja de ser aplicado de forma casi universal (Vedel y Scotti, 2011).

## **1.6 El uso de herramientas bioinformáticas en el análisis de una secuencia promotora**

Una vez identificada la secuencia de un promotor, ésta se analiza mediante las herramientas bioinformáticas propias. El uso de herramientas bioinformáticas permite identificar a los elementos reguladores en *cis* pertenecientes al promotor primordial ó *core promoter*, y a los demás sitios de unión a factores de transcripción (FT), ya que el reconocimiento de estos elementos en una secuencia mediante únicamente inspección visual es casi imposible. Una de las razones que conlleva a la dificultad en la identificación de estos elementos, se debe a la alta variabilidad que los caracteriza.

Además, debido a las características de un promotor, éste puede estar compuesto de módulos que pertenecen a secuencias reguladoras conocidas, pero en una combinación completamente distinta, lo cual frecuentemente sería ignorado por inspección visual ó aún mediante alineamiento de secuencias. Por ello, el uso de distintos algoritmos de reconocimiento se hace aún más indispensable.

Varios programas están actualmente disponibles y son ampliamente usados en la búsqueda y predicción de los motivos de los factores de transcripción, e inclusive en la predicción de promotores completos (Rombauts *et al.*, 2003; Pan *et al.*, 2009; Priest *et al.*, 2009). Algunos son de uso general en todos los organismos, mientras que otros se enfocan exclusivamente a los promotores de plantas, como es el caso de *PLANT CARE* (Plant cis-acting regulatory DNA elements) (Higo *et al.*, 1999) y *Softberry TSSP* (Shahmuradov *et al.*, 2003).

### **1.7 Elementos reguladores en *cis***

En años recientes, la rápida acumulación de grupos de datos a escala genómica, incluyendo la secuenciación de los genomas, *genome annotations*, predicciones de la función genética, y experimentos de patrones de expresión, ha permitido el descubrimiento de los elementos reguladores en *cis*, llevándonos a preguntar sobre sus funciones en la señalización celular de las plantas.

Mediante ensayos de microarreglos, es posible definir a grupos de genes como genes co-regulados, los cuales poseen la característica de compartir patrones de expresión determinados. La búsqueda de los elementos reguladores en *cis* que pudieran ser responsables de su patrón de expresión, se efectúa mediante la búsqueda de elementos de secuencias conservadas, dentro de las regiones de longitud arbitraria río arriba de cada uno de los genes.

Varios estudios se han enfocado en la señalización por fitohormonas, ó en la señalización de la respuesta a estrés. Así por ejemplo, se han analizado a las regiones promotores de los genes de *Arabidopsis* que muestran un patrón de

expresión similar bajo un tratamiento con ácido abscísico (ABA) ó bajo estrés abiótico, es decir condiciones de sequía, frío y distintas concentraciones de sal. De esta manera, se identificó la presencia de un número de elementos de respuesta a ABA (ABREs) que contienen la secuencia ACGT, y también la presencia de elementos acoplados (Zhang *et al.*, 2005).

La mayoría de los eucariontes usan ciclos diarios de luz y oscuridad como señales de tiempo, los cuales permiten que una amplia variedad de procesos biológicos ocurran de manera organizada a diferentes tiempos del día. Estudios recientes han identificado nuevos componentes de la regulación transcripcional diurna y circadiana (Michael<sup>1</sup> *et al.*, 2008; Covington *et al.*, 2008; Michael<sup>2</sup> *et al.*, 2008). A través del análisis de las secuencias promotores de los genes de *Arabidopsis*, que responden a diferencias en la luz, la temperatura ó las condiciones circadianas, se identificaron a tres módulos distintos de elementos reguladores en *cis*, los cuales son capaces de controlar la expresión en determinados momentos del día. Los elementos correspondientes a la mañana, son CRM de la mañana (*cis-regulatory modules*) (ME, CCACAC)/ G-box (CACGTG); los elementos correspondientes a la tarde, CMR de la tarde (EE, AAATATCT)/GATA (GATA); y los elementos de medianoche, los cuales a su vez son CRM de la noche (TBX, AAACCCT)/*starch synthesis box* y (SBX, AAGCCC)/*protein box* (PBX, ATGCCC. Estos tres módulos se encuentran conservados a lo largo de especies distantemente relacionadas, como lo son *Arabidopsis*, el arroz y la papaya sugiriendo que la señalización diurna y circadiana ha moldeado la evolución de las redes transcripcionales en las plantas, permitiéndoles a adaptarse a diversos ambientes, los cuales también cambian a lo largo del día (Michael<sup>1</sup> *et al.*, 2008; Michael<sup>2</sup> *et al.*, 2008; Zdepski *et al.*, 2008).

## 2. MARCO CONCEPTUAL

La gran especificidad de expresión de los genes que controlan y componen la vía genética del rechazo del polen en *Nicotiana*, hacen suponer que estos tienen una regulación muy fina la cual podría abarcar desde eventos pre-transcripcionales hasta eventos post-traduccionales. Sin embargo, el inicio de la transcripción es el punto clave en la regulación de la expresión, por ello aislar a la región del promotor, de un gen modificar esencial en esta vía como lo es *NaStEP*, es de suma trascendencia para iniciar estudios de la regulación de la expresión genética en plantas.

## 3. HIPÓTESIS

La especificidad de expresión de *NaStEP* en estigmas de *N. Alata* hacen suponer que la región promotora de este gen cuenta con elementos reguladores únicos, capaces de inducir su expresión en tejido y tiempo específicos.

#### 4. OBJETIVOS

- I. Identificar y clonar la región promotora del gen *NaStEP*.
- II. Analizar a través de un ensayo bioinformático la secuencia promotora del gen *NaStEP*.
- III. Identificar posibles elementos reguladores en *cis* responsables del patrón de expresión específico de este gen en el estigma.
- IV. Fusionar el cDNA de *NaStEP* a su región promotora.
- V. Generar construcciones *NaStEPp-GFP* para evaluar la función de la región promotora de *NaStEP* en plantas transgénicas de *Nicotiana*.

## 5. MATERIALES Y MÉTODOS

### 5.1 Extracción de ADN genómico a partir de *Nicotiana glauca*.

Purificación de ADN genómico empleando el ChargeSwitch® gDNA Plant Kit

#### **Preparación del lisado proveniente de tejido vegetal**

(Procedimiento para preparar el lisado proveniente de 100 mg de tejido vegetal)

- a) Enfriar el Buffer de Precipitación (N5) en hielo.
- b) En el caso de contar con tejido vegetal duro y fibroso, éste debe congelarse previamente en nitrógeno líquido, para posteriormente ser molido con pistilo y mortero hasta lograr pulverizarlo totalmente. Si el tejido vegetal fuese suave, bastará con cortarlo en fragmentos pequeños.
- c) Adicionar a temperatura ambiente, 1 mL del ChargeSwitch® Lysis Buffer (L18) al tejido del paso anterior.
- d) Adicionar a la muestra 2 µl de RNase A.
- e) Homogeneizar mediante el uso de un molinillo ó de un vortex, hasta obtener a la muestra completamente resuspendida.
- f) Adicionar 100 µl de SDS 10% a 1 ml del lisado vegetal.
- g) Incubar a temperatura ambiente durante 5 minutos.
- h) Adicionar 400 µl ChargeSwitch® Precipitation Buffer (N5) de al lisado. Mezclar por inversión hasta que se observe la formación del precipitado.
- i) Centrifugar a temperatura ambiente y a máxima velocidad durante 5 min. hasta obtener al lisado transparente.
- j) Transferir al lisado claro a un tubo de 1.5 ml limpio y estéril.

#### **Unión del ADN genómico**

- k) Vortexear exhaustivamente a uno de los tubos con ChargeSwitch® Magnetic Beads con la finalidad de lograr su completa resuspensión en el buffer de almacenaje.
- l) Adicionar 100 µl de ChargeSwitch® 10% Detergent (D1) al lisado de ~1.2 ml del inciso j.
- m) Además, adicionar 40 µl de la resuspensión de ChargeSwitch® Magnetic Beads .
- n) Mezclar generosamente mediante el pipeteo arriba y abajo utilizando una micropipeta de 1 ml, ajustada a la cantidad de 900 µl durante cinco veces y evitando la formación de burbujas.
- o) Incubar a temperatura ambiente durante 1 min.
- p) Colocar los tubos en la MagnaRack™, en la posición adaptada al imán hasta observar la formación de un pellet compacto.
- q) Sin remover a los tubos de su posición, aspirar cuidadosamente y desechar al sobrenadante sin destruir al pellet formado en el paso anterior; para ello, se debe apuntar la punta de la pipeta hacia el lado opuesto de la formación de éste.

### **Lavado del ADN genómico**

- r) Remover a los tubos de su posición junto al imán.
- s) Adicionar a cada uno de los tubos 1ml de ChargeSwitch® Wash Buffer (W12) y pipetear gentilmente cinco veces arriba y abajo, evitando la formación de burbujas con una pipeta de 1 ml ajustada a 900 µl, logrando resuspender nuevamente a las microperlas magnéticas.
- t) Colocar los tubos en la MagnaRack™, en la posición adaptada al imán hasta observar nuevamente la formación de un pellet compacto y al sobrenadante transparente.
- u) Sin remover a los tubos de su posición adaptada al imán, aspirar cuidadosamente y descartar nuevamente al sobrenadante.
- v) Remover a los tubos de su actual posición junto al imán.
- w) Realizar por segunda vez el lavado del ADN genómico, mediante la repetición de los incisos s, t y u.

### **Elución del ADN genómico**

- x) Remover a los tubos del inciso anterior, de su posición junto al imán. En este paso no debe observarse sobrenadante en ninguno de los tubos.
- y) Adicionar 150 µl del ChargeSwitch® Elution Buffer (E6).
- z) Pipetear arriba y abajo gentilmente durante 15-30 veces, usando una pipeta de 100 µl, hasta que los ChargeSwitch® Magnetic Beads estén totalmente dispersos y no se observe la formación de ningún cúmulo.
- aa) Incubar a temperatura ambiente durante 1 min.
- bb) Colocar los tubos en la MagnaRack™, en la posición adaptada al imán hasta observar la formación de un pellet compacto.
- cc) Sin remover a los tubos de su posición adaptada al imán, transferir cuidadosamente al sobrenadante, el cual contiene al ADN genómico, a un tubo de microcentrifuga estéril, procurando mantener intacta la estructura del pellet previamente formado.
- dd) Desechar a los ChargeSwitch® Magnetic Beads que se hubieron empleado en la realización del protocolo.

### **5.2. Generación de fragmentos de ADN genómico adyacentes al extremo 5' de la secuencia del gen *NaStEP* mediante la metodología del Kit Universal *Genome Walker*.**

El Kit Universal *Genome Walker*, provee de un método simple para la clonación de una secuencia de ADN genómico desconocida, adyacente a una secuencia conocida de ADN genómico ó cDNA.

El primer paso del Kit Universal *Genome Walker* es la generación de cuatro bibliotecas a partir del ADN genómico de interés. La construcción de las cuatro bibliotecas, consiste primeramente en la digestión por separado de cuatro alícuotas de ADN genómico. Cada alícuota es digerida con una enzima de restricción distinta, *EcoRV*, *Dra I*, *Stu I* y *Pru I*, las cuales son provistas por el Kit. Así, el ADN genómico digerido es ligado en un segundo

paso al Adpatador. Después de la generación de las cuatro bibliotecas, se continúa a realizar dos amplificaciones del producto deseado en cada una de ellas mediante PCR. La primera reacción o PCR primario utiliza un *primer* externo (AP1) que empalma con la secuencia del adaptador y que es provisto por el Kit, y un segundo *primer* externo que empalma con una secuencia específica (GSP1) de ADN adyacente a la región de búsqueda. La mezcla obtenida del PCR primario es diluida y usada como templado de una segunda reacción de PCR, a la cual se le conoce como PCR anidado o secundario. En este segunda reacción se utilizan dos *primers*, el AP2 y el GSP2, los cuales empalman de manera anidada con la secuencia que corresponde al adaptador y a la secuencia específica, respectivamente. El protocolo inmediato a la generación de las cuatro bibliotecas, es descrito a continuación.

### 5.2.1 PCR primario

a. Etiquetar los tubos de PCR de 0.5 ml a manera igual a la que se observa en la Tabla 1. La colección de tubos xA corresponde a la reacción de PCR primario. Los tubos xB corresponden a la reacción de PCR anidado ó secundario. DL1, DL2, DL3 y DL4 se refiere a cada una de las cuatro bibliotecas. La biblioteca control es construida a partir de ADN genómico humano, provisto por el kit. La biblioteca control preconstruida es provista por el kit y consiste de ADN genómico humano, digerido y ligado al adaptador. Estas dos últimas bibliotecas servirán ambas como controles positivos. Tanto el control negativo No. 1 como el control negativo No. 2 consisten en reacciones de PCR sin emplear ADN genómico.

**Tabla1. Etiquetas de los tubos de el PCR primario y secundario**

Biblioteca de DNA	Tubo No.	Primers		
		1º PCR	2º PCR	
DL-1 ( <i>EcoRV</i> )	1A	GSP1 & AP1	1B	GSP2 & AP2
DL-2 ( <i>Dra I</i> )	2A	"	2B	"
DL-3 ( <i>Stu I</i> )	3A	"	3B	"
DL-4 ( <i>Pru I</i> )	4A	"	4B	"
<b>Control negativo No.1</b>	5A	"	5B	"
<i>Nada</i>				
<b>Control positivo No. 1</b>	6A	PCP1 & AP1	6B	PCP2 & AP2
<i>Biblioteca control</i>				
<b>Control negativo No. 2</b>	7A	"	7B	"
<i>Nada</i>				
<b>Control positivo No.2</b>	8A	"	8B	"
<i>Biblioteca control preconstruida</i>				

b. Preparar una mezcla maestra de ocho reacciones más una adicional, para la reacción de PCR primario tal como indica la Tabla 2.

**Tabla 2. Mezcla maestra del PCR primario.**

	<u>9 rxns</u>	<u>Por rxn</u>
H <sub>2</sub> O desionizada	360 µl	40 µl
10X Advantage 2 PCR Buffer	45 µl	5 µl
dNTP (10 mM each)	9 µl	1 µl
AP1 (10 mM)	9 µl	1 µl
Advantage 2 Polymerase Mix (50X)	9 µl	1 µl
Volumen Total	432 µl	48 µl

c. Mezclar mediante vortexeo, buscando no introducir burbujas y centrifugar ligeramente en una microcentrifuga.

d. Adicionar 48 µl de la mezcla maestra de PCR primario a cada tubo de 0.5 ml debidamente etiquetado.

e. A los tubos correspondientes a las reacciones 1A a 5A, adicionar 1µl de GSP1. Y a los tubos que correspondan a las reacciones 6A a 8A, adicionar 1µl de PCP1.

f. Adicionar 1µl de cada una de las bibliotecas de DNA a los cuatro tubos dispuestos para ello; y adicionalmente añadir 1 µl de la biblioteca control y un 1 µl de la biblioteca realizada con DNA control, a cada uno de los controles positivos.

g. Adicionar 1µl de H<sub>2</sub>O a cada control negativo.

h. Centrifugar los tubos en una microcentrifuga.

i. Siguiendo los parámetros de un ciclo de dos pasos, aplicar el ciclo térmico siguiente (Tabla 3).

<b>Tabla 3. Ciclo térmico del PCR primario.</b>		
• 7 ciclos		
	94° C	25 sec
	72° C	3 min
• 32 ciclos		
	94° C	25 sec
	67° C	3 min
• 1 ciclo	67° C	7 min

- j. Analizar 5  $\mu$ l del producto de PCR primario en un gel de agarosa/EtBr al 1.5%. Si no se observa algún producto aumentar cinco ciclos adicionales.
- k. En un tubo de 0.5 ml para cada muestra, incluyendo los controles positivos y negativos, diluir 1  $\mu$ l del producto de la reacción primaria de PCR en 49  $\mu$ l de agua desionizada.

### 5.2.2 PCR secundario

- l. Preparar una mezcla maestra de ocho reacciones más una adicional, tal como se indica en la Tabla 4.

**Tabla 4. Mezcla maestra del PCR anidado ó secundario.**

	<u>9 rxns</u>	<u>Por rxn</u>
H <sub>2</sub> O desionizada	360 $\mu$ l	40 $\mu$ l
10X Advantage 2 PCR Buffer	45 $\mu$ l	5 $\mu$ l
dNTP (10 mM each)	9 $\mu$ l	1 $\mu$ l
AP2 (10 mM)	9 $\mu$ l	1 $\mu$ l
Advantage 2 Polymerase Mix (50X)	9 $\mu$ l	1 $\mu$ l
Volumen Total	432 $\mu$ l	48 $\mu$ l

- m. Mezclar mediante vortexeo, buscando no introducir burbujas y centrifugar ligeramente en una microcentrifuga.
- n. Adicionar 48  $\mu$ l de la mezcla maestra de PCR secundario a cada tubo de 0.5 ml debidamente etiquetado.
- o. De la reacción 1B a la 5B, adicionar 1  $\mu$ l del primer GSP2 a cada tubo. De la reacción 6B a 8B adicionar 1  $\mu$ l de PCP2 a cada tubo.
- p. Adicionar 1  $\mu$ l de cada una de las diluciones de los productos de PCR primario a los tubos correspondientes, por ejemplo 1A a 1B.
- q. Centrifugar ligeramente los tubos en una microcentrifuga.
- r. Siguiendo los parámetros de un ciclo de dos pasos, aplicar el ciclo térmico siguiente (Tabla 5).

<b>Tabla 5. Ciclo térmico del PCR anidado</b>		
• 5 ciclos		
	94° C	25 sec
	72° C	3 min
• 20 ciclos		
	94° C	25 sec
	67° C	3 min
• 1 ciclo	67° C	7 min

s. Analizar 5  $\mu$ l del producto de PCR primario en un gel de agarosa/EtBr al 1.5%. Si no se observa producto aumentar a cuatro ciclos adicionales.

### **5.3 Purificación de los distintos fragmentos amplificados, a partir de gel de agarosa**

Después de llevarse a cabo la segunda reacción de PCR es posible obtener a más de una banda en cada una de las cuatro bibliotecas; por lo que es importante contar con un método de purificación y aislamiento de los amplicones.

#### **5.3.1 Purificación de fragmentos de ADN empleando la metodología del Zymoclean™ Gel DNA Recovery Kit.**

- a) Cortar el fragmento de ADN del gel de agarosa usando navaja y transferirlo a un tubo de microcentrifuga de 1.5 ml.
- b) Adicionar tres volúmenes de ADB por cada volumen de agarosa. Por ejemplo si es que se han tomado 100  $\mu$ l (mg) del gel de agarosa, adicionar 300  $\mu$ l de ADB .
- c) Incubar a 37-55° C por 5-10 minutos hasta que la agarosa sea completamente disuelta.
- d) Transferir la solución de agarosa disuelta a una Columna Zymo-Spin™ en un Tubo Colector.
- e) Centrifugar a >10,000 x g por 30-60 seg. Desechar el sobrenadante.
- f) Adicionar 200  $\mu$ l del Buffer de lavado a la columna y centrifugar >10,000 x g por 30 seg. Desechar el sobrenadante. Repetir el paso de lavado.
- g) Adicionar >6  $\mu$ l de agua directamente a la matriz de la columna. Colocar la columna en un tubo de 1.5 ml y centrifugar >10,000 x g por 30-60 seg con el fin de eluir al ADN.

### **5.4 Clonación de los fragmentos de ADN en el vector pGEM T-Easy**

#### **5.4.1 pGEM T-Easy, un vector de clonación**

El vector pGEM T-Easy es un vector linearizado que cuenta en cada uno de sus bordes con una única timina, la cual se encuentra en el extremo 3' terminal (Fig. 4 ). Estas proyecciones de timina ubicadas en los extremos 3' del vector, flanquean lo que se convertirá en el sitio de inserción de todos aquéllos productos de PCR que posean en sus extremos 3' una cola-poli A. La cola-poli A es una sucesión de adeninas en el extremo 3' de un fragmento de ADN amplificado, la cual es añadida durante la reacción de PCR por ciertas DNA polimerasa, por ejemplo, la *Taq polimerasa*.

El vector pGEM T-Easy posee en su secuencia a las regiones promotoras T7 y SP6 (Fig. 4 y 5), a las cuales se unen respectivamente la RNA polimerasa T7 y SP6. Las regiones promotoras T7 y SP6 flanquean el sitio múltiple de clonación. Tanto el sitio múltiple de clonación como las regiones promotoras T7 y SP6, se encuentran dentro de la región que codifica para el péptido  $\alpha$  de la enzima  $\beta$ -galactosidasa. Cuando la reacción de ligado entre el vector pGEM T-Easy y el producto de PCR es exitosa; la secuencia de lac Z se ve

interrumpida y la traducción del péptido  $\alpha$  de la enzima  $\beta$ -galactosidasa no se lleva a cabo de manera exitosa.

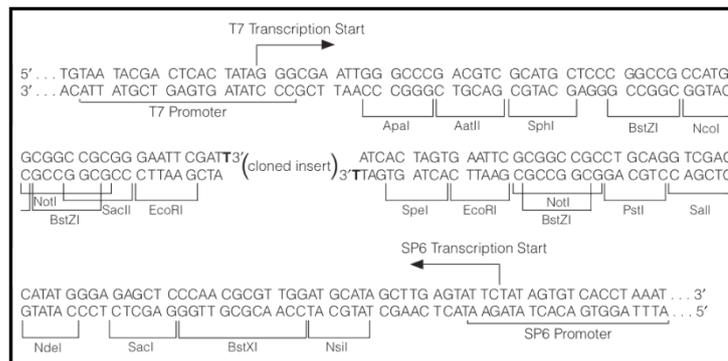


Figura 4. Sitio múltiple de clonación del vector pGEM T-Easy .

La selección de células transformantes con este vector, se efectúa mediante su crecimiento en un medio adicionado con ampicilina; el cual es el antibiótico para el cual confiere resistencia el vector pGEM T-Easy. Sin embargo, las células transformantes pueden pertenecer a dos distintos grupos de colonias. En un primer grupo, se encuentran aquellas células que han incorporado al plásmido recircularizado; mientras que en un segundo grupo, se encuentran aquellas donde el evento de ligación ha sido exitoso, y se ha llevado a cabo finalmente la inserción de la secuencia al vector. Ya que el medio además posee a un análogo a la lactosa, X-Gal, es posible distinguir entre el primer grupo de colonias transformantes las cuales poseen una coloración azul, del segundo grupo de colonias recombinantes las cuales crecen con una coloración blanca. Esto se debe a que X-Gal, es un análogo a la lactosa el cual es hidrolizado mediante la acción de la enzima  $\beta$ -galactosidasa, produciendo un índigo insoluble que proporciona la coloración azul al grupo de colonias con el plásmido recircularizado, ya que éstas poseen la secuencia intacta de la enzima.

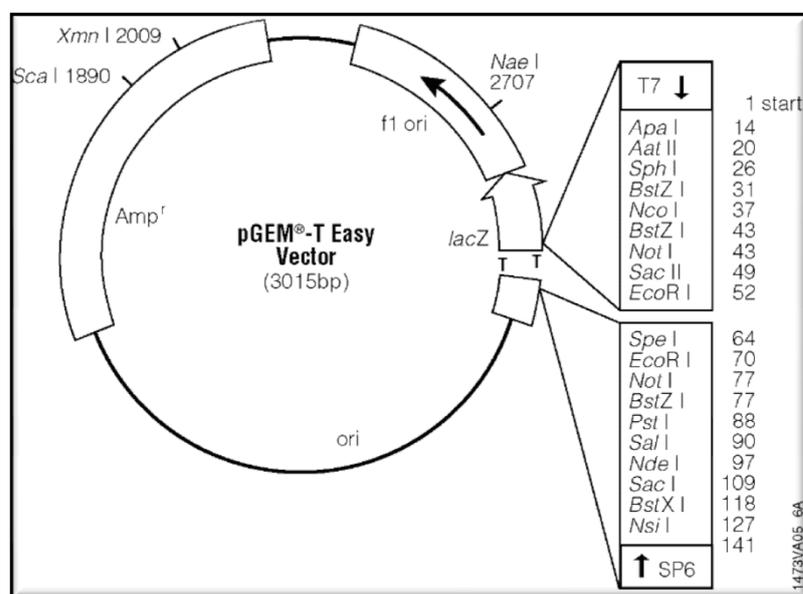


Figura 5. Mapa del vector pGEM T-easy. Donde se observa a la secuencia lac Z interrumpida por el sitio múltiple de clonación y las regiones promotoras T7 y SP6.

### 5.4.2 Reacción de ligado

La reacción de ligado entre el vector pGEM T-Easy y los diferentes fragmentos de ADN purificados a partir del gel de agarosa, se lleva a cabo empleando a la enzima T4 ADN ligasa. El protocolo de esta reacción se muestra a continuación:

- a) Adicionar y mezclar en un tubo las cantidades que se muestran en la Tabla 6.
- b) Incubar la mezcla a temperatura ambiente durante 1 hr. ó bien durante toda la noche a 4° C.

**Tabla 6. Cantidades y reactivos empleados en la reacción de ligado del vector pGEM T-Easy**

Reactivo	Volumen
ADN	1-7 $\mu$ l
Plásmido <i>pGEM T-easy</i> Vector 50 ng/ $\mu$ l	1.0 $\mu$ l
Solución amortiguadora de la reacción de ligado (10x)	1.0 $\mu$ l
Ligasa T4 DNA	1.0 $\mu$ l
H <sub>2</sub> O	c.b.p. 10 $\mu$ l
	10 $\mu$ l de Volumen Total

### 5.4.3 Transformación por choque térmico de células de competentes

El protocolo de la transformación por choque térmico de células competentes es el siguiente:

- a) Agregar 2  $\mu$ l de la reacción de ligado a 100  $\mu$ l de células competentes de *E.coli*.
- b) Incubar 30 min. en hielo.
- c) Incubar a 42° C durante 42-45 seg.
- d) Colocar en hielo 2 min.
- e) Adicionar 400  $\mu$ l de medio LB con glucosa 20 mM.
- f) Incubar 2 h a 37° C con agitación constante (200-225 rpm).
- g) Centrifugar durante 2 min. a 11,000 rpm.
- h) Decantar el sobrenadante.
- i) Resuspender el paquete celular en 400  $\mu$ l de medio LB.
- j) Sembrar en medio LB (Luria Broth) + ampicilina 50 mg/mL + IPTG 0.5 mM + X-gal 80 mg/mL
- k) Incubar al menos 12h a 37° C

## 5.5 Purificación de ADN plasmídico de colonias transformantes por el método de lisis alcalina

Este protocolo es usado para la extracción de ADN plasmídico a partir de una suspensión de células bacterianas; y esta basado en el procedimiento de lisis

alcalina desarrollado por Birnboim y Doly (1979). El cual se basa en la diferencia de topologías entre el DNA plasmídico y el DNA cromosomal.

Los plásmidos son moléculas pequeñas de ADN supercoloidal; mientras que el ADN bacteriano cromosómico es de mayor tamaño y menos supercoloidal. Esta diferencia es lo que permite mantener en suspensión al ADN plásmidico y a las moléculas de RNA, y en el precipitado de manera selectiva al ADN cromosomal y a las proteínas celulares.

Las células bacterianas son primeramente lisadas bajo condiciones alcalinas; lo cual desnatura a ambos ácidos nucleicos y a las proteínas. Luego se procede a neutralizar la solución mediante la adición de acetato de potasio. En este paso se precipita al ADN cromosomal y a las proteínas, por que resulta imposible para ellas plegarse nuevamente de manera correcta debido a su tamaño. En cambio los plásmidos se pliegan nuevamente de manera correcta, por lo que al neutralizar la solución estos permanecerán en solución.

- A. Centrifugar el cultivo con el fin de precipitar a las células.
- B. Usando una pipeta retirar y descartar el sobrenadante. Siendo cuidadoso de no perturbar el precipitado celular.
- C. Adicionar 100 ml de Buffer 1 (50 mM Tris-HCl, 10 mM EDTA; 100 mg/mL RNasa, pH 8.0) y resuspender el precipitado celular mediante vortexeo. Es muy importante que la suspensión de células sea homogénea y que no se observen grumos.

Las células en este paso se encuentran resuspendidas en una solución buffer con RNase. Cuando las células sean lisadas en el siguiente paso, la RNasa catalizará la hidrólisis de las moléculas de RNA en nucleótidos, mientras que el DNA no se verá afectado.

1. Adicionar 200  $\mu$ l de Buffer 2 (1% SDS, 0.2 M NaOH). Mezclar la solución mediante inversión, pero no vortexear. Esto último podría provocar que el DNA cromosomal liberado de las células lisadas se fragmentara y ocasionará una contaminación del producto.

El SDS es un detergente iónico que disrumpe a las membranas celulares y desestabiliza todas las interacciones hidrofóbicas que mantienen a varias macromoléculas en su conformación nativa. El pH alto provocado por la concentración de NaOH 0.2M, desnatura a las macromoléculas cambiando los estados de ionización de los grupos funcionales (ionizando ciertos grupos y desionizando otros).

2. Dejar la solución en hielo por al menos 5 minutos
3. Adicionar 150  $\mu$ l de Buffer 3 (3.0 M Acetato de Potasio, pH 5.5) previamente enfriado. Mezclar la solución invirtiendo rápidamente el tubo un par de veces. Un precipitado blanco se forma en este paso.

Este es el paso clave del procedimiento en la lisis alcalina. El pH bajo de la solución de acetato de potasio neutraliza al NaOH. Cuando el pH regresa casi a la neutralidad las macromoléculas se renaturalizan. Sin embargo las

moléculas de DNA de gran tamaño y las proteínas no se renaturalizan de manera correcta. Las moléculas de DNA plasmídico, sin embargo nunca se desnaturalizan por completo por se trata de moléculas pequeñas y circulares las cuales son supercoloidales. A pesar del rompimiento de los puentes de hidrógeno entre los pares de bases, debido al alto pH; estos se reforman de manera correcta cuando hay una disminución del pH. El DNA cromosómico y las proteínas forman precipitados, debido a que ellas se unen unas o otras en un agregado de gran tamaño; pero los plásmidos no precipitan debido a que ellos renaturalizan de manera correcta y no forman parte de estos agregados multi- moleculares.

5. Dejar la solución en hielo por al menos 5 min.

6. Colocar los tubos en un centrifuga, y centrifugar a máxima velocidad durante 5 min. El precipitado se formará a lo largo de uno de los lados del tubo.

7. Transferir el sobrenadante en un tubo limpio de 1.5 mL; teniendo cuidado de no tomar parte del precipitado. Conservar el tubo con el sobrenadante.

8. Medir el volumen con pipeta, y añadir un tercio del volumen de acetato de amonio y un volumen de isopropanol. Mezclar por inversión y dejar precipitando a  $-20^{\circ}$  C durante toda la noche ó al menos durante 1 hr.

El DNA es una molécula iónica (cargada negativamente debido a los grupos fosfato) la cual es altamente soluble en agua pero no en solventes orgánicos. La concentración de isopropanol causa que el DNA se vuelva insoluble y precipite.

9. Centrifugar a máxima velocidad durante 15 min. Y desechar el sobrenadante.

10. Añadir 200  $\mu$ l de etanol absoluto y mezclar por inversión un par de veces.

11. Centrifugar a máxima velocidad por 5 min.

12. Remover y desechar el sobrenadante de manera cuidadosa .

13. Colocar el tubo en un estufa a  $37^{\circ}$  C con el fin de evaporar el EtOH restante.

14. Resuspender el precipitado en H<sub>2</sub>O desionizada. La cantidad de ésta, dependerá de la concentración a la que se desea obtener la muestra de DNA plasmídico.

## **5.6 Clonación de un fragmento de ADN con posible actividad promotora en un vector de expresión binario**

La Tecnología Gateway® es un método de clonación universal basado en la recombinación sitio específica, el cual tiene por base al sistema de recombinación del bacteriófago lambda (Landy, 1989). La Tecnología Gateway® provee una forma rápida y eficiente de trasladar secuencias de ADN en múltiples sistemas de vectores con la finalidad de realizar el análisis funcional de una secuencia, como puede ser el caso de una región con posible actividad promotora, ó bien la expresión de un proteína bajo condiciones específicas (Hartle *et al.*, 2000).

### 5.6.1 Clonación en el vector de entrada pENTR™/D-TOPO®

El primer paso de la clonación, es la generación de un clona de entrada, una vez generada ésta, el producto de PCR entra por recombinación al vector de selección del Sistema Gateway®.

Existen un gran número de sistemas de clonación que permiten generar a la clona de entrada. Sin embargo, en el laboratorio se decidió optar por el sistema pENTR™ Directional TOPO® Cloning Kits; el cual, permite clonar a los productos de PCR con extremos romos de manera direccional en el vector de entrada pENTR™/D-TOPO®. A través de este sistema los insertos son clonados en la orientación correcta con una alta eficiencia, y esto se logra simplemente añadiendo cuatro bases (CACC) al extremo 5' de la secuencia del *forward*.

El diseño de oligos fue el siguiente:

*Forward* : **FPET32** 5' **CACC** TGG TCC TGA CAC TTA CCA ACA CAT G 3'

*Reverse* : **R32BAMHI** 5' CGC GGATCC GTT TAC GCA AGG **AGT GTG** 3'

En el extremo 3' del oligo *reverse* se añadió un sitio de reconocimiento para la enzima *BamHI*.

La reacción de clonación TOPO®, se efectúa mediante la mezcla del producto de PCR y el vector pENTR™/D-TOPO®, a la cual se añade solución de sales provista por el kit y agua estéril para completar un volumen de reacción final de 6µl. La reacción se deja incubar a temperatura ambiente durante 5 min. Sin embargo, para productos de PCR con una longitud mayor a un 1 Kb, el tiempo de incubación recomendado se extiende hasta un máximo de 30 min.

La selección de las colonias transformantes se realiza mediante su crecimiento en medio semisólido LB a una concentración de kanamicina de 50 µg/ml, antibiótico para el cual confiere resistencia el vector de clonación pENTR™/D-TOPO®. Una vez aisladas y seleccionadas a las colonias transformantes en medio semisólido, éstas son puestas a crecer durante toda la noche, en medio líquido LB más kanamicina, a una concentración de antibiótico de 100 µg/ml. Al día siguiente se realiza la extracción de plásmido con la finalidad de identificar a las colonias que son recombinantes de las que sólo son transformante.

### 5.6.1 Clonación en el vector binario de expresión pBGWFS7

Una vez obtenida la clona de entrada, se continua con la reacción de recombinación (LR). La cual tiene por objeto, facilitar el intercambio entre la secuencia flanqueada por los sitios *attL* de la clona de entrada, y la secuencia flanqueada por los sitios *attR* del vector destino. De esta manera es posible transferir, por ejemplo, a una secuencia con posible actividad promotora, del vector de entrada al vector destino. Además, tras el proceso de recombinación, se obtiene finalmente a la clona de expresión, la cual posee a los sitios *attB*.

El vector destino del Sistema Gateway® seleccionado fue **pBGWFS7**, cuya longitud es de 12,451 pb . Este vector destino, el cual es un vector binario de expresión permite probar la actividad promotora de una secuencia y posee los sitios **attR1** y **attR2**, los cuales recombinan a través de la reacción **LR** con los sitios **attL1** y **attL2** del vector pENTRTM/D-TOPO®. De esta manera ocurre el intercambio de secuencias entre aquella con posible actividad promotora del vector de entrada y la secuencia del gen letal, **ccd**, del vector destino.

El vector pBGWFS7, confiere resistencia a espectinomicina y estreptinomicina. Además, posee a la secuencia del gen **bar**, el cual confiere resistencia al herbicida Basta ®, cuyo ingrediente activo es glufosinato de amonio. En su secuencia río abajo del sitio de inserción del promotor, se encuentra la secuencia de los genes reporteros **GFP** y **GUS**. Mediante la expresión de los genes **GFP** y **GUS** en plantas transgénicas de *Nicotiana*, es posible evaluar la actividad promotora de la secuencia río arriba del gen *NaStEP*.

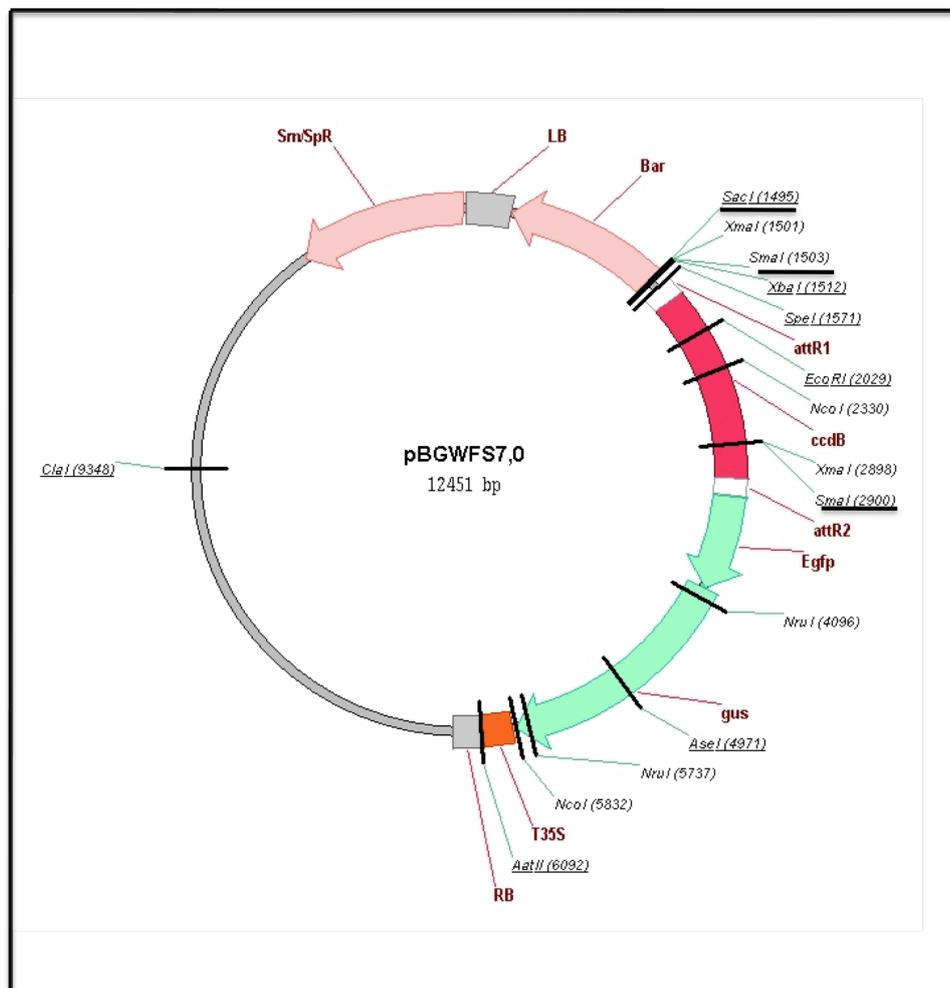


Figura 6. Vector pBGWFS7

La reacción LR se efectuó usando al mix de enzimas de Gateway® LR Clonase™ II, el vector destino pBGWFS7 y el plásmido purificado de la clona de entrada. El protocolo *Gateway® Technology with Clonase™ II*, propone efectuar la reacción LR de acuerdo a lo mostrado en la tabla 7.

**Tabla 7. Reacción LR de acuerdo al protocolo de *Gateway® Technology with Clonase™ II***

Reactivo	Cantidad
Vector de entrada	50 a 150 ng
Vector destino	1 µl a una concentración de 150 ng/ µl
Buffer TE, pH8	c.b.p. 8 µl
Mix de enzimas de Gateway® LR Clonase™ II	2 µl
<b>Volumen final</b>	<b>10 µl</b>

La transformación se realiza mediante el protocolo de choque térmico, sobre la cepa de células competentes *E.Coli* TOP10, empleando 2 µl del producto de la reacción **LR**. Las colonias transformantes, son seleccionadas a través su crecimiento en medio LB semisólido a una concentración de espectinomicina de 100 µg/ml. La selección y el análisis de colonias recombinantes, parte del conocimiento de que el vector pBGWFS7 confiere resistencia para espectinomicina y estreptomycin. El vector pBGWFS7 confiere además resistencia a cloranfenicol, cuando la reacción de recombinación no se ha llevado a cabo. Una vez que el proceso de recombinación ha ocurrido de manera exitosa entre los sitios *attR* del vector destino y los sitios *attL* del vector de entrada pENTRTM/D-TOPO®, la resistencia a cloranfenicol se pierde, junto con el gen letal *ccd*. Debido a que *E.Coli* TOP10 es una sensible a este gen letal, no es posible aislar colonias transformantes no recombinantes, lo cual es un seguro más de este sistema de clonación.

## 6. *RESULTADOS*

## **6.1 CLONACIÓN, SECUENCIACIÓN Y ANÁLISIS DE LOS FRAGMENTOS DE ADN GENÓMICO ADYACENTES AL EXTREMO 5' DE LA SECUENCIA DEL GEN *NaSTEP*.**

### **6.1.1. Obtención de fragmentos de ADN genómico adyacentes al extremo 5' de la secuencia del gen *NaSTEP***

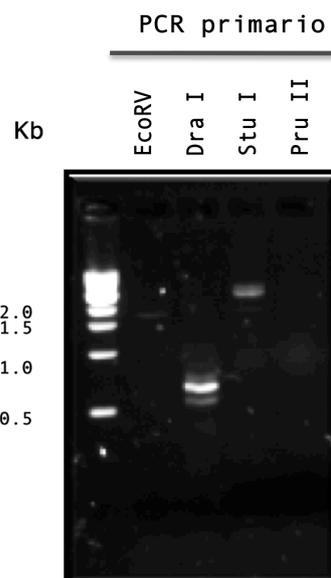
Con la finalidad de obtener fragmentos de ADN genómico adyacentes al extremo 5' de la secuencia del gen *NaSTEP*, se construyeron cuatro bibliotecas a partir de ADN genómico de *Nicotiana alata*. Cada una de estas bibliotecas consistió en fragmentos de ADN genómico ligados a dos adaptadores específicos.

El nombre de cada una de éstas, se derivó de la enzima de restricción con la cual fue digerido el ADN, así las cuatro bibliotecas fueron nombradas como: *EcoRV*, *Dra I*, *Stu I* y *Pru I*.

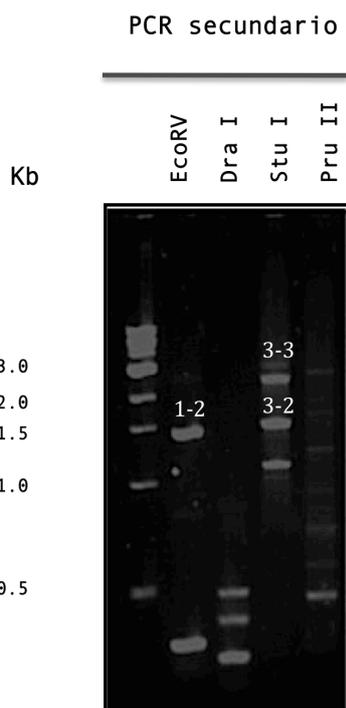
A partir de cada biblioteca se realizaron dos amplificaciones consecutivas. Los resultados de la primera reacción de amplificación, PCR primario, se muestran en la Figura 7. En esta figura, es posible observar la amplificación de un fragmento en cada una de las cuatro bibliotecas, los cuáles varían en cuanto a su tamaño e intensidad. El producto de esta reacción primaria se utilizó posteriormente como templado de la segunda reacción de amplificación, ó PCR secundario (Figura 8).

En la Figura 8 se observa la amplificación de una gama de fragmentos nuevamente de intensidad variable que van desde pesos moleculares por debajo de 0.5 Kb hasta el mayor de 3.0 Kb. Sin embargo, a diferencia del PCR primario donde sólo se observó un fragmento en cada una de las cuatro bibliotecas, la amplificación por PCR secundario produjo más de un fragmento de cada una de éstas (*EcoRV*, *Dra I*, *Stu I* y *Pru I*). Así, de la amplificación por PCR secundario de la biblioteca *PvuII* se obtuvieron un total de tres amplicones. A los dos productos de mayor tamaño, cuyos pesos moleculares correspondieron a 1.5 y 3.0 Kb, se les nombró como 3-2 y 3-3 respectivamente. De la amplificación mediante PCR secundario de la

biblioteca *EcoRV* se obtuvieron dos fragmentos, teniendo el de mayor tamaño un peso molecular de 1.5 Kb, al cual finalmente se le nombró como fragmento 1-2.



**Fig. 7. PCR primario.** Fraccionamiento electroforético de los productos de PCR primario de cada una de las cuatro bibliotecas . Se observa la amplificación de un fragmentos en cada una de las bibliotecas, los cuáles varían en cuanto a su tamaño e intensidad. La separación se realizó en un gel de agarosa al 1.5%, teñido con bromuro de etidio; M: Marcador de 1kb, *EcoRV*: biblioteca digerida con *EcoRV*, *Dra I*: biblioteca digerida con *Dra I*, *Stu I*: biblioteca digerida con *Stu I*, *Pru II*: biblioteca digerida con *Pru II* .

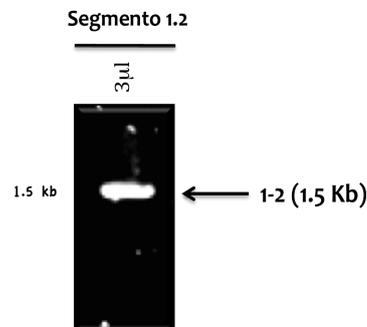


**Fig. 8. PCR secundario ó anidado.** Fraccionamiento electroforético de los productos de PCR utilizando como templado al producto de la reacción de PCR primario de cada una de las bibliotecas. 1-2, 3-2 y 3-3: fragmentos que muestran mayor intensidad y tamaño igual o mayor a 1.5 kb. La separación se realizó en un gel de agarosa al 1.5%, teñido con bromuro de etidio; M: Marcador de 1kb, *EcoRV*: producto amplificado de la biblioteca *EcoRV*, *Dra I*: producto amplificado de la biblioteca *Dra I*, *Stu I*: producto amplificado de la biblioteca *Stu I*, *Pru II*: producto amplificado de la biblioteca *Pru II* .

### 6.1.2. Clonación de los fragmentos 1-2, 3-2 y 3-3.

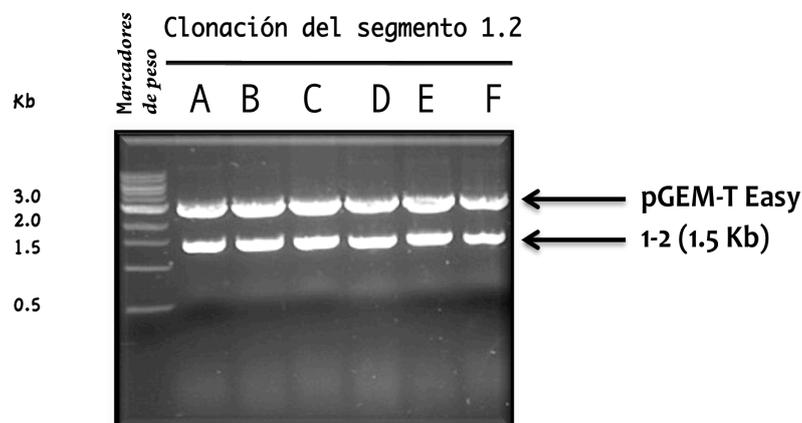
Sólo los fragmentos de mayor tamaño 1-2, 3-2 y 3-3 se clonaron posteriormente en el vector pGEM T- easy.

La purificación del fragmento 1-2 cuyo tamaño correspondió aproximadamente a 1.5 Kb se muestra en la Figura 9 .



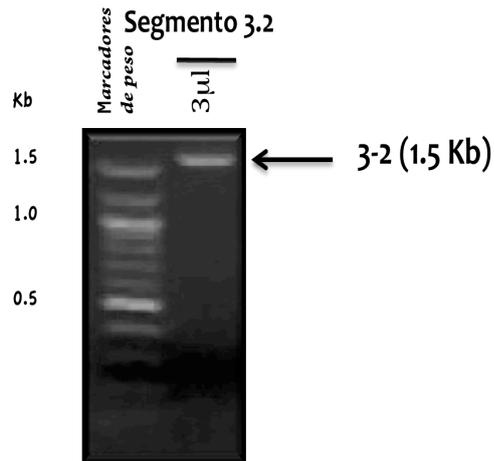
**Fig. 9. Fragmento 1-2 purificado.** Electroforesis del fragmento 1-2 purificado a partir de gel de agarosa. La banda observada se encuentra a una altura de 1.5 KB, tamaño correspondiente al segmento 1-2. Se cargaron 3 µl de un volumen final de resuspensión del 12 µl a un gel de agarosa al 1% teñido previamente con bromuro de etidio.

Con la finalidad de comprobar la inserción del segmento 1.2 en el vector pGEM T-Easy, se llevó a cabo la reacción de digestión con la enzima *EcoRI* del plásmido purificado de seis de las colonias transformantes. En la Figura 10, se observa la liberación en el plásmido proveniente de cada una de las 6 colonias de un fragmento de ADN con un tamaño de 1.5 Kb, el cual corresponde al del fragmento 1-2.



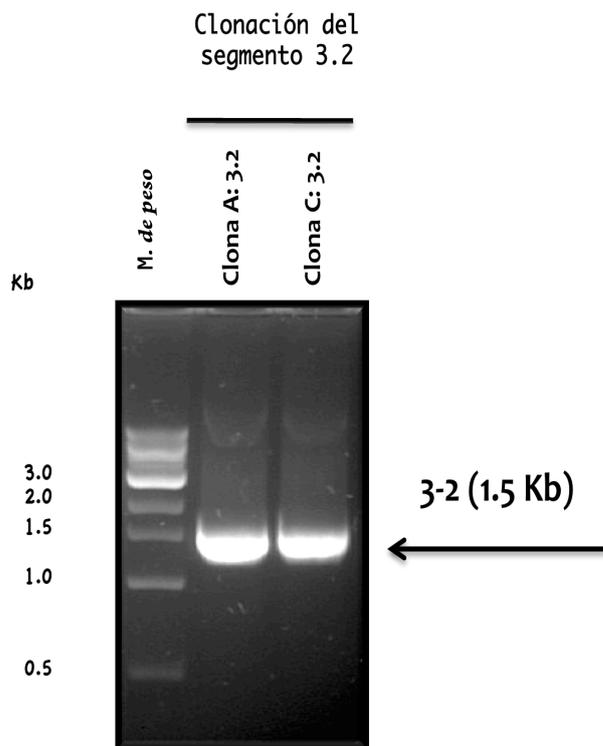
**Fig. 10. Clonación del fragmento 1-2.** Análisis de restricción y liberación del inserto 1-2 del vector pGEM-T-Easy. Reacción de digestión con las enzima *EcoRI* de 6 distintas colonias transformantes. Se observa la liberación de un fragmento de 1.5 Kb en cada una de las colonias analizadas. Electroforesis en gel de agarosa al 1.0% teñido previamente con bromuro de etidio. A, B, C, D, E, F: Clonas transformantes y recombinantes del fragmento 1-2 ligado al vector pGEM-T-Easy . Marcadores de peso de 1kb.

La purificación del fragmento 3-2 cuyo tamaño correspondió a 1.5 Kb se muestra en la Figura 11.



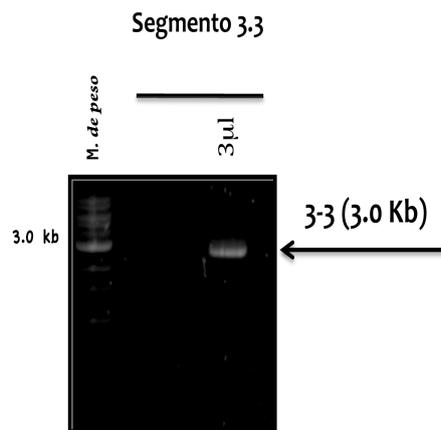
**Fig. 11. Fragmento 3-2 purificado.** Electroforesis del fragmento 3-2 purificado a partir de gel de agarosa. La banda observada se encuentra a una altura de 1.5 Kb, tamaño correspondiente al del segmento 3-2. Se cargaron 3 µl de un volumen final de resuspensión del 12 µl a un gel de agarosa al 1% teñido previamente con bromuro de etidio. Marcadores de peso de 100 pb.

Con la finalidad de comprobar la inserción del segmento 3-2 en el vector pGEM T-Easy se llevó a cabo una reacción de PCR, la cual se muestra en la figura 12. En la cual se observa la amplificación de un fragmento a la altura de 1.5 Kb, correspondiente al tamaño del fragmento 3-2.



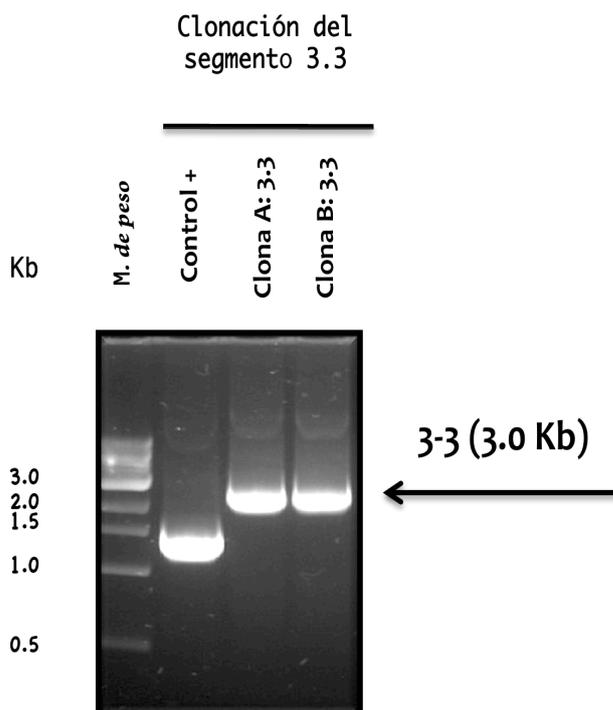
**Fig. 12. Clonación del fragmento 3-2.** PCR confirmatorio de la presencia del inserto 3-2 en el vector pGEM T-Easy. Electroforesis en gel de agarosa al 1.0% teñido previamente con bromuro de etidio. Clona A:3-2, Clona C:3-2; clonas positivas para el fragmento 3-2 ligado al vector pGEM T-Easy. Marcadores de peso de 1kb.

La purificación del fragmento 3-3 cuyo tamaño correspondió aproximadamente a 3.0 Kb. se muestra en la Figura 13.



**Fig. 13. Fragmento 3-3 purificado.** Electroforesis del fragmento 3-3 purificado a partir de gel de agarosa. La banda observada se encuentra a una altura de 3.0 KB, tamaño correspondiente al del segmento 3-3. Se cargaron 3 µl de un volumen final de resuspensión del 12 µl a un gel de agarosa al 1% teñido previamente con bromuro de etidio. Marcadores de peso de 1kb.

Con la propósito de comprobar la inserción del segmento 3-3 en el vector pGEM T-Easy se llevó a cabo una reacción de PCR (Figura 14). En esta figura es posible observar la amplificación de un fragmento a la altura aproximada de 3.0 Kb, correspondiente al tamaño del fragmento 3-3.



**Fig. 14. Clonación del fragmento 3-3.** PCR confirmatorio de la presencia del inserto 3-3 en el vector pGEM-T-Easy. Electroforesis en gel de agarosa al 1.0% teñido previamente con bromuro de etidio. Clona A:3-3, Clona B:3-3; clonas positivas para el fragmento 3-3 ligado al vector pGEM-T-Easy. Marcadores de peso de 1kb.

### **6.1.3. Secuenciación y análisis de los fragmentos 1-2, 3-2 y 3-3.**

Las secuencias de los fragmentos de ADN genómico 1-2, 3-2 y 3-3 adyacentes al extremo 5' de la secuencia de cDNA del gen *NaStEP* y clonados en el vector pGEM-T Easy, se obtuvieron a través de la Unidad de Secuenciación de ADN del Instituto de Biotecnología de la UNAM.

El análisis de las secuencias de los distintos fragmentos, 1-2, 3-2 y 3-3, se realizó mediante el uso del servidor 2.2.12 NCBI GenBank BLAST (Altschul *et al.*, 1997). En la búsqueda de identidad entre secuencias se utilizó el programa BLASTX del Centro de Supercómputo de San Diego (<http://workbench.sdsc.edu>). Para el alineamiento entre secuencias se hizo uso del programa ClustalX (Thompson *et al.*, 1997).

#### **6.1.3.1 Secuencia y análisis del segmento 1-2 clonado en el vector pGEM T-Easy**

Mediante la secuenciación con el oligo T7 del fragmento 1-2 clonado en el vector pGEM T-Easy se obtuvo una secuencia con una longitud de 1,480 pb. La cual muestra en su extremo 3' identidad con 52 pb de la secuencia de cDNA del gen *NaStEP*. Además, de presentarse en el mismo extremo la secuencia correspondiente al oligo *reverse* GSP2. Sin embargo, estos hallazgos no son suficientes para confirmar que la secuencia del segmento 1-2 corresponde a la región río arriba del gen *NaStEP*.

#### **6.1.3.2 Secuencia y análisis del segmento 3-2 clonado en el vector pGEM T-Easy**

La secuencia del segmento 3.2, posee una longitud total de 1,573 pb. Esta secuencia de ADN genómico, posee en su extremo 5' a la secuencia del oligo *forward*, AP2, y en su extremo 3' a la del oligo *reverse*, GSP2 (Figura 15). Además, la secuencia del fragmento 3-2 posee identidad al 100% en su extremo 3' con el extremo 5' de la secuencia de cDNA del gen *NaStEP* en una longitud de 154pb. Este resultado aseguró que la secuencia del fragmento 3.2 es la región contigua y río arriba del gen *NaStEP*. Siendo así que el segmento 3.2, poseen una alta probabilidad de contener parte de la secuencia promotora del gen *NaStEP*.

```

Segmento_3-2      GAATTCGATTACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGACACTTACCAACACATGGGGTCGGTTGTGATGATG
S.inicial_de_NaStEP -----
Segmento_3-2      CTACACTTTGCACATATGTGCAGATACCGAAGCGGCATTTGAGCAGCAGCACTAGATCGGGAGCCAACCTTCAATCCACCGAGAC
S.inicial_de_NaStEP -----
Segmento_3-2      ACCGAGGTAGCCTTGCAGGCGTCCGAGGCCGACGTTTCTCTATCTTACATTTCAATTTATGTTATCTCATGTATCCGAGACAA
S.inicial_de_NaStEP -----
Segmento_3-2      ATAGTGTATTATTCTTTCAAACGGTTGTATTAGTACTCTTAGTGGTCTGTGAGTAAATGTGACACCAGTCTTGGGTAGAGGC
S.inicial_de_NaStEP -----
Segmento_3-2      AAATGATGAGTTTCGCATTATTGTCCAGTTTATTTAATTTAAGTCTTCCGTTTTATTTTTATGTTTCGTGGTTATTAATTTGTTT
S.inicial_de_NaStEP -----
Segmento_3-2      AACACGTGTTGTTATAACTGATAAAAAGGATTTAAAATAAAAAGTGTGGAAGTTTATGAAATATGGCTTGCCTGGCTTCTACGAGT
S.inicial_de_NaStEP -----
Segmento_3-2      AGGCGCCATCACGACTCCCGAAGGTGGGAATTCAGATCGTGACAAGAAATGGGTCAATCTTTTTGGAACATACCAAAAAGGTAA
S.inicial_de_NaStEP -----
Segmento_3-2      TAGGTTACATAAACTGGAACAGAGGAATGTTATTTATATATTTAGATATAATTTAACTTTGAACTTTATATTTAATATTAATG
S.inicial_de_NaStEP -----
Segmento_3-2      ACATAATTTTATAAATGCAGAAATTTCTATGGCCATATTTACAACCTTAGTCCCATAGGGATGGTGAGAGTTCATAAGAGTCCGG
S.inicial_de_NaStEP -----
Segmento_3-2      AATTCAGATTTTCTTTGGGCGCTGGCCATCTGCCGGAGTTTCTTGTTTTAAATACCGATATTTTAAAAACAAATCCAAAAACAAC
S.inicial_de_NaStEP -----
Segmento_3-2      TTTACTGACAATTATATTTTTTTTTTTTATTTATGTCAGAAGAGTCATCTGAATATCTACAACAATTAATGGCAAGATTAACAATC
S.inicial_de_NaStEP -----
Segmento_3-2      AATATGCGCCACATAAACTAAAAGAGATAGACTATATGATTCATCGTTTGCCCATTAATAAATGAACGCAACATAGACTCT
S.inicial_de_NaStEP -----
Segmento_3-2      ACTACTCTCTCCGTTTCAATTTATGTTAACACATTTGACTGAGCAGGAGTTTAAAGAAAAGAGAGAAGACTTTTGAACCTTGTGGT
S.inicial_de_NaStEP -----
Segmento_3-2      GTAAAATAAAGCACATATATCTTTGTGTGGCTATAAATAGTTGCATAAAGGTAATGTTTCCAAATAGGAAAAGACGTTATTTTT
S.inicial_de_NaStEP -----
Segmento_3-2      TTTTGCACAGACTAAAAGAAAATAGTTTCACATAAATGAAACGGAGAAAAGTATAAAAACAATACACATATTTTAAACGTTCCAA
S.inicial_de_NaStEP -----
Segmento_3-2      TGTGAAGGATGGTACATGCAAGTATTGTAGCATCTTATACTCCTGCTAGTCATCTACGCACCTCTATAAATTCAGCTTCAGTAA
S.inicial_de_NaStEP -----
Segmento_3-2      TGGATTTGAGTATCAACAATACGCAACACACTCCTTGCGTAAACATGAAATCCTTTATTTTCAGCTTCCTCTTGCTTTCAACTA
S.inicial_de_NaStEP -----ATGAAATCCTTTATTTTCAGCTTCCTCTTGCTTTCAACTA
*****
Segmento_3-2      CTCCTCCTTGCTTCCTTTGTAGTCTTTTCATCATCTTTCACCTCCACCAATCCATTTGCTTCCCACTACTACTGATGATGA
S.inicial_de_NaStEP CTCCTCCTTTGCTTCCTTTGTAGTCTTTTCATCATCTTTCACCTCCACCAATCCATTTGCTTCCCACTACTACTGATGATGA
*****
Segmento_3-2      CAACGATTACCTGTACTTTCCGAAGTCAATC
S.inicial_de_NaStEP CAACGATTACCTGTACTTTCCGAAGTGC-----
*****

```

**Fig. 15. Secuencia río arriba de la región codificadora del gen *NaStEP* de 1,573 pb.** **Segmento\_3-2:** corresponde a la secuencia del fragmento 3.2 clonado en el vector pGEM-T Easy, con una extensión total de 1,573pb. **S.inicial\_de\_NaStEP:** corresponde a la secuencia con una longitud de 154pb del extremo 5' del cDNA del gen *NaStEP*. El alineamiento entre secuencias se realizó usando el ClustalW. Mediante este alineamiento es posible observar, la homología de secuencias entre el extremo 3' del fragmento 3-2 y el extremo 5' del gen *NaStEP*. En verde se indica a la secuencia donde se ancla el oligo *forward AP2*, el cual es empleado durante la reacción de PCR secundario junto con el oligo *reverse GSP2*, en amarillo. Los nucleótidos donde se observa homología entre secuencias están representados por (\*). Los guiones en la secuencia se introdujeron para mantener un buen alineamiento.

### **6.1.3.3 Secuencia y análisis del segmento 3-3 clonado en el vector pGEM T-Easy**

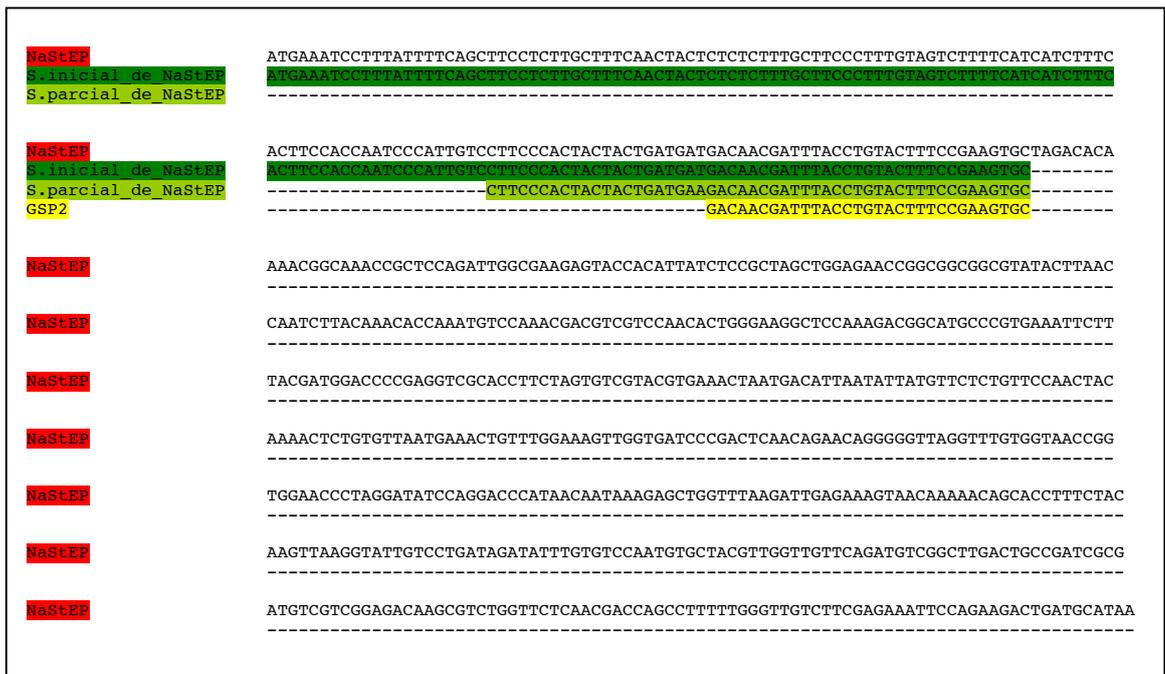
El fragmento 3.3 clonado en el vector pGEM T-Easy, posee un tamaño de aproximadamente 3.0 Kb, tal como se observa en el corrimiento electroforético de la Figura 14. No obstante, mediante su secuenciación con el oligo T7, sólo se obtuvo un fragmento de 1,166 pb. Este fragmento posee en su extremo 3' la secuencia del oligo *reverse GSP2*; sin embargo, sólo se observó identidad con 52 pb de la secuencia de cDNA del gen *NaStEP*, tal como ocurrió para la secuencia del fragmento 1.2. No siendo posible confirmar que la secuencia del segmento 3-2 pertenece en realidad a la región río arriba del gen *NaStEP*.

### **6.1.3.4 Identidad entre la secuencia de cDNA del gen *NaStEP* y los fragmentos de ADN 1-2, 3-2 y 3-3**

El alineamiento de los fragmentos 1-2, 3-2 y 3-3 con la secuencia del gen *NaStEP* respectivamente, da como resultado dos distintos patrones de identidad.

Cuando la secuencia de cDNA del gen *NaStEP* fue alineada con la secuencia del segmento 3-2, se presentó entre ambas una identidad de 154 pb. Estos 154 pb en el extremo 5' de la secuencia del gen *NaStEP* están indicados en la figura 16 con el nombre de *S. inicial de NaStEP*.

Al alinear la secuencia de los fragmentos 1-2 y 3-3 con la secuencia de cDNA del gen, sólo 52 pb de la secuencia de *NaStEP* presentó identidad con estos fragmentos. A este segmento de 52 pb se le nombró en la figura 16 como *S. parcial de NaStEP*.



**Fig. 16. Identidad entre la secuencia de cDNA del gen *NaStEP* y los fragmentos de ADN 1-2, 3-2 y 3-3.** NaStEP corresponde a la secuencia de cDNA del gen *NaStEP*. La S. inicial de NaStEP se refiere al segmento de la secuencia de cDNA del gen *NaStEP* que empalma con la secuencia del fragmento 3-2. La S. parcial de NaStEP se refiere al segmento de la secuencia de cDNA del gen *NaStEP*, que empalma con las secuencias de los fragmentos 1-2 y 3-3. En amarillo, se muestra la secuencia del primer *reverse GSP2*, el cual empalma de manera anidada con la secuencia del gen *NaStEP*. El alineamiento entre las distintas secuencias se realizó usando el ClustalW. Los guiones en la secuencia se introdujeron para mantener un buen alineamiento.

### 6.1.3. Los segmentos 1-2 y 3-3 se tratan probablemente de la misma región en el genoma de *N. alata*

Los segmentos 1-2 y 3-2 mostraron el mismo patrón de identidad con la secuencia de cDNA del gen *NaStEP*, de tal forma que se pensó en realizar un alineamiento entre ambas secuencias (Figura 17), con el propósito de averiguar si se trataban de una misma región del genoma de *N. alata*. La identidad que se observó mediante el alineamiento sugiere fuertemente que ambas secuencias podrían tratarse de una misma región en el genoma de *Nicotiana*.



**6.2 LA SECUENCIA 1.4 Kb RÍO ARRIBA DE LA REGIÓN CODIFICADORA DEL GEN *NaStEP*, POSEE VARIOS ELEMENTOS DE UNA REGIÓN PROMOTORA.**

**6.2.1 Análisis bioinformático de la región 1.4 Kb río arriba del gen *NaStEP*, mediante el uso del servidor *PLANT CARE***

Al analizarse la secuencia 1.4 Kb río arriba de la región codificadora del gen *NaStEP* mediante el servidor bioinformático *PLANT CARE*, se obtuvieron a 59 motivos correspondientes a la caja TATA; debido a que varios motivos aparecieron más de una vez en la secuencia, el conjunto de cajas TATAs se englobó en un total de 19 distintas secuencias, como se observa en la tabla 8. Este conjunto de cajas TATAs pertenece a un total 11 organismos distintos, entre los cuales se encuentra *N. tabacum*.

En la primer columna de la tabla 8 se muestra el organismo al cual pertenece la caja TATA, la segunda al número de nucleótidos conservados y la tercera a la secuencia misma; mientras que el número señalado entre paréntesis corresponde al número de veces que se encuentra presente ese motivo.

Además, el análisis bioinformático de la secuencia del promotor potencial de *NaStEP* mediante el servidor bioinformático *PLANT CARE*, no sólo resultó en la identificación de elementos reguladores propios del *core promoter*, sino también de motivos reguladores en *cis*; los cuales se muestran en la tabla 9. En esta tabla, la primer columna indica el nombre del elemento regulador en *cis*, la segunda la secuencia del motivo, la tercera el número de nucleótidos conservados (NC), la cuarta el organismo, la quinta la función del elemento regulador, y finalmente la última se refiere al número de identificación del gen en la base de datos del GENBANK.

**Tabla 8. Caja TATA.**

Organismo	Nucleótidos conservados	Secuencia	No. de veces presente
<i>Glycine max</i>	5	TAATA	(8)
<i>Lycopersicon esculentum</i>	5	TTTTA	(16)
<i>Arabidopsis thaliana</i>	9	TAAAAATAA	(1)
<i>Arabidopsis thaliana</i>	5	TATAA	(4)
<i>Arabidopsis thaliana</i>	4	TATA	(10)
<i>Oryza sativa</i>	8	TACATAAA	(1)
<i>Arabidopsis thaliana</i>	9	taTATAAAtc	(1)
Ac	7	TATAAAT	(3)
<i>Arabidopsis thaliana</i>	6	TATAAA	(4)
<i>Arabidopsis thaliana</i>	7	TATATAA	(1)
<i>Brassica napus</i>	6	ATATAT	(2)
<i>Brassica oleracea</i>	7	ATATAAT	(1)
<i>Pisum sativum</i>	8	TATAAAAT	(1)
<i>Arabidopsis thaliana</i>	7	TATAAAA	(1)
<i>Zea mays</i>	8	TTTAAAAA	(1)
<i>Brassica napus</i>	6	ATTATA	(1)
<i>Pisum sativum</i>	7	TATATG	(1)
<i>Nicotiana tabacum</i>	9	tcTATAAAta	(1)
<i>Arabidopsis thaliana</i>	9	tcTATATAtt	(1)

La posición de los distintos elementos reguladores en *cis* descritos en la tabla 9, se presentan en la figura 18 de acuerdo a su posición dentro de la secuencia de 1.4 Kb río arriba del gen *NaStEP*. El color con el que los distintos motivos fue indicado dentro de la secuencia de la figura 18, corresponde al mismo color con que el cual se marco a cada motivo en la tabla 9.

**Tabla 9. Elementos reguladores en cis presentes en la región 1.4 Kb río arriba de la secuencia codificadora del gen *NaStEP*, estimados mediante el servidor *PLANTCARE*.**

Nombre	Secuencia	N.C	Organismo	Función	ID del gen
4cl-CMA2b	TCTCACCAACC	10	<i>Petroselinum crispum</i>	Elemento regulador en <i>cis</i> , similar a la caja L y asociado con las cajas A y P, involucrado en la respuesta a la luz.	X05351
AAGAA-motif	GAAAGAA	7	<i>Avena sativa</i>		phyA3
ABRE	CACGTG	6	<i>Arabidopsis thaliana</i>	Elemento regulador en <i>cis</i> , involucrado en la respuesta a ácido abscísico.	U01377
AC-I	TCTCACCAACC	10	<i>Populus tremuloides</i>	Regulador negativo de la expresión en polen. Responsable de restringir la expresión vascular a el xilema.	AF041051
Box 4	ATTAAT	6	<i>Petroselinum crispum</i>	Elemento que forma parte del arreglo modular conservado, pal-CMA1, involucrado en la respuesta a luz.	X15473
Box I	TTTCAAA	7	<i>Pisum sativum</i>	Elemento de respuesta a luz	M21356
CTAG-motif	ACTAGCAGAA	9	<i>Avena sativa</i>		
G-Box	CACGTG CACATGG CACGTG ACACGTGT CACGAC	6 7 6 8 6	<i>Pisum sativum</i> <i>Solanum tuberosum</i> <i>Arabidopsis thaliana</i> <i>Brassica napus</i> <i>Zea mays</i>	Elemento regulador en <i>cis</i> , involucrado en la respuesta a luz.	
GAG-motif	AGAGAGT	7	<i>Arabidopsis thaliana</i>	Elemento que forma parte del arreglo modular conservado, rbcA-CMA1, involucrado en la respuesta a luz.	A86720
GCN4-motif	CAAGCCA	7	<i>Oryza sativa</i>	Motivo altamente conservado en todos los genes <i>glutelin</i> involucrado en la expresión específica del endosperma.	X54314
L-box	TCTCACCAACC	10	<i>Petroselinum crispum</i>	Elemento regulador en <i>cis</i> , asociado con la cajas A y P involucrado en la inducción de la actividad transcripcional pero no suficiente para conferir respuesta a la luz.	L37355
MBS	TAACTG	6	<i>Arabidopsis thaliana</i>	Sitio de unión de MYB involucrado en la inducción por drogas.	D13044
MSA-like	TCAAACGGT	9	<i>Catharanthus roseus</i>	Elemento regulador para el ciclo celular durante la fase M.	D86388
Skn-1 motif	GTCAT	5	<i>Oryza sativa</i>	Elemento regulador en <i>cis</i> requerido para la expresión en el endosperma en altos niveles en cooperación con otros motivos (AACA; GCN4; ACGT).	X54314
TCA-element	TCAGAAGAGG	9	<i>Brassica oleracea</i>	Elemento de actuación en <i>cis</i> , involucrado en la respuesta a ácido salicílico.	X98521
TCT-motif	TCTTAC	6	<i>Arabidopsis thaliana</i>	Parte de un elemento de respuesta a luz	L14749

```

+ GACGGCCCGG GCTGGTCTG ACACTTACCA ACACATGGGG TCGGTTGTGA TGATGCTACA CTTTGACAT ATGTGCAGAT
- CTGCCGGGCC CGACCAGGAC TGTGAATGGT TGTGTACCCC AGCCAACACT ACTACGATGT GAAACGTGTA TACACGTCTA

+ ACCGAAGCGG CATTGAGCA GCAGCACTAG ATCGGGAGCC AACCTCAAT CCACCGAGAC ACCGAGGTAG CCTTGCAGGC
- TGGCTTCGCC GTAAACTCGT CGTCGTGATC TAGCCCTCGG TTGGAAGTTA GGTGGCTCTG TGGTCCATC GGAACGTCCG

+ GTCCGAGGC CCGACGTTT CTCTATCCTA CATTTCATT ATGTTATCTC ATGTATCCGA GACAAATAGT GTATTATTTT
- CAGGCGTCCG GGCTGCAAAG GAGATAGAAT GTAAAGTAAA TACAATAGAG TACATAGGCT CTGTTTATCA CATAATAAAG

+ TTTCAAACGG TTGTATTTAG TACTCTTAGT GGTCTGTGAG TAAATGTGAC ACCAGTTCTT GGGTAGAGGC AAATGATGAG
- AAAGTTTGCC AACATAAATC ATGAGAATCA CCAGACACTC ATTTACACTG TGGTCAAGAA CCCATCTCCG TTTACTACTC

+ TTTTCGATTA TTGTCCAGTT TATTTAATTT AAGTCTCCG TTTTATTTTT ATGTTTCGTG GTTATTAAT TGTTTAACAG
- AAAGCGTAAT AACAGGTCAA ATAAATTTAA TTCAGAAGGC AAAATAAAAA TACAAAGCAC CAATAATTTA ACAAATTTGTG

+ GTGTTGTTAT AACTGATAAA AAGGATTTAA AATAAAAGTG TTGGAAGTTT ATGAATATGG CTTGCCTGGC TTCTACGAGT
- CACAACAATA TTGACTATTT TTCTAAATTT TATTTTTCAC AACCTTCAA TACTTATACC GAACGGACCG AAGATGCTCA

+ AGGCGCCATC ACGACTCCCG AAGGTGGGAA TTCCAGATCG TGACAAGAAA TGGGTCATTT TTTTGGAAC ATACCAAAAA
- TCCGCGGTAG TGCTGAGGGC TTCCACCCTT AAGGTCTAGC ACTGTTCTTT ACCCAGTAAG AAAAACCTTG TATGGTTTTT

+ GGTAATAGGT TACATAAAT GGAACAGAGG AATGTATT TATATATTTA GATATAATTT AACTTTGAAC TTTATATTTT
- CCATTATCCA ATGTATTTGA CCTGTCTCC TTAACAATAA ATATATAAAT CTATATTTAA TTGAAACTTG AAATATAAAA

+ AATATTAATG ACATAATTT ATAATTGCG AAATTTCTAT GGCCATATTT ACAACCTTAG TCCCATAGGG ATGGTGAGAG
- TTATAATTAC TGTATTTAAA TATTAACGTC TTAAAGATA CCGTATAAA TGTTGGAATC AGGGTATCC TACCACCTC

+ TTCATAAGAG TCCGGAATTC AGATTTTCTT TGGGCCGCTG GCCATCTGCC GGGAGTTTCT TGTTTAAATA CCGATATTTT
- AAGTATTCTC AGGCCTAAG TCTAAAAGAA ACCCGCGAC CGGTAGACGG CCCTCAAAGA ACAAATTTAT GGCTATAAAA

+ AAAAACAAT CAAAACAAC TTTACTGACA ATTATATTTT TTTTATTTA TGTCAGAAGA GTCATCTGAA TATCTACAAA
- TTTTGTTTA GGTTTGTG AAATGACTGT TAATATAAAA AAAATAAAT ACAGTCTTCT CAGTAGACTT ATAGATGTTT

+ CAATTAATGG CAAGATTTTA TCTCTTAAC AAATCAATAT GCGCCACATA AACTAAAAGA GATAGACTAT ATGATTCTCA
- GTTAATTACC GTTCTAAAAT AGAGAGATTG TTTAGTTATA CGCGGTGTAT TTGATTTTCT CTATCTGATA TACTAAGAGT

+ TCGTTTGCC ATTAATAATA ATGAACGCAA CATAGACTCT ACTACTCTCT CCGTTTCAAT TTATGTAAAC ACATTTGACT
- AGCAAACGG TAATTATTAT TACTTGCGTT GTATCTGAGA TGATGAGAGAGGCAAAGTTA AATACAATTG TGTAACCTGA

+ GAGCACGGAG TTTAAGAAA GAGAGAAGAC TTTTGAACCT GTGGTGATAA ATAAAGCACA TATATCTTGT GTGGCTATAA
- CTCGTGCCTC AAATCTTTT CTCTCTCTG AAAACTTGAA CACCACATTT TATTTCTGTT ATATAGAACA CACCGATATT

+ ATAGTTGCAT AAAGGTAAT TGTTCCAAA TAGGAAAAGA CGTTATTTTT TTTGACAG ACTAAAAGA AAATAGGTTT
- TATCAACGTA TTCCATTTA ACAAAGGTTT ATCCTTTTCT GCAATAAAAA AAAACGTGTC TGATTTTCT TTTATCCAAG

+ ACATAAATTG AAACGGAGAA AGTATAAAC AATACACATA TTTAAACGT TCCAATGTGA AGGATGGTAC ATGCAAGTAT
- TGTATTTAAC TTGCCTCTT TCATATTTTG TTATGIGTAT AAAATTTGCA AGGTTACTT TCCTACCATG TACGTTTATA

+ TGTAGCATCT TACTCTCTG CTAGTCATCT ACGCACCTCT ATAAATCAA GCTTCAGTAA TGGATTTGAG TATCAAACAA
- ACATCGTAGA ATATGAGGAC GATCAGTAGA TCGGTGGAGA TATTTAAGTT CGAAGTCATT ACCTAAACTC ATAGTTTGTG

+ TACGCAACAC ACTCCTTGCG TAAACAT
- ATGCGTTGTG TGAGGAACGC ATTTGTA

```

**Fig. 18. Posición de los distintos elementos reguladores en cis, dentro de la secuencia 1.4 Kb río arriba de la región codificadora del gen *NaStEP*, estimados por el servidor *PLANT CARE*.**

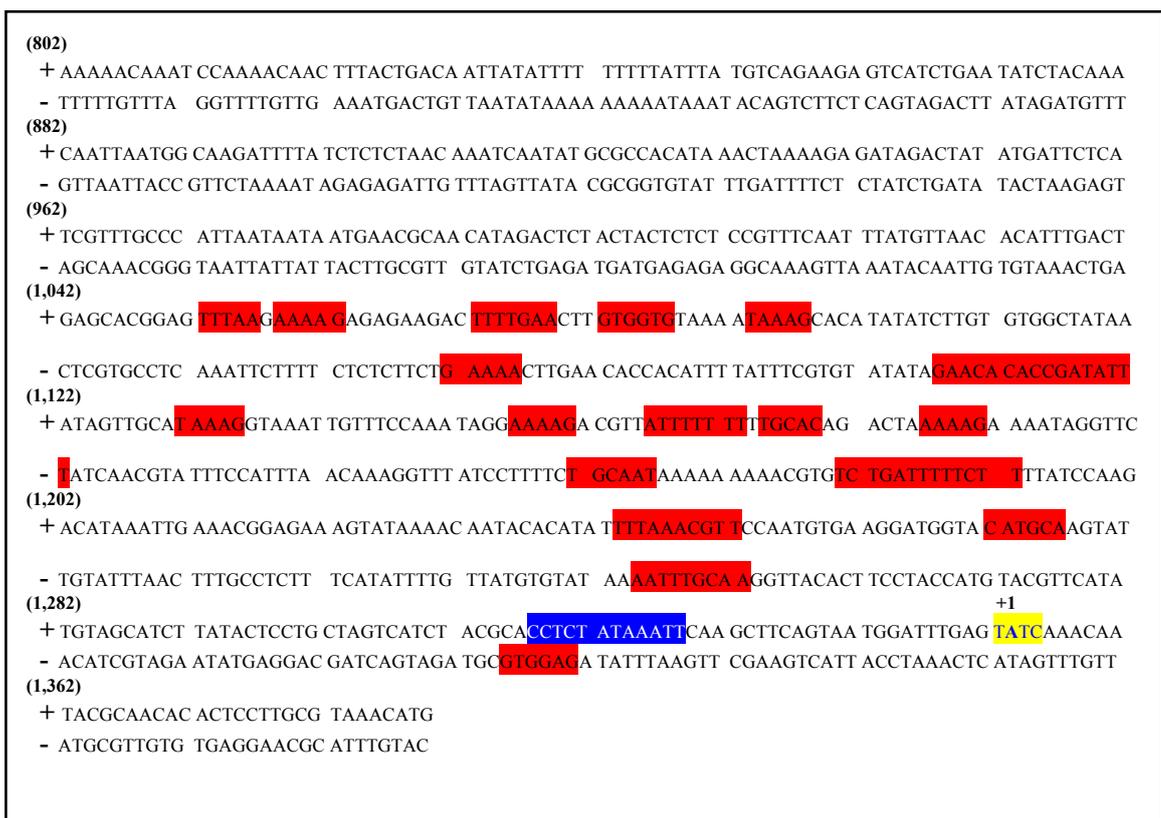
### **6.2.2 Análisis bioinformático de la región 1.4 Kb río arriba del gen *NaStEP*, mediante el uso del servidor *Softberry***

Con el propósito de ampliar el análisis, se utilizó un segundo servidor, Softberry Inc. (<http://linux1.softberry.com>); dentro del cual se empleó el programa TSSP/Prediction of PLANT Promoters (Solovyev *et al.*, 2001), que hace uso de la base de datos *Registe Plant DB* (Shahmuradov *et al.*, 2003). Así, al analizarse la secuencia del promotor potencial de *NaStEP*, se identificó la presencia de motivos correspondientes a distintos elementos reguladores en *cis*, los cuales se muestran en la tabla 10. La primera columna de esta tabla, indica la secuencia del motivo, la segunda el número de identificación del elemento regulador en la base de datos REGSITE DB: *Plant Transcription Regulatory Sites (AC)*, la tercera la posición dentro de la secuencia, la cuarta el número de veces que está presente, la quinta el nombre del elemento regulador (RE), la sexta al nombre de los factores de transcripción (BF), la séptima el gen al cual el elemento co-regula, y la última el nombre del organismo al cual pertenece.

Los distintos elementos reguladores en *cis* descritos en la tabla 10, se muestran en la figura 19 de acuerdo a su ubicación dentro de la secuencia río arriba del gen *NaStEP*. Además, se indica la posición de la caja TATA, y el elemento Inr, donde se ubica el sitio de inicio de la transcripción (+1), ó TSS, por sus siglas en inglés (*Transcription start site*).

**Tabla 10. Elementos reguladores en *cis* estimados por el programa  
TSSP de Softberry.**

SECUENCIA	AC	POSICIÓN	# SITIOS	RE	BF	GEN	Organismo
caTGCAC	RSP00016	1173 (+)	(1)	RY	U.N.F	<i>β</i> -conglycinin	<i>Glycine max</i>
AACGTT	RSP00076	1247 (+) 1252 (-)	(2)	T-box	bZIP factor	<i>rifa-7-P-glucuronidasa transgen</i>	<i>Oryza sativa</i>
WAAAG	RSP00161	1058 (+) 1093 (+) 1131 (+) 1156 (+) 1186 (+) 1075 (-)	(6)	Dof1 BSopt	Dof1	Oligonucléotidos sintéticos	<i>Zea mays</i>
ACACccGagctaaccacaac	RSP00162	1113 (-)	(1)	E1	DPBF-1, DPBF-2	<i>Dc3</i>	<i>Daucus carota</i>
CATGCA	RSP00327	1271 (+)	(1)	RY	ABI3	<i>napA</i>	<i>Brassica napus</i>
TTTGAA	RSP00398	1073 (+)	(1)	CHR	U.N.F	<i>RNP2</i>	<i>Nicotiana tabacum</i>
TAACGT	RSP00401	1166 (-)	(1)	Myba'	Myb	<i>RNP2</i>	<i>Nicotiana tabacum</i>
TTAGTCT	RSP00402	1186 (-)	(1)	AP1	AP1-like leucine zipper factor	<i>RNP2</i>	<i>Nicotiana tabacum</i>
GNGGTG	RSP00469	1082 (+) 1320 (-)	(2)	Alfin1 BS1o	Alfin1	Oligonucléotidos sintéticos	<i>Medicago sativa</i>
GTGGNG	RSP00470	1082 (+)	(1)	Alfin1 BS2	Alfin1	Oligonucléotidos sintéticos	<i>Medicago sativa</i>
TTTAA	RSP00477	1052 (+) 1243 (+) 1248 (-)	(3)	S-box	U.N.F	<i>rbcSF1</i>	<i>Hordeum vulgare,</i> <i>Brassica napus</i>
gcaTTTTTatca	RSP00508	1164 (+) 1165 (+) 1166 (+) 1167 (+) 1168 (+) 1192 (-)	(6)	SEF4 BS	SEF4	<i>β</i> -conglycinin $\alpha'$	<i>Glycine max</i>
tcctACACgcGtcacaattc	RSP00518	1118 (-)	(1)	E3	DPBF-1, DPBF-2	<i>Dc3</i>	<i>Daucus carota</i>
caattcaggACACgtGccctctca	RSP00519	1122 (-)	(1)	E4	DPBF-1, DPBF-2	<i>Dc3</i>	<i>Daucus carota</i>
ACACccG	RSP00521	1113 (-)	(1)	E1- core	DPBF-1, DPBF-2	<i>Dc3</i>	<i>Daucus carota</i>
ACACgcG	RSP00523	1113 (-)	(1)	E3- core	DPBF-1, DPBF-2	<i>Dc3</i>	<i>Daucus carota</i>
ACACgtG	RSP00524	1113 (-)	(1)	E4- core	DPBF-1, DPBF-2	<i>Dc3</i>	<i>Daucus carota</i>
GCCACA	RSP00596	1116 (-)	(1)	box II EE2	unknown nuclear factor	<i>rbcS3A</i>	



**Fig. 19. Posición de los distintos elementos reguladores en cis, dentro de la secuencia 1.4 Kb río arriba de la región codificadora del gen *NaStEP*, estimados por el programa TSSP de Softberry. En rojo se indican los distintos elementos reguladores en cis. En azul se muestra la caja TATA, y en amarillo al elemento Iniciador *Inr*, donde se ubica el sitio de inicio de la transcripción (+1), ó TSS, por sus siglas en inglés (*Transcription start site*).**



### 6.3.2 Elementos reguladores en *cis* involucrados en la expresión específica del estigma y el estilo identificados en la secuencia del promotor de *NaStEP*.

Dzelzkalns *et al.* (1993) además encontraron, que al comparar a las secuencias promotoras de los genes *SLG<sub>13</sub>*, *SLG<sub>2</sub>*, *SLG<sub>8</sub>* y *SLR<sub>1</sub>* se presentaban 5 elementos altamente conservados, lo cual sugería que estos podrían actuar como elementos reguladores en *cis*, implicados en la expresión específica de los diferentes tejidos del pistilo y la antera. Esto fue lo que nos motivó a buscar a estas secuencias en la región promotora del gen *NaStEP* y en la de los genes *SRK<sub>2</sub>* y *SRK<sub>6</sub>* (Figura 21); resultando finalmente en la identificación exitosa de estos elementos en las secuencias promotoras de estos genes.

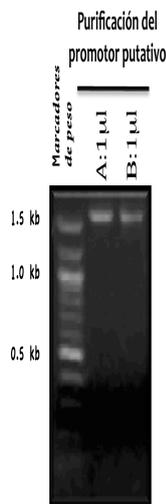
GEN	CAJA I	CAJA II	CAJA III	CAJA IV	CAJA V
	GACNAATGATA	GTTTGT	TGANTTAATCG	TGAAAAAGTCATNGA	ATTTNCTTGCTGCT
<i>SLG<sub>13</sub></i>	GACNAATGATA GACAAATGATA	GTTTGT GTTTGT	TGANTTAATCG TGAATTAATCG	TGAAAAAGTCATNGA TGAAAAAGTCATGGA	ATTTNCTTGCTGCT ATTTTACTTGCTGCT
<i>SLR<sub>1</sub></i>	GACNAATGATA GACCAATGATA	GTTTGT GTTTGT	TGANTTAATCG TGAGTTAATCG	TGAAAAAGTCATNGA TGAAAGTCATAGA	ATTTNCTTGCTGCT ATTTTCTTGCTGCT
<i>SRK<sub>2</sub></i>	GACNAATGATA GAAAGATGATA	GTTTGT GTTTGT	TGANTTAATCG TGAATTAATCG	TGAAAAAGTCATNGA TGAAAAAGTCATGGA	ATTTNCTTGCTGCT ATTTTGCTTGCTGCT
<i>SRK<sub>6</sub></i>	GACNAATGATA GACAAATGATA	GTTTGT GTTTGT	TGANTTAATCG TAGGTCATCG	TGAAAAAGTCATNGA TGAAAGTCATGGA	ATTTNCTTGCTGCT ATTTTGCTTGCTGCT
<i>NaStEP</i>	GACNAATGATA GCAAAATGATA	GTTTGT GTTTGT	TGANTTAATCG TAAATTAATCG	TGAAAAAGTCATNGA CAAAAAAGTCATAGG	GTTTAAATTTAATGTT ATTTNCTTGCTGCT ATTTCTCATCGTTGCT

**Fig. 21. Elementos reguladores en *cis* involucrados en la expresión específica del estigma y el estilo identificados en la secuencia del promotor de *NaStEP*.** En las secuencias de las regiones reguladoras del los genes *SRK<sub>2</sub>* y *SRK<sub>6</sub>*, se observa la presencia de las cinco distintas cajas con ligeras variaciones en sus secuencias. La presencia de las Cajas IV y V es más difícil de asegurar en la secuencia de la región promotora del gen *NaStEP*; sin embargo es posible suponer el hallazgo de las Cajas I, II y III en esta misma secuencia promotora, debido una alta similitud entre secuencias. EL alineamiento entre secuencias se realizó usando el CLUSTAL W. La secuencias resaltadas representan a las **secuencias consenso** de cada una de las cajas. Con una línea debajo, se indica a los nucleótidos donde se observa homología. Mientras que los **nucleótidos no conservados** aparecen indicados en azul.

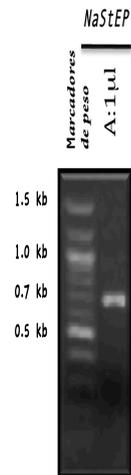
## 6.4. *NaStEP* FUSIONADO A SU PROMOTOR POTENCIAL: CONSTRUCCIÓN Y CLONACIÓN.

### 6.4.1 Construcción del segmento 3-2, fusionado al gen *NaStEP*.

Tanto la secuencia de cDNA del gen *NaStEP* como el fragmento 3-2 se amplificaron mediante PCR. Con la finalidad de fusionar ambos segmentos, el producto de ambas reacciones se purificó a partir de gel de agarosa. En la figura 22 se observa un banda a una altura aproximadamente de 1.5 Kb, la es una alícuota del volumen total de resuspensión del fragmento 3-2 purificado. En la figura 23 se observa una banda a una altura aproximada de 700 pb, la cual correspondió a una alícuota del volumen total de resuspensión del gen *NaStEP* purificado.

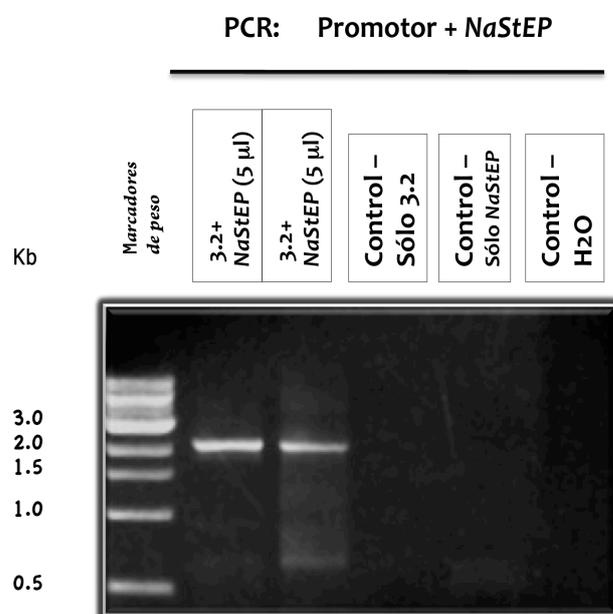


**Fig 22. Fragmento 3.2.** Electroforesis del fragmento 3.2 purificado a partir de gel de agarosa. Se cargó 1 µl de un volumen final de resuspensión de 10 µl, a un gel de agarosa al 1% teñido previamente con bromuro de etidio. A, B: Muestras del total del producto purificado. Marcadores de peso de 1kb.



**Fig 23. *NaStEP*.** Electroforesis del cDNA del gen *NaStEP* purificado a partir de gel de agarosa. Se cargó 1 µl de un volumen final de resuspensión de 10µ, a un gel de agarosa al 1% teñido previamente con bromuro de etidio. A: Muestra del total del producto purificado Marcadores de peso de 1kb.

Posteriormente a la purificación de los productos de PCR, se llevó a cabo otra reacción (PCR) de tipo fusión para unir al fragmento 3-2 cuya secuencia correspondió a aquella ubicada río arriba del gen *NaStEP*, y la cual posee una alta probabilidad de contener a la región promotora de éste gen, y la secuencia de cDNA del gen *NaStEP*. Debemos recordar que el extremo 3' de la secuencia del segmento 3.2 empalma con 140 pb del extremo 5' de la secuencia de *NaStEP*, por lo que el empalme entre el extremo 5' de una secuencia con el extremo 3' de la otra, proporcionó la direccionalidad deseada a la construcción; orientando a la secuencia promotora putativa al extremo 5' y a la secuencia del gen *NaStEP* al extremo 3'. En la figura 24 se muestra al producto de la reacción de fusión, el cual posee una longitud aproximada de 2.1 Kb que corresponde al tamaño esperado de la construcción.



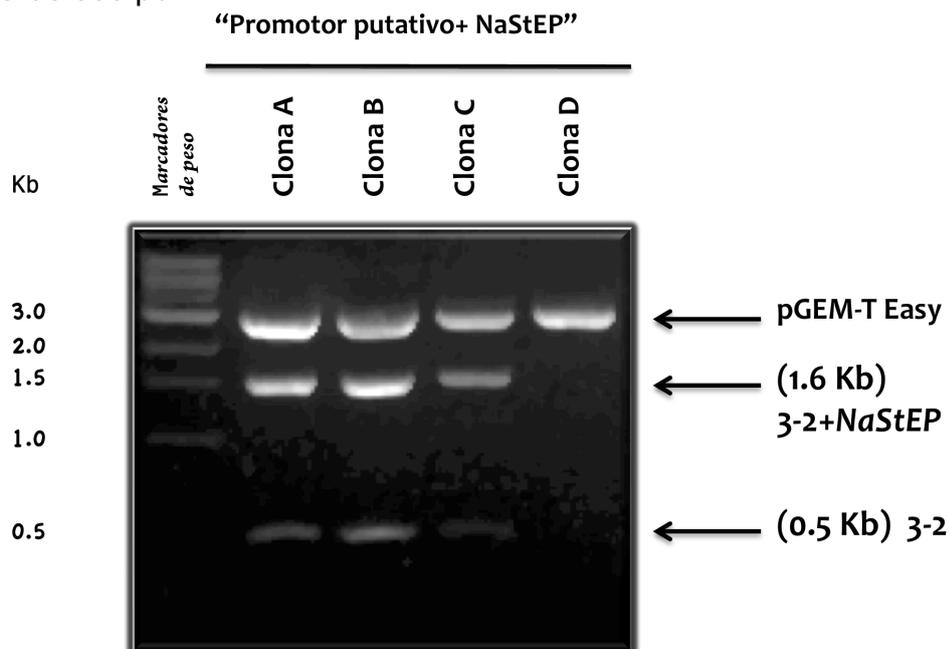
**Fig 24. PCR de fusión entre el gen *NaStEP* y la secuencia de su promotor potencial.**

Fraccionamiento electroforético de los productos de PCR utilizando como templado al ADN cuya secuencia corresponde a la del ADNc del gen *NaStEP* y al de la región promotora putativa de este gen. El fragmento amplificado cuenta con un tamaño aprox. de 2.1 Kb. El tamaño del fragmento amplificado corresponde a la suma hipotética del gen *NaStEP* más su región promotora putativa. La separación se realizó en un gel de agarosa al 1.5%, teñido con bromuro de etidio. Control (-) Sólo 3.2: Reacción de PCR utilizando únicamente como templado el fragmento 3.2, Control (-) Sólo *NaStEP*: Reacción de PCR utilizando únicamente como templado al gen *NaStEP*, Control (-) H<sub>2</sub>O: PCR sin ADN, Marcador de peso de 1 Kb.

#### 6.4.2 Clonación del segmento 3-2 fusionado al gen *NaStEP*

La construcción en la cual se fusionó al gen *NaStEP* con a su promotor potencial se clonó posteriormente en el vector pGEM T-Easy. Con la finalidad de comprobar la inserción de la construcción en el vector de clonación, se realizó un análisis de restricción con la enzima *EcoRI* del plásmido purificado de cuatro de las colonias transformantes.

En la figura 25 , se observa la presencia de tres bandas. La banda que se localiza a una altura aproximada de 3.0 Kb se piensa corresponde al vector linearizado, ya que el vector pGEM T- Easy posee un longitud de 3015 pb. La construcción del gen *NaStEP* fusionado a su promotor potencial, posee una longitud total aproximada de 2.1 Kb, sin embargo debido a que la región promotora del gen *NaStEP* , ó fragmento 3.2, cuenta con un sitio interno de corte para la enzima *EcoRI* a aprox. 500 pb del inicio de la secuencia en el extremo 5' , no fue posible observar la liberación de un fragmento a la altura de 2.1 Kb. Por el contrario, se observó la presencia de dos bandas, la de mayor tamaño con un peso aproximado de 1.6 Kb, y la segunda a una altura aproximada de 500 pb.



**Fig 25. Clonación del gen *NaStEP* fusionado a su promotor potencial.** Análisis de restricción del gen *NaStEP* fusionado a su promotor potencial clonado en el vector pGEM-T-Easy, mediante la reacción de digestión con la enzima *EcoRI*. La banda ubicada aprox. a 3Kb corresponde al vector pGEM-T-Easy linearizado; mientras que las bandas ubicadas a 1.5 Kb y 500 pb. aprox. corresponden a la digestión del gen *NaStEP* fusionado a su promotor putativo. Electroforesis en gel de agarosa al 1.5% teñido previamente con bromuro de etidio. Clonas A, B y C : Colonias transformantes con la presencia del inserto en el vector pGEM-T-Easy. Clona D: Colonia transformante sin la presencia del inserto en el vector. Marcadores de peso de 1kb.

**6.5. GENERACIÓN DE UN VECTOR BINARIO DE EXPRESIÓN PARA EVALUAR LA FUNCIÓN COMO UNA REGIÓN PROMOTORA, DE LA SECUENCIA 1.4 Kb RÍO ARRIBA DE LA REGIÓN CODIFICADORA DEL GEN *NaStEP***

**6.5.1 Generación de la clona de entrada, pENTR/D-TOPO® : 3.2**

Mediante los oligos FPET32 y R32BAMHI se amplificó a la secuencia promotora potencial del gen *NaStEP* mediante una reacción de PCR. El producto de la reacción de PCR se purificó a partir de un gel de agarosa, y posteriormente se efectuó la reacción de clonación en el vector TOPO.

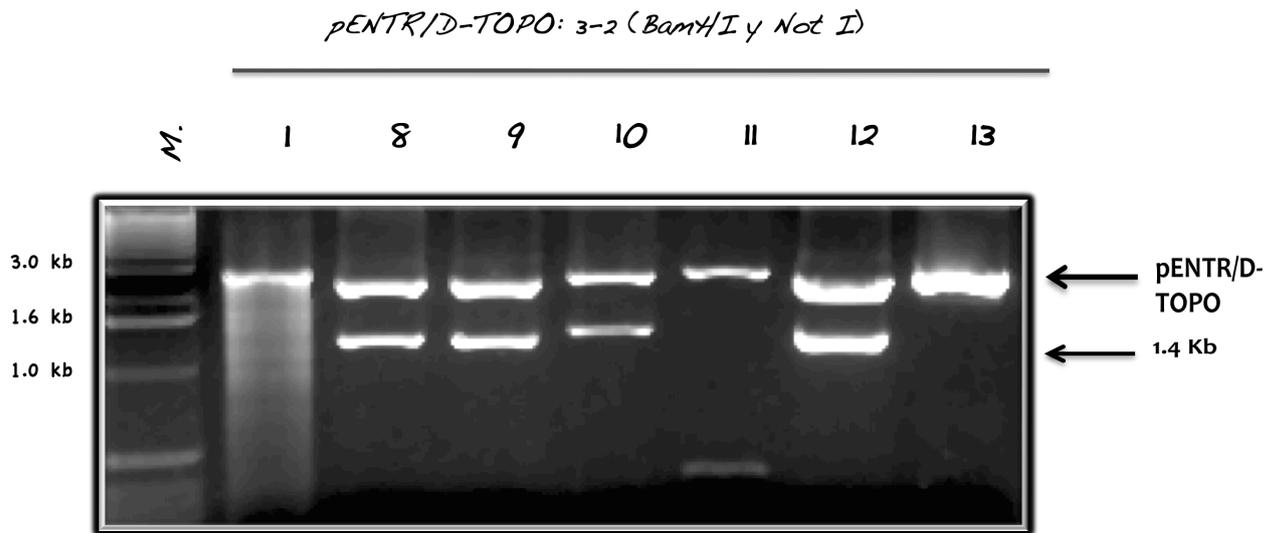
Con la finalidad de confirmar la presencia del segmento 3.2, el cual posee una longitud de aprox. 1.4 Kb, en el vector pENTR/D-TOPO®, se realizó un análisis de restricción del plásmido purificado de 25 colonias transformantes. Para ello, se efectuó una doble digestión empleando a las enzimas *Not I* y *BamHI*. La secuencia de reconocimiento de la enzima *Not I*, estuvo presente en el vector cercana al sitio de recombinación. La secuencia de reconocimiento de la enzima *BamHI* estuvo presente únicamente en la secuencia del extremo 3' del fragmento 3.2, amplificado con el oligo *reverse* R32BAMHI.

En la figura 26 se muestra el resultado de la reacción de doble digestión con las enzimas *Not I* y *BamHI* del plásmido extraído de 7 de las 25 colonias transformantes; donde se observa que donde ha ocurrido la inserción del fragmento 3.2, ocurre la liberación de un segmento a una altura aproximada de 1.4 Kb.

**6.5.2 Generación de la clona de expresión, pBGWFS7:3.2**

Una vez generada la clona de entrada, pENTR/D-TOPO:3.2, se continuó con la reacción de recombinación (LR). A través de esta reacción, se logró transferir a la región promotora potencial del gen *NaStEP* ó segmento 3.2 de la construcción de entrada al vector destino, con la finalidad de generar una clona capaz de dirigir la expresión específica de los genes *GFP* y *GUS* en plantas transgénicas; lo que podría determinar más adelante la actividad

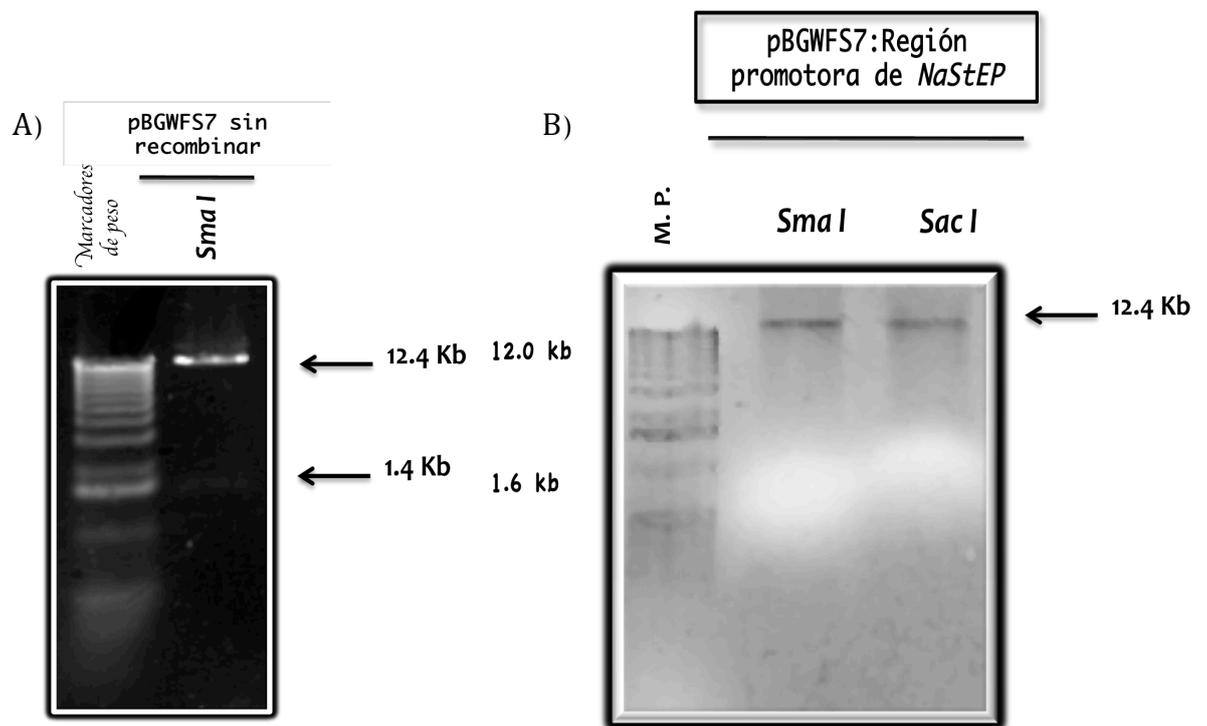
promotora *in vivo* de la región 1.4 Kb río arriba de la secuencia codificadora del gen *NaStEP*.



**Fig. 26. pENTR/D-TOPO: 3.2** Análisis de restricción y liberación del inserto 3.2 del vector pENTR/D-TOPO®. Corrimiento electroforético de la reacción de digestión con las enzimas *Not I* y *BamHI* del plásmido de siete distintas clonas. Donde se puede observar, en algunas casos, la liberación de una banda a la altura de 1.4 Kb aproximadamente; longitud que corresponde a la secuencia de ADN de la región promotora potencial del gen *NaStEP* ó fragmento 3.2. En el carril correspondiente a la clona señalada con el número *11*, se observa la liberación de un fragmento, sin embargo este no corresponde a la longitud esperada de 1.4 Kb. *1*: Control negativo. *8, 9, 10 y 12*: Clonas todas positivas para el fragmento 3.2 clonado en la orientación correcta en el vector de entrada pENTR/D-TOPO. *13*: Clona transformante con plásmido vacío. Electroforesis en gel de agarosa al 1.5% teñido previamente con bromuro de etidio. Marcadores de peso de 1kb.

Una vez seleccionada la clona de expresión, mediante su crecimiento en distintos medios de selección, y aislado el plásmido, se procedió a realizar un análisis de restricción empleando las enzimas *SmaI* y *SacI*. El vector destino pBGWFS7 posee un único sitio de reconocimiento para la enzima *SacI* y dos sitios internos de reconocimiento para la enzima *SmaI*, los cuales se ubican a 1503 pb y 2,900 pb. En la figura 27 (parte A) se observa que de una digestión del vector con la enzima *SmaI* se libera un fragmento de aproximadamente 1.4 Kb.

Una vez que el vector pBGWFS7 recombinó, el sitio de reconocimiento de la enzima *SmaI* a 2,900 pb se perdió, y debido a que el fragmento 3.2 no poseía sitios internos para *SmaI*, el producto esperado de la digestión del plásmido recombinante con esta enzima era una banda a la altura aproximada de 12.4 Kb; tal como se observa en la parte B de la figura 27 donde se aprecian dos bandas a más de 12.0 Kb (probablemente 12.4 Kb). La primer banda en esta figura, fue el resultado de la digestión del plásmido de la colonia recombinante con la enzima *SmaI*, y la segunda de idéntico peso, de la segunda digestión con la enzima *SacI*. Estos datos nos confirmaron finalmente, que el plásmido de la colonia aislada pertenecía al vector pBGWFS7 recombinante, con la secuencia promotora del gen *NaStEP*.



**Fig. 27. pBGWFS7:3.2.** A) Corrimiento electroforético de la reacción de digestión con las enzima *SmaI* del plásmido pBGWFS7 sin recombinar. El cual fue propagado en la cepa *E.Coli* DB3.1. Se esperaba la liberación de un fragmento de aprox. 1.4 Kb, tal como se observa en la figura. Corrimiento en gel de agarosa al 1.0% teñido previamente con bromuro de etidio. Marcadores de peso de 12 Kb. B) Electroforesis de la reacción de digestión con las enzimas *SmaI* y *SacI* del plásmido pBGWFS7, extraído de una colonia aislada de un medio con espectinomicina. No se observa la liberación de un fragmento a la altura de 1.4 Kb en la digestión con la enzima *SmaI*, lo cual indica favorablemente que el plásmido analizado, se trata del vector pBGWFS7 recombinante. Electroforesis en gel de agarosa al 1.0% teñido previamente con bromuro de etidio. M.P. Marcadores de peso de 12 Kb.

*Discusión, Perspectivas y  
Conclusiones.*

## 7. DISCUSIÓN

### 7.1 Clonación de la región promotora del gen *NaStEP*

En la búsqueda de la secuencia promotora del gen *NaStEP*, se apostó por la clonación de la secuencia río arriba de la región codificadora, la cual contaba con una alta probabilidad de contener a la región reguladora de este gen. La metodología universal del *Genome Walker*, permitió recuperar más de un fragmento de ADN en cada una de sus cuatro bibliotecas. No obstante, sólo se decidió clonar aquéllos amplicones de mayor tamaño e intensidad, debido a que éstos poseían la mayor probabilidad de contener a la región reguladora de este gen. Los amplicones clonados y nombrados como fragmentos 1.2, 3.2 y 3.3 poseían una longitud de secuencia mayor ó igual a 1.5 Kb. La secuenciación de estos distintos fragmentos demostró que sólo la secuencia 3.2 posee en su extremo 3' identidad del 100% con la secuencia del extremo 5' del cDNA del gen *NaStEP*; por lo que, se determinó únicamente a la secuencia de ADN del fragmento 3.2 como la región río arriba de la secuencia codificadora de éste gen. Asimismo, al llevarse a cabo un alineamiento entre las secuencia restantes, 1-2 (1.5 Kb) y 3-3 (3.0 Kb), se mostró que éstas dos podían tratarse de una misma región en el genoma de *N. alata*, sugiriendo fuertemente que el oligo *reverse* se ancló en algunos casos, a una secuencia distinta a la del cDNA del gen *NaStEP*, lo que produjo finalmente la amplificación de productos de manera inespecífica.

### 7.2 Análisis de la secuencia 1.4 Kb río arriba de la región codificadora del gen *NaStEP* mediante el uso de dos distintos tipos de algoritmos bioinformáticos

El total de los motivos encontrados en la región 1.4 Kb río arriba de la región codificadora del gen *NaStEP*, mediante el uso del algoritmo PLANT CARE, pertenecen en su totalidad a un conjunto de posibles sitios de unión a factores de transcripción (FT) identificados uno del otro de manera independiente. Programas como PLANT CARE, no buscan la integración de los diferentes motivos en la estructura de un promotor, ni consideran en su

búsqueda la necesidad de una asociación potencial de éstos con otros sitios, ó verifican que su posición, por ejemplo con respecto al sitio de inicio de la transcripción, sea congruente. Por lo que el análisis de la región 1.4 Kb río arriba de la región codificadora del gen *NaStEP* a través de PLANT CARE, no permite determinar con certeza que esta región posea a un complejo activo de transcripción en *cis*.

Sin embargo, existe otro tipo de software al cual pertenece Softberry TSSP, el cual es capaz de contestar preguntas básicas cómo ¿Está secuencia contiene a un promotor? Y si es así, ¿Dónde es que se ubican los sitios de unión a FT? (Shahmuradov *et al.*, 2003). Fue así, que mediante este tipo de algoritmo se determinó a la región 1.4 Kb río arriba de la región codificadora del gen *NaStEP*, como la secuencia potencial de un promotor.

Algoritmos como Softberry TSSP, comparan el número de motivos de 3-6 pb encontradas en una secuencia, con el número de estos que se esperaría encontrar por azar de acuerdo a la composición nucleotídica de la secuencia. Aquéllas que presenten una salida significativa, esto es un exceso de motivos, se consideran pues como secuencias de promotores verdaderos. Además de la predicción de promotores verdaderos, este tipo de programas también son capaces de determinar el posible sitio de inicio de la transcripción (TSS) y los distintos sitio de unión a FT. Además, a diferencia de programas como PLANT CARE, este tipo de software toma en consideración la dimensión combinatoria de un promotor evitando así la predicción de motivos potenciales de unión a FT, los cuales no poseen una función transcripcional real (Shahmuradov *et al.*, 2003).

No obstante, debido a la gran complejidad que caracteriza a los sitios de unión a FT, este tipo de algoritmos no escapa a los falsos positivos y negativos, lo cual genera una considerable tasa de error. Por lo tanto, muchos autores no consideran siquiera a este tipo de algoritmos lo suficientemente poderosos aún para ser usados como herramientas únicas e independientes en la búsqueda de promotores (Pan *et al.*, 2009; Priest *et al.*, 2009; Rombauts *et al.*, 2003; Prestridge, 1995; Mustonen and Lassig, 2005). Así, autores como Vedel y Scotti (2011) aseguran que una mezcla de los dos

tipos de software es necesario, para confirmar la presencia de un promotor potencial y la de sitios de unión a FT.

### **7.3 Elementos del promotor primordial de *NaStEP* (caja TATA)**

La presencia de la caja TATA en la región 1.4 Kb río arriba de la secuencia codificadora del gen *NaStEP* se determinó mediante ambos algoritmos, PLANT CARE y Softberry TSSP. Sin embargo, el primero predijo a un total de 59 motivos correspondientes a la caja TATA, mientras que el segundo sólo lo hizo para un único sitio ubicado a 34 pb río arriba del sitio de inicio de la transcripción. Entre las muchas especies a las cuales pertenecen los motivos de la caja TATA predichos por PLANT CARE se encuentra *N. tabacum*, la cual posee la mayor proximidad genética con respecto a *N. alata* entre estas distintas especies. La caja TATA de un promotor de *N. tabacum* muestra en su secuencia una alta identidad con la secuencia de la caja TATA predicha por el programa de Softberry TSSP. Este resultado sugiere que verdaderamente esta secuencia ubicada a -34pb, podría tratarse de la caja TATA del promotor de *NaStEP*.

### **7.4 Elementos reguladores en *cis* del promotor de *NaStEP***

Los elementos reguladores en *cis* localizados en la región 1.4 Kb río arriba de la región codificadora del gen *NaStEP* corresponden a elementos de respuesta a la luz (L-Box, A y P-Box, Box 4, Box I, G-Box, GAG-motif, TCT-motif , 4cl-CMA2b, G-Box 6 y GATA 2T), elementos de respuesta al ácido abscísico y al ácido salicílico, elementos de respuesta al estrés por sequía, y elementos relacionados a la expresión específica del endosperma y a la regulación negativa de la expresión en el polen.

Sin embargo, sería prematuro pensar que cada uno de estos elementos reguladores en *cis* predichos por el análisis bioinformático, funciona en verdad como un sitio de unión a un factor de transcripción *in vivo*. Debido a que los motivos relativamente cortos, por ejemplo el elemento GATA, el cual aparece en promedio varias veces en la región promotora de un gen, pueden estar presentes en cada una de las regiones del genoma de una planta, como lo son las regiones reguladoras ó promotoras, las regiones intergénicas, las

regiones codificadoras y por último los intrones. Otro ejemplo es el elemento corto y palindrómico conocido como caja G (CACGTG). La caja G es uno de los elementos en *cis* mejor estudiados, el cual es capaz de dirigir la expresión genética de las plantas en respuesta a la luz (Giuliano *et al.*, 1988). Varios estudios han demostrado que la caja G está frecuentemente sobrerrepresentada en las secuencias promotoras de ciertos genes co-expresados y en las secuencias intragenómicas conservadas no codificadoras (CNSs), además de estar presente en aproximadamente 29,000 locaciones del genoma de *Arabidopsis* (Freeling *et al.*, 2007; Freeling *et al.*, 2009). Estos últimos autores, analizaron 14,944 CNSs de *Arabidopsis* y demostraron que muchos motivos de unión a FT conocidos, tales como la caja G, están sobrerrepresentados en estas regiones. Es por tanto que aún hoy continua siendo un reto poder distinguir elementos potenciales en *cis*, del ruido genómico de fondo, que sirvan como auténticos sitios de unión a FT.

Un estudio basado en las tecnologías recientes, tales como el *high-throughput sequencing* facilitaría enormemente los esfuerzos por identificar a los elementos reguladores en *cis* verdaderamente funcionales, del conjunto de elemento potenciales. La inmunoprecipitación de cromatina asociada con una secuenciación a fondo (ChIP-seq) (Mikkelsen *et al.*, 2007; Johnson *et al.*, 2007) serviría en la identificación de factores de transcripción blancos, y finalmente el mapeo de la localización de los nucleosomas en el genoma completo (Shivaswamy *et al.*, 2008), podría asociar la organización de la cromatina con la actividad transcripcional del promotor de *NaStEP*.

### **7.5 Elementos de regulación en *cis* implicados en la expresión específica de estigma y estilo del promotor del gen *SLG*, identificados en la región 1.4 Kb río arriba del gen *NaStEP*.**

En *Brassica*, el gen de la glicoproteína del Locus S, *SLG*, se expresa únicamente en los órganos reproductivos de la flor, el pistilo y las anteras. En plantas transgénicas de *Brassica* y *Nicotiana* el promotor de este gen es activo en células del estigma y en el tejido de transmisión del estilo y el

ovario, además de que en la primera, la actividad de este promotor es evidente en el tapete, un tejido esporofítico de la antera y en las microsporas (Sato *et al.*, 1991; Thorsness *et al.*, 1991).

Dzelzkalns *et al.* (1993) establecieron a los elementos mínimos de expresión en el promotor del gen *SLG<sub>13</sub>*, mediante la transformación en *N. tabacum* con versiones truncadas del promotor y fusionados al gen reportero *GUS*. De esta manera, se demostró que el promotor *SLG<sub>13</sub>* posee una organización modular que consiste de elementos separados, que de manera independiente especifican la expresión tanto en el polen como en el pistilo. El módulo mínimo necesario para conferir especificidad en la expresión del estilo y estigma del gen *SLG<sub>13</sub>*, posee una longitud de 196 pb ubicado entre los nucleótidos -339 y -143. Esta región de 196 pb, delimita la expresión en el tejido de transmisión a lo largo del estilo, en la zona de secreción del estigma y en las células papilares. Dos dominios distintos, pero funcionalmente redundantes, delimitan la expresión en el polen (-415 a -291 y -117 a +8). Además de que un elemento silenciador en *cis* actúa sobre el dominio ubicado a -117, regulando negativamente la expresión en el polen.

Durante el mismo trabajo además de la secuencia promotora del gen *SLG<sub>13</sub>*, se reportaron dos secuencias más correspondientes al promotor de distintos alelos de este gen, *SLG<sub>2</sub>* de *B. oleracea* y *SLG<sub>8</sub>* de *B. Campestris*, así como también a la región promotora del gen *SLR<sub>1</sub>*. El gen *SLR<sub>1</sub>* (*S Locus Related gene*) es un miembro de la familia de los genes *S* que no está ligado al Locus *S* y que es expresado de manera similar al gen *SLG* (Lalonde *et al.*, 1989; Trick and Flavell, 1989; Hackett *et al.*, 1992). Al ser alineadas estas cuatro regiones promotoras, *SLG<sub>13</sub>*, *SLG<sub>2</sub>*, *SLG<sub>8</sub>* y *SLR<sub>1</sub>*, se observó un alto grado de identidad entre ellas, con la presencia de 5 elementos de secuencias altamente conservados. Estas 5 cajas ó elementos de secuencias con alta identidad, poseen la siguientes secuencias consenso: caja I, GACNAATGATA, caja II, GTTTGT, caja III, TGANTTAATCG, caja IV, TGAAAAAGTCATNGA y caja V, ATTTNCTTGTCTGCT.

Tres de los elementos de las 5 secuencias conservadas, caja I, II y III, se encuentran ubicadas dentro de la secuencia de 196 pb del dominio de

especificidad de estigma y estilo. Las cajas IV y V no son necesarias para la expresión en el pistilo; sin embargo, se propone a estas dos últimas cajas como moduladores de la expresión, ó bien como a elementos de actuación en *cis* cuya función pudiera estar relacionada a las de un promotor mínimo. No obstante, la caja V se encuentra presente dentro del dominio del promotor a -117, por lo que este elemento podría jugar un papel en la dirección de la expresión específica del polen (Dzelzkalns *et al.*, 1993).

El gen *NaStEP*, al igual que *SLG*, se encuentra fuertemente regulado y cuya expresión es específica del estigma, por lo que se decidió comparar a la secuencia de los dominios funcionales del promotor del gen *SLG*<sub>13</sub> con la secuencia 1.4 Kb río arriba del gen *NaStEP*, la cual es altamente probable de ser la región promotora de este gen. Sin embargo, no se observaron regiones con un alto grado de identidad entre ambas secuencias. Este resultado es congruente con la literatura, que sugiere que debido a la naturaleza modular que caracteriza y define a la secuencia de un promotor, no se presentan grandes regiones de identidad entre las secuencias de los promotores de los genes *SLG*<sub>13</sub>, *SLG*<sub>2</sub>, *SLG*<sub>8</sub> y *SLR*<sub>1</sub>, y las secuencias de los promotores reportados de otros genes que son expresados específicamente en el pistilo ó en el polen (Dzelzkalns *et al.*, 1993).

No obstante, la búsqueda de los elementos de secuencias altamente conservados ó cajas, dentro de la secuencia 1.4 Kb río arriba de la región codificadora del gen *NaStEP* fue exitosa. Además de buscarse a estas cajas dentro de la posible secuencia promotora del gen *NaStEP*, el análisis se extendió a las secuencias promotoras de los alelos *SRK*<sub>2</sub> y *SRK*<sub>6</sub>. El gen *SRK* (*S receptor kinase*) fue identificado en *B. Oleracea* por Stein *et al.* (1991), siendo el segundo gen ligado al locus *S* descubierto después del gen *SLG*. El transcripto de *SRK* es expresado principalmente en el tejido estigmático, donde sus niveles de expresión son extremadamente bajos relativos a los niveles de *SLG* (Stein *et al.*, 1991; Glavin *et al.*, 1994; Watanabe *et al.*, 1994; Delorme *et al.*, 1995). Debido a que los genes *SLG*, *SLR*, *SRK* y *NaStEP* comparten un patrón similar de expresión, el hallazgo de secuencias conservadas en las regiones reguladoras de estos genes, como la caja I, II y III, primeramente identificados en las regiones promotoras de los

genes *SLG* y *SLR*, sugiere fuertemente que estas secuencias podrían tratarse de elementos reguladores en *cis* implicados en la expresión específica del estigma y del estilo.

### **7.6 Clonación del gen *NaStEP* fusionado a su promotor potencial**

En el entendimiento de manera más integral del mecanismo bioquímico y celular a través del cual *NaStEP* participa en el rechazo del polen en *N. alata*, es preciso conocer el mecanismo mediante el cual ésta proteína se transporta desde el estigma al tubo polínico. Una manera de enfrentar el problema, es construyendo plantas transgénicas de *Nicotiana* que expresen a *NaStEP* fusionada a una proteína fluorescente, la cual nos permita monitorear *in vivo* su transporte a los tubos polínicos. Sin embargo, debido a que el patrón de expresión de este gen en el estigma de especies incompatibles es altamente específico, se requirió primeramente de la clonación del gen *NaStEP* junto a su secuencia promotora; con el propósito de llevar a cabo en el futuro un estudio *in vivo* de esta proteína bajo condiciones reales de su expresión.

### **7.7 Generación de un vector binario de expresión para evaluar a la secuencia 1.4 Kb río arriba de la región codificadora del gen *NaStEP*, como región reguladora de la expresión.**

Mediante el análisis bioinformático, se pudo establecer la presencia de elementos propios a una región promotora en la secuencia, de longitud de 1.4 Kb, río arriba de la región codificadora del gen *NaStEP*. Además, de hallarse distintas secuencias las cuales podrían tratarse de elementos reguladores en *cis* implicados en la expresión específica del estigma y del estilo.

Sin embargo, aún queda pendiente realizar el análisis funcional de ésta secuencia con la finalidad de comprobar *in vivo* su actividad como una región promotora. Por lo cual, se llevó a cabo la clonación de este fragmento de ADN en el vector binario de expresión, pBGWFS7. Lo que permitirá, que a través de la generación de plantas transgénicas de *Nicotiana*, se evalúe a la secuencia promotora potencial del gen *NaStEP*, mediante la expresión de la proteína verde fluorescente.

## 8. PERSPECTIVAS

### 8.1 Interacción del promotor de el gen *NaStEP* con sus distintos factores de transcripción (FTs).

Claramente, la descripción de la secuencia del promotor de *NaStEP* no es el punto final de la historia; un conocimiento preciso de sus interacciones con los reguladores transcripcionales, es también necesario. Un procedimiento actualmente usado en la investigación de interacciones entre proteínas y ADN es la inmunoprecipitación de cromatina (ChIP) (Johnson *et al.*, 2007); así mismo, técnicas de inmunoprecipitación altamente sensibles están siendo desarrolladas en muchos organismos tales como las levaduras (Valouev *et al.*, 2008), los mamíferos (Nal *et al.*, 2001; Buck *et al.*, 2004) y las plantas (Misson *et al.*, 2005; Schaffer *et al.*, 2000). Además, de que estudios recientes han demostrado la posibilidad de describir el mecanismo a través del cual los diferentes factores de transcripción (FTs) eligen de manera específica a sus sitios de unión (Pan *et al.*, 2009).

Se ha descubierto también, que las tasas de degradación de estos FTs y de la RNA polimerasa, por ejemplo, mediante estudios combinados entre el ciclo de transcripción de la polimerasa con el mapeo de FT en el genoma completo, son factores que actúan como pasos reguladores claves en la control de la expresión genética (Priest *et al.*, 2009).

*¿Quiénes son los FTs que se unen a las secuencias reguladoras del gen NaStEP? ¿Cómo es que estos distintos FTs reconocen y se unen a estas secuencias de manera específica? ¿Cómo es que los distintos estados celulares regulan las tasas de degradación de estos factores, modulando así finalmente la expresión de este gen?.* Son todas preguntas que aún quedan por contestar. En la medida en que podamos responder a éstas, seremos capaces de integrar el conocimiento sobre la estructura del promotor de *NaStEP* y los diferentes sitios de unión a FTs presentes en ésta, con el mecanismo final a través del cual se efectúa la regulación fina de la expresión de este gen.

## **8.2 El estudio de la regulación de la expresión genética de *NaStEP*, más allá de la regulación en *cis***

Además de la regulación en *cis*, la cual juega un papel fundamental en el control de la expresión genética, existen varios pasos más que controlan la regulación de la expresión, además de numerosos y complejos puntos de retroalimentación. Por ejemplo, poco se conoce acerca de la regulación de la accesibilidad de los genes para la transcripción causada por las modificaciones en el estado de condensación de la cromatina.

Los cambios transitorios observados en el patrón de expresión de genes que no dependen de la variación genética, pueden deberse a efectos de la epigenética. La epigenética se define como la herencia de variaciones fenotípicas causada por factores distintos a la variación en la secuencia de ADN, por ejemplo la metilación del ADN (Waddington, 1957). Las secuencias reguladoras son el blanco perfecto de estos mecanismos, en la manera en que estas controlan la expresión genética y a las redes reguladoras. Sabiendo que la actividad de los complejos de transcripción depende de las estructuras secundarias y terciarias de la doble cadena de ADN y de la cromatina, el estudio de los mecanismos finos de regulación en *cis* dependerá en gran parte de la investigación paralela de los procesos epigenéticos (Zhu y Thiele, 1996; Liu y Stein, 1997; Shapiro *et al.*, 2004). Sólo así, mediante el entendimiento integral de los efectos epigenéticos sobre la regulación en *cis*, es que se alcanzará finalmente un conocimiento amplio de cómo es que la regulación de la expresión genética, conduce a patrones de expresión altamente específicos como el del gen *NaStEP*.

## **8.3 El alcance evolutivo de las modificaciones en el control de la expresión genética.**

Estudios evolutivos recientes, aplicando las nuevas tecnologías de secuenciación y las poderosas herramientas bioinformáticas, indican claramente que la evolución está principalmente conducida por cambios en el complejo mecanismo de regulación genética y no únicamente por los polimorfismos presentes en las secuencias codificadoras de las proteínas

(Carroll, 2005; Wray, 2003; Wagner and Lynch, 2010). Además, la mayoría, sino todos los estudios de las secuencias promotoras se enfocan en los mecanismos que gobiernan sus funciones, sin atender las fuerzas evolutivas que pudieron haber determinado su estructura y función a lo largo del tiempo.

Así, el enfoque primordial de un estudio a escala fina sobre evolución genotípica y fenotípica del sistema de incompatibilidad en las plantas, deberá estar basado en un futuro sobre el estudio de las redes reguladoras en *cis* y no únicamente en la presencia ó ausencia de genes ó mutaciones en la región codificadora. Por ejemplo, se podría empezar por entender cuáles son las mutaciones a nivel de la secuencia promotora del gen *NaStEP*, que llevaron a éste, a expresarse únicamente en las especies de *Nicotiana* autoincompatibles.

## 9. CONCLUSIONES

- La secuencia de 1.4 Kb río arriba de la región codificadora del gen *NaStEP*, posee varios elementos propios a los de un promotor, lo que sugiere fuertemente que en esta zona podría ubicarse la región reguladora de este gen.
- La secuencia promotora potencial de *NaStEP*, posee elementos reguladores en *cis*, implicados en la expresión específica del estigma.

## 10. BIBLIOGRAFÍA

- Allen E, Z Xie, A M Gustafson, J C Carrington (2005)** miRNA-directed phasing during trans-acting siRNA biogenesis in plants, *Cell* 121 207–221.
- Alonso C R, A S Wilkins (2005)** The molecular elements that underlie developmental evolution, *Nat. Rev. Genet.* 6 709–715.
- Anderson M A, E C Cornish, S-L Mau, E G Williams, R Hoggart, A Atkinson, I Boning, B Grego, R Simpson, P J Roche, J D Haley, J D Penschow, H D Nial, G W Tregear, J P Coghlan, R J Crawford and A E Clarke (1986)** Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. *Nature*, 321, 38–44.
- Altschul S. F. et al. (1997)** Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.
- Blackwood E M, J T Kadonaga (1998)** Going the distance: a current view of enhancer action, *Science* 281 60–63.
- Birnboim H C y Doly J (1979)** A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7 (6): 1513-1523.
- Broothaerts W, G A Janssens, P Proost and W F Broekaert (1995)** cDNA cloning and molecular analysis of two self-incompatibility alleles from apple. *Plant Mol Biol.* 27, 499–511.
- Buck M J, J D Lieb (2004)** ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments, *Genomics* 83 349–360.
- Busot G, B McClure, C Ibarra-Sánchez, K Jiménez-Duran, S Vázquez Santana, F Cruz-García (2008)** Pollination in *Nicotiana alata* stimulates synthesis and transfer to the stigmatic surface of NaStEP, a vacuolar Kunitz proteinase inhibitor homologue. *J. Exp. Bot.* 59:3187-3201.
- Butler J E F, J T Kadonaga (2002)** The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev.* 16:2583–2592
- Carroll S B (2005)** Evolution at two levels: on genes and forms, *PLoS Biol.* 3e 245
- Cornish E C, J M Pettitt, I Böning, and A E Clarke (1987)** Developmentally controlled expression of a gene associated with self-incompatibility in *Nicotiana alata*. *Nature*, 326, 99–102.
- Cruz-García F, C N Hancock, D Kim, B McClure (2005)** Stylar glycoproteins bind to S-Rnase in vitro. *The Plant J.* 42:295-304.
- Covington M F, J N Maloof, M Straume, S A Kay, S L Harmer (2008)** Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development. *Genome Biol.* 9:R130.

- de Nettancourt, D. (1977)** Incompatibility in Angiosperms. Springer-Verlag, Berlin.
- de Nettancourt, D. (2001)** Incompatibility in Angiosperms: Monographs on Theoretical and Applied Genetics 3. Springer-Verlag, New York. 312 p.
- Dzelzkalns V A, M K Thorsness, K G Dwyer, J S Baxter, M A Balent, M E Nasrallah and J B Nasrallah (1993)** Distinct *cis*-Acting Elements Direct Pistil-Specific and Pollen-Specific Activity of the Brassica *S* Locus Glycoprotein Gene Promoter. *Plant Cell*, Vol. 5, 855-863.
- Delorme V, J-L Giranton, Y Hatzfeld, A Friry, P Heizmann, M J Ariza, C Dumas, T Gaude, J M Cock (1995)** Characterization of the *S* locus genes, *SLG* and *SRK*, of the *Brassica S<sub>3</sub>* haplotype: identification of a membrane-localized protein encoded by the *S* locus receptor kinase. *Plant J.* 7: 429-440.
- Franklin- Tong N V, y F C Franklin. (2003)** Gametophytic self-incompatibility inhibits pollen tube growth using different mechanisms. *Trends Plant Sci.* 8: 598-605.
- Freeling M, S Subramaniam (2009)** Conserved noncoding sequences (CNSs) in higher plants. *Curr Opin Plant Biol*, 12:126-132.
- Freeling M, L Rapaka, E Lyons, B Pedersen, B CThomas (2007)** G- Boxes, Bigfoot genes, and environmental response: characterization of intragenomic conserved noncoding sequences in *Arabidopsis*. *Plant Cell*.19:1441-1457.
- Filipowicz W, S N Bhattacharyya, N Sonenberg (2008)** Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.* 9 102–114.
- Fassler J S, G N Gussin (1996)** Promoters and basal transcription machinery in eubacteria and eukaryotes: concepts, definitions, and analogies, RNA polymerase and associated factors. *Methods Enzymol.* (1996) 3–29.
- Giuliano G, Pechersky E, Malik V S, Timko M P, Scolnik P A, Cashmore A R (1988)** An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. *Proc Natl Acad Sci U S A*, 85:7089-7093.
- Glavin T L, Goring D R, Schafer U, Rothstein S J (1994)** Features of the extracellular domain of the S-locus receptor kinase from *Brassica*. *Molecular and General Genetics* 244: 630-637.
- Goldraij A, K Kondo, C B Lee, C N Hancock, M Sivaguru, S Vazquez-Santana, S Kim, T E Phillips, F Cruz-García, B McClure (2006)** Compartmentalization of S-RNase and HT-B degradation in self-incompatible *Nicotiana*. *Nature* 439:805-810.
- Gray, J E, McClure B A, Bönig I, Anderson M A, and Clarke A E (1991)** Action of the style product of the self-incompatibility gene of *Nicotiana glauca* (S-RNase) on in vitro-grown pollen tubes. *Plant Cell*, 3, 271–283.
- Hackett R M, Lawrence M J, and Franklin F C H (1992)** A *Brassica S*-locus related gene promoter directs expression in both pollen and pistil of tobacco. *Plant J.* 2, 613-617.
- Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999)** Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999, *Nucleic Acids Res.* 27 297–300.
- Huang S, Lee H S, Karunanandaa B, and Kao T-h. (1994)** Ribonuclease activity of *Petunia inflata S* proteins is essential for rejection of self-pollen. *Plant Cell*, 6, 1021–1028.

- Hobert O (2008)** Gene regulation by transcription factors and microRNAs, *Science* 319 1785–1786.
- Ikeda K, I Boris, K Ushijima, H Yamame, N R Haunck, R Nakano, H Sassa, A F Iezzoni, J R Kohn, R Tao (2004)** Primary structural features of the S-haplotype-specific F box proteína, *SFB*, in *Prunus*. *Sexual Plant Reprod.* 16:235-243.
- Jiménez-Durán K, F Cruz-García (2011)** Sexual incompatibility, a genetic mechanism that prevents self-fertilization and contributes to plant diversity. *Rev. Fitotec. Mex.* Vol. 34 (1): 1-9.
- Johnson D S, A Mortazavi, R M Myers, B Wold (2007)** Genome-wide mapping of *in vivo* protein–DNA interactions. *Science*, 316:1497-1502.
- Juárez-Díaz J, B Mc Clure, S Vázquez-Santana, A Guevara-García, P León-Mejía , J Márquez-Guzmán, F Cruz-García (2006)** A novel Thioredoxin h is secreted in *Nicotiana glauca* and reduces S-RNase *in vitro*. *J. Biol. Chem.* 281:3418-3424.
- Kao T-h. and Tsukamoto T (2004)** The molecular and genetic bases of SRNase-based self-incompatibility. *Plant Cell*, 16, 72–83.
- Lind J L, A Bacic, A E Clarke, M A Anderson (1994)** A style-specific hydroxyproline-rich glycoprotein with properties of both extensins and arabinogalactan proteins. *Plant J.* 6: 491-502.
- Lalonde B A, Nasrallah M E, Dwyer K G, Chen C.-H, Barlow B, and Nasrallah J B (1989)** A highly conserved *Brassica* gene with homology to the S- locus-specific glycoprotein structural gene. *Plant Cell* 1, 294-258.
- Liu K, Stein A (1997)** DNA sequence encodes information for nucleosome array formation. *J. Mol. Biol.* 270 (1997) 559–573.
- <sup>1</sup>Michael T P, T C Mockler, G Breton, C McEntee, A Byer, J D Trout, S P Hazen, R Shen, H D Priest, C M Sullivan (2008)** Network discovery pipeline elucidates conserved time of day specific *cis*-regulatory modules. *PLoS Genet*, 4:e14.
- <sup>2</sup>Michael T P, G Breton, S P Hazen, H D Priest, T C Mockler, S A Kay, J Chory (2008)** A morning specific phytohormone gene expression program underlying rhythmic plant growth. *PLoS Biol*, 6:e225.
- Mikkelsen TS, M Ku, D B Jaffe, B Issac, E Lieberman, G Giannoukos, P Alvarez, W Brockman, T K Kim, R P Koche (2007)** Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, 448:553-560.
- Misson J, K G Raghothama, A Jain, J Jouhet, M A Block, R Bligny, P Ortet, A Creff, S Somerville, N Rolland, P Dumas, P Nacry, L Herrera-Estrella, L Nussaume, M.-C. Thibaud (2005)** A genome-wide transcriptional analysis using *Arabidopsis thaliana* affymetrix gene chips determined plant responses to phosphate deprivation, *Proc. Natl. Acad. Sci. U.S.A.* 102 11934–11939.
- Mustonen V, Lassig M (2005)** Evolutionary population genetics of promoters: predicting binding sites and functional phylogenies, *Proc. Natl. Acad. Sci. U.S.A.* 102 15936–15941.
- Nal B, Mohr E, Ferrier P, (2001)** Location analysis of DNA-bound proteins at the whole-genome level: untangling transcriptional regulatory networks, *Bioessays* 23 473–476.

- McClure B (2006)** New views of S-RNase-based self-incompatibility. *Curr. Op. Plant Biol.* 9:639-646.
- Mc Clure B (2005)** The stylar 120 kDa glycoprotein is required for S-specific pollen rejection in *Nicotiana*. *Plant J.* 43:716-723.
- Mc Clure, B. and Franklin-Tong V, (2006)** Gametophytic self-incompatibility: understanding the cellular mechanism involved in “self” pollen tube inhibition. *Planta*, 224, 233–45.
- McClure B (2004)** S-RNase and SLF determine S-haplotype-specific pollen recognition and rejection. *Plant Cell*, 16, 2840–2847.
- McClure B, F Cruz-García, B Beecher, W Sulaman (2000)** Factors affecting inter and intra specific pollen rejection in *Nicotiana*. *Ann. Bot. Suppl.* 85A: 113-123.
- McClure B, B Mou, S Canevascini, R Bernatzky (1999)** A small asparagine-rich protein required for S-allele-specific pollen rejection in *Nicotiana*. *Proc. Nat. Acad. Sci.* 96:13548-13553.
- McClure B, H Du, Y H Lui, and A E Clarke (1993)** S-locus products in *Nicotiana alata* pistils are subject to organ-specific post-transcriptional processing but not post-translational processing. *Plant Mol. Biol.*, 22, 177–187.
- McClure B, J E Gray, M A Anderson and A E Clarke (1990)** Self-incompatibility in *Nicotiana alata* involves degradation of pollen rRNA. *Nature*, 347, 757–760.
- McClure B, V Haring, P R Ebert, M A Anderson, R J Simpson, F Sakiyama y A Clarke (1989)** Style self-incompatibility gene products of *Nicotiana alata* are ribonucleases. *Nature* 342: 955-957.
- McCubbin A G and Kao T H (1996)** Molecular mechanisms of self-incompatibility. *Curr. Op. Biotechnol.* 7, 150–154.
- O’Brien M, G Major, S Chanta, D Matton (2004)** Isolation of S-RNase protein from *Solanum chacoense*: identification of an SBP1 (Ring finger protein) ortholog. *Sexual Plant Reprod.* 17:81-87.
- Ohler U, Liao G, Niemann H, Rubin G M (2002)** Computational analysis of core promoters in the *Drosophila* genome. *Genome Biol.* 3:0087– 87.12
- Pan Y P , Tsai C J, Ma B Y, Nussinov R (2009)** How do transcription factors select specific binding sites in the genome? *Nat. Struct. Biol.* 16 1118–1120.
- Prestridge D S (1995)** Predicting Pol-II promoter sequences using transcription factor-binding sites, *J. Mol. Biol.* 249 923–932.
- Priest H D, Filichkin S A, Mockler T C (2009)** cis-Regulatory elements in plant cell signaling, *Curr. Opin. Plant Biol.* 12 (2009) 643–649.
- Rhoades M W, Reinhart B J, Lim L P, Burge C B, Bartel B, Bartel D P (2002)** Prediction of plant miRNA targets, *Cell* 110 513–520.
- Rombauts S, Dehais P, Van Montagu M, Rouze P (1999)** *PlantCARE*, a plant *cis*-acting regulatory element database. *Nucleic Acids Res*, 27:295-296.

- Sato, T., Thorsness, M.K., Kandasamy, M.K., Nishio, T., Hiral, M., Nasrallah, J.B., and Nasrallah, M.E. (1991)** Activity of an *S* locus gene promoter in pistils and anthers of transgenic Brassica. *Plant Cell* 3, 867-876.
- Sassa H, Nishio T, Kowyama Y, Hirano H, Koba T and Ikehashi H (1996)** Self-incompatibility (*S*) alleles of the Rosaceae encode members of a distinct class of the T2/S-ribonuclease superfamily. *Mol Gen Genet.*, 250, 547–557.
- Schaffer R, J Landgraf, M Pérez-Amador, E Wisman (2000)** Monitoring genome-wide expression in plants, *Curr. Opin. Biotechnol.* 11 162–167.
- Shahmuradov I A, A J Gammerman, J M Hancock, P M Bramley, V V Solovyev (2003)** PlantProm: a database of plant promoter sequences. *Nucleic Acids Res.*, 31, 114-117.
- Shapiro M D, M E Marks, C L Peichel, B K Blackman, K S Nereng, B Jonsson, D Schluter, D M Kingsley (2004)** Genetic and developmental basis of evolutionary pelvic reduction in three spine sticklebacks. *Nature* 428 717–723.
- Shivaswamy S, Bhinge A, Zhao Y, Jones S, Hirst M, Iyer VR (2008)** Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol*, 6:e65.
- Sijacic P, X Wang, A L Skirpa, Y Wang, P E Dowd, A G McCubbin, S Huang, T-h Kao (2004)** Identification of the pollen determinant of *S*-RNase-mediated self-incompatibility. *Nature* 429:302-305.
- Smale S T, Kadonaga J T (2003)** The RNA polymerase II Core Promoter. *Annu. Rev. Biochem.* 2003. 72:449–79.
- Smale S T (2001)** Core promoters: active contributors to combinatorial gene regulation. *Genes Dev.* 15:2503– 2508
- Smale S T (1994)** *Transcription: Mechanisms and Regulation*. Ed. RC Conaway, JW Conaway, pp. 63–81. New York: Raven.
- Solovyev V V (2001)** Statistical approaches in Eukaryotic gene prediction. In *Handbook of Statistical genetics* (eds. Balding D. et al.), John Wiley & Sons, Ltd., p. 83-127.
- Solovyev V V (2002)** Structure, Properties and Computer Identification of Eukaryotic genes. In *Bioinformatics from Genomes to Drugs. V.1. Basic Technologies*. (ed. Lengauer T.), p. 59 – 111
- Stein J C, B Howlett, D C Boyes, M E Nasrallah, J B Nasrallah (1991)** *Proc. Natl. Acad. Sci. U.S.A.* 88, 8816
- Sullivan J A, K Shirasu, X W Deng (2003)** The diverse roles of ubiquitin and the 26S proteasome in the life of plants. *Nature Rev. Genet.* 4: 948-958.
- Thompson J D, Gibson T J, Plewniak F, Jeanmougin F, and Higgins D G (1997)** The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, 24, 4876–4882.
- Thompson J D, Higgins D G, Gibson T J (1994)** CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, 22, 4673–4680.

- Thorsness M K, Kandasamy M K, Nasrallah M E, and Nasrallah J B (1991)** A *Brassica S*-locus gene promoter targets toxic gene expression and cell death to the pistil and pollen of transgenic *Nicotiana*. *Dev. Biol.* 143, 173-184.
- Trick M, and Flavell R B. (1989)** A homozygous *S* genotype of *Brassica oleracea* expresses two *S*-like genes. *Mol. Gen. Genet.* 218, 112-117.
- Valouev A, D S Johnson, A Sundquist, C Medina, E Anton, S Batzoglou, R M Myers, A Sidow (2008)** Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data, *Nature* 5 829–834.
- Vedel V, Scotti I (2011)** Promoting the promoter. *Plant Science* 180: 182–189.
- Waddington C H (1957)** *The Strategy of the Genes*, Allen & Umwin, London.
- Wagner G P, V.J. Lynch (2010)** Evolutionary novelties, *Curr. Biol.* 20 48–52.
- Watanabe M, Takasaki T, Toriyama K, Yamakawa S, Isogai A, Suzuki A, Hinata K (1994)** A high degree of homology exists between the protein encoded by *SLG* and the *S* receptor domain encoded by *SRK* in self-incompatible *Brassica campestris* L. *Plant and Cell Physiology* 35: 1221-1229.
- Watanabe Masao, Suzuki Go, Takayama Seiji, Isogai Akira and Hinata Kokichi (2000)** Genomic organization of the *SLG/SRK* region of the *S* Locus in *Brassica* species. *Annals of Botany* 85 (Supplement A): 155-160, 2000.
- B. Wasylyk (1988)** Transcription elements and factors of RNA polymerase B promoters of higher eukaryotes, *CRC Crit. Rev. Biochem.* 23 77–119.
- Wray G A (2003)** Transcriptional regulation and the evolution of development, *Int. J. Dev. Biol.* 47 675–684.
- Wagner G P, Lynch V J (2010)** Evolutionary novelties, *Curr. Biol.* 20 48–52.
- Wasserman W W, A Sandelin (2004)** Applied bioinformatics for the identification of regulatory elements. *Nature Reviews Genetics* 5, 276-287.
- Woodward, J. R., A. Bacic, W. Jahnen y A. E. Clarke (1989)** N-Linked Glycan Chains on S-Allele-Associated Glycoproteins from *Nicotiana glauca*. *Plant Cell* 1: 511-514.
- Xue Y, Carpenter R, Dickinson H G, and Coen E S (1996)** Origin of allelic diversity in *Antirrhinum S* locus RNases. *Plant Cell*, 8, 805–814.
- Yamame H, R Tao (2009)** Molecular basis of self-(in)compatibility and current status of *S*-genotyping in Rosaceous fruit trees. *J. Japa. Soc. Hort. Sci.* 78:137-157.
- Zdepski A, W Wang, H D Priest, F Ali, M Alam, T C Mockler, T P Michael (2008)** Daily transcriptional programs in papaya. *Trop Plant Biol*,1:236-245.
- Zhang W, J Ruan, T H Ho, Y You, T Yu (2005)** Quatrano RS: cis- Regulatory element based targeted gene finding: genome- wide identification of abscisic acid- and abiotic stress-responsive genes in *Arabidopsis thaliana*. *Bioinformatics*, 21:3074-3081.
- Zhao Y, Srivastava D (2007)** A developmental view of microRNA function, *Trends Biochem. Sci.* 32 189–197.
- Zhu Z W, D J Thiele (1996)** A specialized nucleosome modulates transcription factor access to a *C-glabrata* metal responsive promoter. *Cell* 87, 459–470.

## Apéndice

**Tabla 1.** Preparación del medio de cultivo LB.

Reactivo	LB líquido	LB sólido
Extracto de levadura	5.0 g	5.0 g
Peptona/Triptona/Caseína hidrolizada <sup>1</sup>	10.0 g	10.0 g
NaCl	5g	5g
H2O	X ml	X ml
Agar	-	15 g
TOTAL *	1000 ml	1000 ml

\*Autoclavear el medio por 20 min a 1.2 kg/cm<sup>2</sup>

<sup>1</sup>Cualquiera de estos reactivos puede ser usado como fuente de nitrógeno

**Tabla 2.** Preparación de los antibióticos.

Antibióticos*	Solución stock	Concentración de trabajo	Disolvente
Ampicilina	100 mg/ml	100-300 mg/l	H2O
Espectinomicina	100 mg/ml	100 mg/l	H2O
Cloranfenicol	50 mg/ml	25 mg/l	Etanol
Kanamicina	100 mg/ml	90 mg/l	H2O

\*Una vez preparados, los antibióticos son esterilizados por filtración (0.22 µm) y adicionados a los medios de cultivo a temperaturas por debajo de los 45o oC.

**Tabla 3.** Preparación de los suplementos de los medios de cultivo.

Suplemento	Concentración de trabajo	Disolvente
XGAL	80 mg/ml	N-N dimetil formamida
IPTG	0.5 mM	H2O