

## Aislamiento y clonación parcial de un gen de terpeno sintasa de orégano mexicano (*Lippia graveolens* Kunth)

### Isolation and partial cloning of a terpene synthase gene from Mexican oregano (*Lippia graveolens* Kunth)

Helani Yazmin Ortiz Macias, René Isaías Loeza Román, Osvaldo Adrián Castellanos Hernández, Araceli Rodríguez Sahagún, Gustavo Javier Acevedo Hernández

Laboratorio de Biología Molecular Vegetal, Centro Universitario de la Ciénega, Universidad de Guadalajara. Av. Universidad 1115, Col. Linda Vista, Ocotlán 47820, Jalisco, México

#### Resumen

El orégano mexicano (*Lippia graveolens*) es una planta aromática que es ampliamente utilizada como condimento, pero que además posee diversas aplicaciones tanto tradicionales como industriales. Las propiedades que hacen tan atractiva a esta especie se deben principalmente a la composición de su aceite esencial, en el cual los monoterpenos timol y carvacrol son los constituyentes mayoritarios. La síntesis de los terpenos involucra la participación de enzimas denominadas genéricamente como terpeno sintasas, las cuales han sido ampliamente estudiadas en un gran número de especies vegetales debido a la gran diversidad de compuestos que generan, dando lugar a la mayor familia de productos naturales conocidos. Debido a la carencia de información sobre los genes involucrados en la síntesis de terpenos en *L. graveolens*, el propósito de este estudio fue la obtención y análisis de un gen de terpeno sintasa de esta especie, por lo que se procedió a la amplificación por PCR de un fragmento de cDNA empleando primers degenerados diseñados a partir de regiones conservadas de genes de este tipo caracterizados en otras plantas aromáticas. El fragmento obtenido fue clonado, y el análisis de su secuencia muestra una alta similitud a nivel de nucleótidos con genes de terpeno sintasa, indicando que corresponde a la clona parcial de un gen potencialmente involucrado en la síntesis de terpenos en esta especie.

#### Palabras clave

orégano, terpenos, timol, carvacrol

**Autor de correspondencia:** Gustavo Javier Acevedo Hernández, e-mail: gustavo.acevedo@cuci.udg.mx

#### Abstract

Mexican oregano (*Lippia graveolens*) is an aromatic plant extensively used as condiment, that also has many traditional and industrial uses. The properties that make this plant so attractive are mainly due to the composition of its essential oil, with monoterpenes thymol and carvacrol being the major constituents. The enzymes involved in the biosynthesis of terpenes are generically known as terpene synthases, and these enzymes have been widely studied in many plant species due to the diversity of compounds they are able to produce, leading to the largest group of known natural products. In view of the lack of information regarding the genes involved in the biosynthesis of terpenes in *L. graveolens*, the purpose of this study was to obtain and analyze the sequence of a terpene synthase gene from this species. With this aim, a cDNA fragment was amplified through PCR using degenerate primers designed from conserved regions of related genes characterized in other aromatic plants. The amplified fragment was cloned, and the nucleotide sequence analysis reveals a high similarity to terpene synthase genes, suggesting it corresponds to the partial clone of a gene potentially involved in the synthesis of terpenes in this species.

#### Keywords

oregano, terpenes, thymol, carvacrol

#### Introduction

Oregano is the generic name given to more than 40 aromatic plant species belonging to at least four botanical families that are grown in many regions worldwide. These plants are used mainly as

## Introducción

Orégano es el nombre genérico con el que se conoce a más de 40 especies de plantas aromáticas pertenecientes a por lo menos cuatro familias botánicas y que son cultivadas en varias regiones del mundo donde son utilizadas principalmente como condimento, aunque también son ampliamente reconocidas por sus propiedades medicinales. Todas estas plantas presentan un olor característico y son ricas en monoterpenos fenólicos, principalmente en los isómeros timol y carvacrol. Las especies de orégano de mayor importancia económica debido a su distribución geográfica, volumen de producción y propiedades aromáticas, son el orégano europeo (*Origanum vulgare*) perteneciente a la familia Lamiaceae y el orégano mexicano (*Lippia graveolens*), de la familia Verbenaceae (García-Pérez et al. 2012). En México, *L. graveolens* se distribuye en sitios áridos y semiáridos y se produce de manera silvestre en 24 estados del país, localizándose las poblaciones con fuerte aroma principalmente en los estados de Jalisco, Zacatecas, Durango, Coahuila y San Luis Potosí (Sánchez et al. 2007; García-Pérez et al. 2012; Granados-Sánchez et al. 2013).

Además de tener un uso muy extendido como condimento, el orégano mexicano ha sido utilizado tradicionalmente como antiinflamatorio, analgésico, antipirético, y se ha empleado en el tratamiento de la diarrea, disentería, desórdenes menstruales y enfermedades respiratorias, entre otras dolencias (Pascual et al. 2001). Debido a su uso tradicional para el tratamiento de enfermedades infecciosas, se han realizado numerosos estudios para analizar las propiedades antimicrobianas del aceite esencial de esta planta, y se ha encontrado que posee una importante actividad contra bacterias Gram positivas y Gram negativas, así como contra algunos hongos y virus (Salgueiro et al. 2003; Hernández et al. 2009; Pilau et al. 2011).

Se ha reportado que los valores de rendimiento del aceite esencial obtenido de orégano mexicano, así como la abundancia relativa de los metabolitos que lo componen, dependen de diversos factores climáticos como la altitud, precipitación pluvial, la época de recolección e incluso la calidad espectral de la luz. Esta amplia variabilidad en la composición del aceite esencial ha llevado a la clasificación de distintos quimiotipos de orégano con base en el metabolito predominante, generalmente timol o carvacrol (Bueno-Durán et al. 2014; Calvo-Irabién et al. 2014). Es probable que algunas de estas

condiment, but their medicinal properties are also widely recognized. They have a characteristic aroma and are particularly rich in monoterpenoid phenols, especially thymol and its structural isomer carvacrol. Based on their geographic distribution, production volume and aromatic properties, the most economically important oregano species are European oregano (*Origanum vulgare*), which belongs to the family Lamiaceae, and Mexican oregano (*Lippia graveolens*), a member of the family Verbenaceae (García-Pérez et al. 2012). In Mexico, *L. graveolens* is distributed in arid and semi-arid regions, and is obtained through exploitation of wild populations in 24 states. Populations with a strong aroma are generally found in the states of Jalisco, Zacatecas, Durango, Coahuila and San Luis Potosí (Sánchez et al. 2007; García-Pérez et al. 2012; Granados-Sánchez et al. 2013).

Besides its extensive use as a condiment, Mexican oregano has been traditionally used as an anti-inflammatory, analgesic or antipyretic remedy, and has also been used to treat diarrhea, dysentery, menstrual disorders, respiratory diseases and other ailments (Pascual et al. 2001). By cause of its traditional use for the treatment of infectious diseases, a number of studies have been conducted to investigate the antimicrobial properties of oregano essential oil, and it has been demonstrated that it possesses significant activity against Gram-positive and Gram-negative bacteria, as well as some fungi and viruses (Salgueiro et al. 2003; Hernández et al. 2009; Pilau et al. 2011).

It has been reported that the yield and the relative abundance of the components of the essential oil of Mexican oregano depend on multiple climatic factors such as altitude, precipitation, time of harvesting, and even the spectral quality of light. This great variability in the composition of the essential oil has led to the classification of chemotypes based on the dominant metabolite, usually thymol or carvacrol (Bueno-Durán et al. 2014; Calvo-Irabién et al. 2014). This compositional variation is likely to be due at least in part to the effect of environmental factors on the expression of genes encoding enzymes involved in the synthesis of terpenes (Majdi et al. 2017). Thus, obtaining the sequence of the corresponding genes in *L. graveolens* is a key step in further studying their regulation. Several terpene synthases (TPSs) are responsible for the synthesis of the great diversity of terpene molecules found in plants, using five-carbon (5C) units as precursors to produce isoprene (5C),

variaciones en la composición se deban al efecto de los factores ambientales sobre la expresión de los genes que codifican para las enzimas involucradas en la síntesis de terpenos (Majdi et al. 2017), por lo que contar con la secuencia de los genes correspondientes en *L. graveolens*, sería un paso importante para poder analizar su regulación. Varias enzimas terpeno sintasas (TPSs) son responsables de sintetizar la gran diversidad de moléculas de terpenos que se encuentran en las plantas, empleando para ello precursores de 5 átomos de carbono (5C) a partir de los cuales se generan isoprenos (5C), monoterpenos (10C), sesquiterpenos (15C) o diterpenos (20C). En *O. Vulgare* y en *Thymus caespititius* se ha propuesto que la síntesis de los monoterpenos timol y carvacrol se realiza a partir del precursor geranil difosfato (GPP), una conversión catalizada inicialmente por una TPS, seguida por un proceso de hidroxilación catalizado por monooxigenasas P450 (Figura 1) (Crocoll et al. 2010; Lima et al. 2013).

monoterpenos (10C), sesquiterpenos (15C) or diterpenos (20C). It has been proposed that the synthesis of the monoterpenes thymol and carvacrol in *O. vulgare* and *Thymus caespititius* occurs through the initial conversion of the precursor geranyl diphosphate, a reaction catalyzed by a TPS, followed by a hydroxylation process catalyzed by cytochrome P450 monooxygenases (Figure 1) (Crocoll et al. 2010; Lima et al. 2013). The study of the wide diversity of TPSs in plants has demonstrated their rapid evolution with such a divergence that in *Arabidopsis*, for instance, several sesquiterpene synthases are more similar to diterpene synthases from the same species than they are to sesquiterpene synthases found in other plant species. It is also frequent to observe cases of convergent evolution of these enzymes, where species belonging to different taxonomic families are able to produce the same compounds, a phenomenon that can be observed in the plants grouped under the denomination of oregano. These

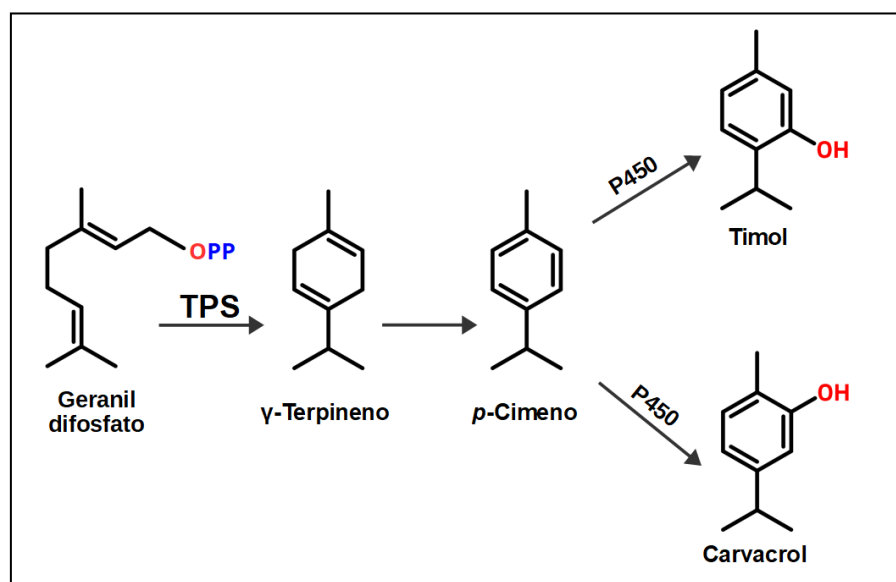


Figura 1. Vía biosintética de los terpenos timol y carvacrol. A partir del precursor de geranil difosfato (10C) ocurren procesos de ciclación, oxidación e hidroxilación para la formación de ambos isómeros. TPS: terpeno sintasa, P450: citocromo P450 monooxigenasa.

Figure 1. Biosynthesis pathway of terpenes thymol and carvacrol. Starting from precursor geranyl diphosphate (10C), cyclization, oxidation and hydroxylation processes occur to form both isomers. TPS: terpene synthase, P450: cytochrome P450 monooxygenase.

El estudio de la amplia diversidad de TPSs en plantas ha mostrado que poseen una rápida evolución con una divergencia tal que por ejemplo, en *Arabidopsis*, varias sesquiterpeno sintasas son

analyses also show that besides the presence of a short aspartate-rich motif in the enzyme active site (motif DDXXD), which is involved in binding of  $Mg^{2+}$  ions, there is no clear sequence conservation

más parecidas a diterpeno sintasas de la misma especie que a otras sesquiterpeno sintasas encontradas en otras especies vegetales. También es frecuente observar casos de evolución convergente de estas enzimas, donde especies pertenecientes a distintas familias taxonómicas son capaces de producir los mismos compuestos, como observamos en las plantas agrupadas en la denominación de orégano. Estos análisis también muestran que más allá de la presencia en el sitio activo de un motivo corto rico en aspartato relacionado con la unión de iones  $Mg^{2+}$  (motivo DDXXD), no hay una clara conservación de secuencias entre las distintas TPSs, y las similitudes que se observan entre ellas están dominadas por la relación entre especies más que por la especificidad del producto o sustrato (Tholl 2006; Crocoll et al. 2010; Chen et al. 2011; Lima et al. 2013). Esta gran variación en las secuencias de TPSs de distintas plantas supone un reto para su clonación a partir de una especie como *L. graveolens*, para la que no hay hasta el momento ninguna secuencia de TPS reportada. Con esto en mente, la finalidad de este trabajo fue lograr la amplificación y clonación de una secuencia correspondiente a una TPS de orégano mexicano empleando primers degenerados diseñados a partir de regiones conservadas de genes TPS de distintas plantas aromáticas.

## **Materiales y métodos**

### Material vegetal, extracción de RNA total y síntesis de cDNA.

Se utilizaron plantas de *L. graveolens* crecidas en invernadero, las cuales fueron originalmente germinadas a partir de semillas maduras colectadas de una población silvestre ubicada en el municipio de Colotlán, localizado al norte del estado de Jalisco. La colecta fue realizada durante el mes de mayo de 2015, época en la que se observaba la presencia de frutos secos. A partir de estas plantas se colectaron hojas frescas sin daño aparente, y 100 mg de tejido fueron molidos con nitrógeno líquido y utilizados para el aislamiento de RNA total empleando el reactivo TRIzol® (Thermo Fisher Scientific, Waltham, EUA), siguiendo el protocolo sugerido por el fabricante, y posteriormente la preparación fue tratada con el gDNA Removal Kit (Jena Bioscience, Jena, Alemania) para la eliminación del DNA genómico que pudiera estar presente. La calidad del RNA obtenido fue evaluada observando su integridad por electroforesis en un gel de agarosa al 1.2% desnaturizante con formal-

among different TPSs, and that the observed similarities are more dominated by the taxonomic relatedness of species than the substrate or product specificity (Tholl 2006; Crocoll et al. 2010; Chen et al. 2011; Lima et al. 2013). The great variation in the sequence of TPSs in different plant species poses a challenge for the cloning of a TPS gene from *L. graveolens*, a species for which there are not TPS sequences reported to date. Hence, the aim of this work was to amplify and clone the sequence corresponding to a TPS from Mexican oregano, using degenerate primers designed from conserved regions of TPS genes from different aromatic plants.

## **Materials and methods**

### Plant material, RNA extraction and cDNA synthesis.

*L. graveolens* seedlings were obtained by germination of seeds collected from a wild population located in the municipality of Colotlán in northern Jalisco. Collection of the seeds was performed during May 2015, when the presence of dry fruits could be observed. Plants were allowed to grow under greenhouse conditions and fresh leaves with no apparent damage were collected. A sample of 100 mg tissue was grinded in liquid nitrogen and used for the isolation of total RNA with TRIzol® reagent (Thermo Fisher Scientific, Waltham, USA), following the protocol suggested by the provider. The RNA preparation was then treated with the gDNA Removal Kit (Jena Bioscience, Jena, Germany) to eliminate contaminating genomic DNA. The quality of isolated RNA was assessed by visualization of integrity after denaturing 1.2% agarose/formaldehyde gel electrophoresis and quantified using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, USA). cDNA was synthesized using 500 ng of total RNA using the Superscript III First-Strand Reverse Transcription Kit (Thermo Fisher Scientific, Waltham, USA) adding oligo (dT)<sub>20</sub> as the primer.

### Sequences and design of degenerate primers.

Nucleotide sequence of the entire coding region of *Origanum vulgare* and *Thymus caespitosus* TPSs, for which their role in the synthesis of thymol has been demonstrated, as well as the sequence of monoterpene synthases from other aromatic plants, were obtained from the database GenBank. The sequences used and their GenBank ID are: *Origanum majorana* TPS (KC432651), *Origanum*

dehído, y fue cuantificado por absorción UV utilizando un espectrofotómetro NanoDrop 2000 (Thermo Fisher Scientific, Waltham, EUA). Se utilizaron 500 ng del RNA total para la síntesis del cDNA utilizando el Superscript III First-Strand Reverse Transcription Kit (Thermo Fisher Scientific, Waltham, EUA), para lo que se empleó como primer el oligo (dT)<sub>20</sub>.

#### Secuencias y diseño de primers degenerados.

La secuencia de nucleótidos de la región codificante de TPSs para las que ya se ha demostrado su papel en la síntesis de timol en *Origanum vulgare* (Crocoll et al. 2010) y *Thymus caespititius* (Lima et al. 2013), así como la secuencia de monoterpene sintasas de otras plantas aromáticas, fueron obtenidas a partir de la base de datos GenBank. Las secuencias utilizadas y su identificador de GenBank son: *Origanum majorana* TPS (KC432651), *Origanum vulgare* TPS2 (GU385977), *Salvia officinalis* (+)-bornil difosfato sintasa (AF051900), *Salvia officinalis* sabineno sintasa (AF051901), *Salvia rosmarinus* pineno sintasa (EF495245), *Thymus caespititius* TPS2 (KC181099), *Thymus serpyllum* TPS (KF494251) y *Thymus vulgaris* TPS2 (JX997981). El alineamiento de las secuencias fue realizado en el programa CLC Sequence Viewer v. 7.6 (Quiagen Bioinformatics, Redwood City, EUA), haciendo ajustes manuales en caso de ser necesario. A partir del alineamiento, se identificaron aquellas regiones con mayor grado de similitud entre los distintos genes, y se seleccionaron aquellas en las posiciones más extremas de manera que al diseñarse los primers pudiera obtenerse un fragmento del mayor tamaño posible. A partir de estas regiones se diseñaron dos primers degenerados en los que se incorporaron las variaciones observadas entre las distintas secuencias: TPS\_FW (5'-MARTTGGAGTTGATTGA-3') y TPS\_RV (5'-CTTTVARAAAYRTCRTANGCCAKYTC-3'), este último de mayor longitud, ya que tiene un mayor grado de degeneración (Figura 2).

#### Amplificación por PCR, clonación y secuenciación.

La primera cadena de cDNA fue amplificada por PCR utilizando los primers degenerados, realizándose las reacciones de amplificación en un volumen final de 20 µL conteniendo 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 2 mM MgCl<sub>2</sub>, 0.2 mM mezcla de dNTPs, 1 µM de cada primer, 2 unidades de Taq DNA polimerasa (Jena Bioscience, Jena, Alemania) y adicionando 2 µL de la reacción de

*vulgare* TPS2 (GU385977), *Salvia officinalis* (+)-bornyl diphosphate synthase (AF051900), *Salvia officinalis* sabinene synthase (AF051901), *Salvia rosmarinus* pinene synthase (EF495245), *Thymus caespititius* TPS2 (KC181099), *Thymus serpyllum* TPS (KF494251) and *Thymus vulgaris* TPS2 (JX997981). Sequence alignment was performed using the CLC Sequence Viewer software v. 7.6 (Quiagen Bioinformatics, Redwood City, USA), with manual adjustments when needed. From this alignment, regions with the highest similarity among genes were identified, and those closer to the edges were selected for primer design in order to obtain the longest fragment possible. Degenerate primers were designed from the selected regions, incorporating into the primers the observed variations among sequences: TPS\_FW (5'-MARTTGGAGTTGATTGA-3') y TPS\_RV (5'-CTTTVARAAAYRTCRTANGCCAKYTC-3'), the latter being longer due to its higher degeneracy (Figure 2).

#### PCR amplification, cloning and sequencing.

First-strand cDNA was amplified by PCR using degenerate primers. Amplification reactions were performed in a final volume of 20 µL containing 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 2 mM MgCl<sub>2</sub>, 0.2 mM dNTPs mix, 1 µM each primer, 2 units of Taq DNA polymerase (Jena Bioscience, Jena, Germany) and adding 2 µL of the cDNA synthesis reaction. Amplification was performed in a Multigene Thermocycler (Labnet International, Edison, USA) using the program: 1 step at 94°C for 5 min, 35 cycles consisting of 30 s at 94°C, 30 s at 50°C, and 1 min at 72°C, and a final extension step of 5 min at 72°C. Amplification products were analyzed by electrophoresis in 0.8% agarose gels to verify the presence of a single fragment, and the PCR product was cloned using the TA Cloning® Kit containing plasmid pCR®2.1 (Thermo Fisher Scientific, Waltham, USA), which allows the direct cloning of PCR fragments obtained with the enzyme Taq DNA polymerase. Insert-containing clones were recovered and plasmid DNA was extracted according to the method reported by Zhou et al. (1990). Purified plasmid was sent to the Laboratory of Genomic Services at Langebio-Cinvestav (Irapuato, Mexico) for sequencing of the insert. The obtained sequence was compared against the nucleotide collection (nr) of GenBank database using BLAST (Basic Local Alignment Search Tool, <https://blast.ncbi.nlm.nih.gov/Blast.cgi>) with the

síntesis del cDNA.

La amplificación se llevó a cabo en un termociclador Multigene (Labnet International, Edison, EUA) empleando el programa: 1 paso a 94°C por 5 min, 35 ciclos que consistieron de 30 s a 94°C, 30 s a 50°C, y 1 min a 72°C, con un paso de extensión final de 5 min a 72°C. Los productos de la amplificación fueron analizados por electroforesis en un gel de agarosa al 0.8% para verificar la presencia de un solo fragmento, y se procedió a su clonación empleando el TA Cloning® Kit conteniendo el plásmido pCR®2.1 (Thermo Fisher Scientific, Waltham, EUA), el cual permite la clonación directa de productos de PCR obtenidos con la enzima Taq DNA polimerasa. Las clonas que contenían inserto fueron recuperadas y se realizó extracción de DNA plasmídico de acuerdo con el protocolo reportado por Zhou et al. (1990). El plásmido purificado fue enviado al Laboratorio de Servicios Genómicos del Langenbio-Cinvestav (Irapuato, México) para la secuenciación del inserto.

program blastn, and after analysis was submitted to the public database GenBank.

## Results and discussion

The cDNA obtained from *L. graveolens* was successfully amplified using the degenerate primers designed to anneal to the conserved regions in different sequences of TPS genes. As a result, the presence of a single PCR product of about 900 bp was observed, the fragment was cloned directly into the plasmid pCR®2.1 and sequenced. Sequence analysis showed the cloning of a fragment of 880 bp, and translation of one of the reading frames displays a continuous succession of amino acids in which the presence of the conserved motif DDXXD, a distinctive feature of TPSs, can be observed.

This strongly suggests that the cloned sequence corresponds to a TPS gene from *L. graveolens* (Figure 3). The obtained sequence was deposited into the GenBank database (ID: MK214683).

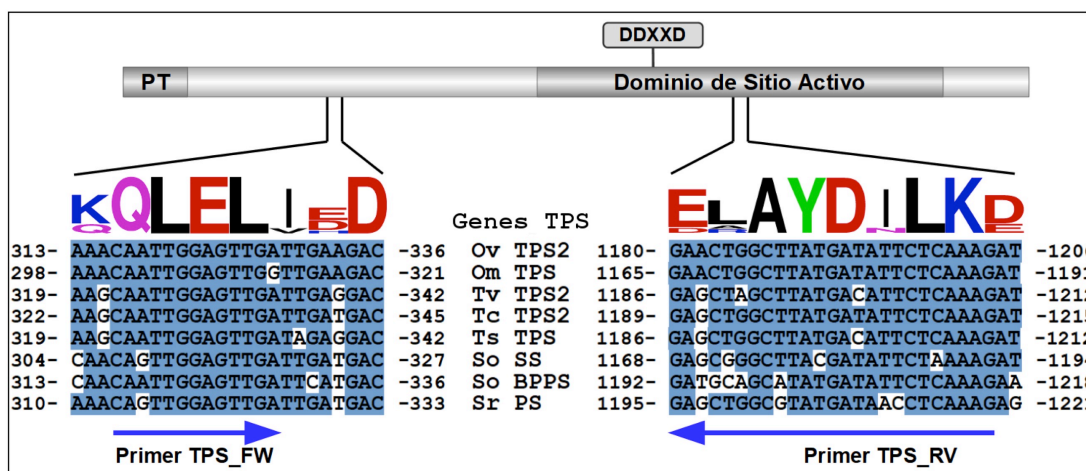


Figura 2. Localización de los primers diseñados a partir de regiones conservadas en genes de terpeno sintasa de distintas especies de plantas aromáticas. Se muestran las secuencias de nucleótidos y la conservación de aminoácidos en ambas regiones. PT: péptido tránsito, DDXXD: motivo rico en aspartato involucrado en la actividad catalítica. Los genes y especies de plantas están indicados en Materiales y Métodos.

Figure 2. Localization of primers designed from the conserved regions of terpene synthase genes from different aromatic plant species. Conservation of nucleotide and amino acid sequences in both regions is shown. PT: transit peptide, DDXXD: aspartate-rich motif involved in the catalytic activity. Genes and plant species are described in Materials and Methods.

La secuencia obtenida fue comparada contra la colección de nucleótidos (nr) de la base de datos GenBank utilizando la herramienta BLAST (Basic Local Alignment Search Tool, <https://blast.ncbi.nlm.nih.gov/Blast.cgi>) con el

Comparison of this sequence against the GenBank database using BLAST confirms its identity as part of a TPS gene, since the obtained sequence shows a high similarity to different TPSs from several aromatic plant species. Considering

programa blastn, y una vez analizada fue enviada para su incorporación a la base de datos pública de GenBank.

**Resultados y discusión**

Empleando los primers degenerados que fueron diseñados para unirse a regiones conservadas en distintas secuencias de TPSs, fue posible llevar a cabo la amplificación del cDNA obtenido de *L. graveolens*. Como resultado se observó la presencia de un solo producto de alrededor de 900 pb, el cual fue clonado directamente en el plásmido pCR®2.1 y posteriormente secuenciado. El análisis de la secuencia del inserto muestra la obtención de un fragmento de 880 pb, cuya traducción de uno de los marcos de lectura pone de manifiesto una sucesión continua de aminoácidos en las que puede observarse la presencia del motivo conservado DDXXD, una característica distintiva de las TPSs, por lo que muy probablemente esta secuencia corresponde a una región de un gen TPS de *L. graveolens* (Figura 3).

only the 20 sequences with the highest similarity to the cloned fragment, identity at the nucleotide level ranges from 69 to 76%. These 20 sequences were selected to represent the similarity as a dendrogram displaying the distance between individual sequences, a feature found in BLAST (Figure 4). It is noteworthy that all of the sequences correspond to monoterpene synthases, and the highest percent identity is obtained with sequences of TPSs from *Phyla dulcis* (a synonym with *Lippia dulcis*), even though the products of the enzymes are quite different, thus confirming the above-mentioned stronger dominance of taxonomic relatedness on the sequence similarity of TPSs, prevailing over the substrate or product specificity (Tholl 2006; Lima et al. 2013).

By inspection of the tree generated from the calculated distances between sequences, it becomes clear the vast majority are grouped according to their taxonomic relationship and not to the product formed.

<u>AAG TTG GAG TTG ATT GAT</u> GAC TTG CGG AGG ATG GGA ATA TCT TAT CAC TTT CAA GAT GAA	< 60
K L E L I D D L R R M G I S Y H F Q D E	
ATT AAT CAA ATT CTA AAT TCT GTA TAT AAT AGC AAG AAT TCA GAC AAT GAA CGA GAT TTG	< 120
I N Q I L N S V Y N S K N S D N E R D L	
TAC TCA ACA GCT CTT ACA TTC AGA CTT CTC AGA CAA CAT GGC TTT GGT ATA TCC AAA GAC	< 180
Y S T A L T F R L L R Q H F G I S K D	
GTG TTC GAT TGT TTC AAA AAT GAC AAG GGC GAA TTT AAG GCA TGT CTC GAA GAT GAT ACC	< 240
V F D C F K N D K G E F K A C L E D D T	
AGG GGA TTG TTG CAA CTT TAT GAA GCT TCA TTC TTA CTA AAA GAT GGC GAA GAA ACT TTG	< 300
R G L L Q L Y E A S F L L K D G E E T L	
GTG CAA GCA AAA GAA TTT TCC ACT AAA TTT CTT AAG AAA AAT CTC GAT GAA GGA AAA AAT	< 360
V Q A K E F S T K F L K K N L D E G K N	
ATT GAT GAG AAT CTT TCG TTA TTG GTT CGC CAT GAA TTG GAG ATG CCA ATT CAT TGG AGG	< 420
I D E N L L V R H E L E M P I H W R	
GCT CAA AGA CTA AAT GCA AGA TGG TTC ATC GAT GCA TAC GAG AAA AGA TCA GAT AGA AAT	< 480
A Q R L N A R W F I D A Y E K R S D R N	
CCA GTC GTA CTG GAG CTT GCA AAG CTG GAC TTC AAC ATT ATT CAA GCA ACG TAC CAA GAA	< 540
P V V L E L A K L S F N I I Q A T Y Q E	
GAA TTA AAA ATT GTC TCA AGT TGG TGG AAG GAA ACA TAC CTT GCT GAG ACG TTG CCA TTT	< 600
E L K I V S S W W K E T Y L A E T L P F	
GCA AGG GAT AGG TTG GTG GAA GCT TAT TAT TGG GCC GTT GGT TGG CTT AAT GAG CCT CAA	< 660
A R D R L V E A Y Y W A V G W L N E P Q	
TAT GGC TAT GCA AGA GCA TTG GTC GCC AAA TCA ATT GCC TTC ATA ACA GTA ATT GAT GAT	< 720
Y G Y A R A L V A K S I A F I T V I <b>D D</b>	
<b>ATA TAT GAT</b> GTC TAT GGT ACT TTG GAA GAG CTA TAT CAG TTT CAA GAT GCT GTT GAT AGA	< 780
<b>I Y D</b> V Y G T L E E L Y Q F Q D A V D R	
TGG GAC ATT GCA GCA ATC GAC CAT CTA CCA CAG TAC ATG CAG ATA TGT TTC TTA GCA CTC	< 840
W D I A A I D H L P Q Y M Q I C F L A L	
AAC AAC TTC GTC GAT <u>GAG CTG GCA TAC GAT ATT CTC AAA G</u>	< 880
N N F V D E L A Y D I L K X	

Figura 3. Secuencia de nucleótidos y traducción a aminoácidos del cDNA parcial de una terpeno sintasa de *Lippia graveolens*. Los sitios de unión de los primers están subrayados y el motivo DDXXD del sitio activo está resaltado.

Figure 3. Nucleotide sequence and translation of the amino acid sequence of the partial cDNA corresponding to a terpene synthase from *L. graveolens*. Primer annealing positions are underlined and the DDXXD motif is highlighted.



La secuencia fue depositada en la base de datos de GenBank con el identificador MK214683. La comparación de esta secuencia contra la base de datos de GenBank empleando BLAST confirma su identidad como parte de un gen de TPS, pues revela su alta similitud a nivel de nucleótidos con diversas TPSs de varias especies de plantas aromáticas. Tomando en cuenta las 20 secuencias que presentan mayor similitud con el fragmento clonado, se observa una identidad a nivel de nucleótidos que va del 69 al 76%. Estas 20 secuencias fueron seleccionadas para la presentación de los resultados de similitud en un dendrograma de distancias dentro de la herramienta BLAST (Figura 4).

Es interesante observar que todas estas secuencias corresponden a monoterpene sintasas, y que el mayor porcentaje de identidad se obtiene con las secuencias de TPSs de *Phyla dulcis* (cuya sinonimia es *Lippia dulcis*), aun cuando los

Although all of the sequences showing the highest percent identity correspond to monoterpene synthases, it can not be established from these similarity data that the cloned fragment also corresponds to a monoterpene synthase, it can not be established from these similarity data that the cloned fragment also corresponds to a monoterpene synthase until the sequence of the entire gene is obtained, since a distinctive feature of this group of enzymes is the presence of a transit peptide (PT) for proper targeting into the chloroplast. Furthermore, its putative role in the biosynthetic pathway of carvacrol can only be elucidated through its functional characterization, usually achieved by characterization, usually achieved by expression of the gene in a heterologous system (e.g *Escherichia coli*), addition of the precursor GPP and evaluation of the reaction products by gas chromatography coupled to mass spectrometry (GC-MS).

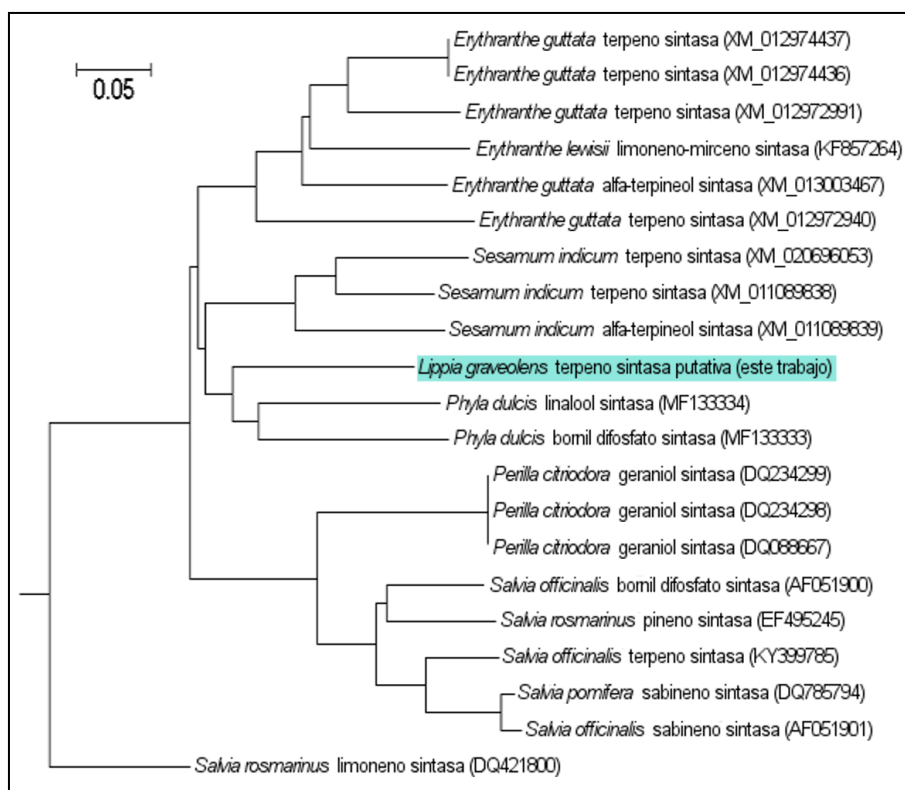


Figura 4. Dendrograma generado por la herramienta BLAST que representa la distancia calculada a partir del alineamiento entre la secuencia de nucleótidos del fragmento clonado y las secuencias en la base de datos de GenBank. Se muestran sólo las 20 secuencias con mayor porcentaje de identidad. Para cada secuencia se muestra el GenBank ID correspondiente.

Figure 4. Dendrogram generated with BLAST, displaying the distance calculated from the alignment of the cloned fragment and the sequences in the GenBank database. Only the 20 sequences with the highest percent identity are shown. The GenBank ID for each sequence is displayed.



productos de esas enzimas son muy distintos, confirmando lo que se ha mencionado previamente acerca de que la similitud entre las distintas TPSs está mayormente influenciada por la relación entre especies más que por la especificidad del producto o sustrato (Tholl 2006; Lima *et al.* 2013). Al observar el árbol generado a partir de la distancia calculada entre las secuencias, resulta evidente que la gran mayoría están agrupadas de acuerdo con su relación taxonómica y no por el producto de su catálisis. Aunque todas las secuencias con las que se observa mayor porcentaje de identidad corresponden a monoterpenos sintetas, no puede establecerse a partir de los datos de similitud que el fragmento clonado corresponde también a una monoterpeno sintasa sin contar con la secuencia completa del gen, ya que una característica distintiva de estas enzimas es la presencia de un péptido tránsito (PT) para su localización en el cloroplasto. De igual manera, su posible participación en la vía de biosíntesis de los terpenos timol y carvacrol sólo podrá dilucidarse a través de su caracterización funcional, realizada comúnmente expresando el gen en un sistema heterólogo (como *Escherichia coli*), añadiendo el precursor GPP y evaluando los productos de la reacción por cromatografía de gases acoplada a espectrometría de masas (GC-MS).

### **Conclusiones**

En resumen, en este trabajo se presenta la clonación de un fragmento de cDNA de *L. graveolens*, cuya secuencia de nucleótidos corresponde, por análisis de similitud, a un gen TPS. Por lo tanto, este se trataría del primer reporte de una secuencia de TPS en esta planta aromática. La obtención de esta secuencia es un paso necesario para el diseño de primers específicos que permitan evaluar la expresión del gen correspondiente bajo distintas condiciones de cultivo y observar si esta expresión muestra alguna correlación con la composición de terpenos en el aceite esencial. De igual manera, esto posibilita la obtención de la secuencia completa del gen a través del uso de herramientas para la amplificación de extremos del cDNA como son 5' y 3' RACE.

### **Agradecimientos**

Los autores H. Y. O. M. y R. I. L. R. agradecen al Consejo Nacional de Ciencia y Tecnología por el otorgamiento de una beca para la realización de estudios de Maestría.

### **Conclusions**

In summary, this work describes the cloning of a cDNA fragment from *L. graveolens* which nucleotide sequence corresponds, by analysis of similarity, to a TPS gene. Hence, this is the first report of a TPS sequence from this aromatic plant. Obtaining this sequence is a necessary step for designing specific primers in order to evaluate expression of this gene under different growth conditions and to establish any correlation of this expression with the terpene composition of the essential oil. In addition, this partial sequence facilitates obtaining the entire sequence of the gene through the amplification of cDNA ends using tools such as 5' and 3' RACE.

### **Acknowledgements**

Authors H.Y.O.M. and R.I.L.R. wish to thank the National Council of Science and Technology for the fellowship granted for pursuing Masters studies.

### **References**

- Bueno-Durán, A.Y., Cervantes-Martínez, J. & Obledo-Vázquez, E.N., 2014. Composition of essential oil from *Lippia graveolens*. Relationship between spectral light quality and thymol and carvacrol content. *Journal of Essential Oil Research*, 26(3), pp.153–160. Available at: <https://www.tandfonline.com/doi/abs/10.1080/10412905.2013.840808>.
- Calvo-Irabién, L.M. et al., 2014. Phytochemical diversity of the essential oils of Mexican oregano (*Lippia graveolens* Kunth) populations along an edapho-climatic gradient. *Chemistry and Biodiversity*, 11(7), pp.1010–1021. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/cbdv.201300389>.
- Chen, F. et al., 2011. The family of terpene synthases in plants: A mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *The Plant Journal*, 66(1), pp.212–229. Available at: <https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-3113X.2011.04520.x>.

- Crocoll, C. et al., 2010. Terpene synthases of oregano (*Origanum vulgare* L.) and their roles in the pathway and regulation of terpene biosynthesis. *Plant Molecular Biology*, 73(6), pp.587–603. Available at: <https://link.springer.com/article/10.1007%2Fs11103-010-9636-1>.
- García-Pérez, E. et al., 2012. Revisión de la producción, composición fitoquímica y propiedades nutraceuticas del orégano mexicano. *Revista Mexicana de Ciencias Agrícolas*, 3(2), pp.339–353. Available at: [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S2007-09342012000200010](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-09342012000200010).
- Granados-Sánchez, D. et al., 2013. Ecología, aprovechamiento y comercialización del orégano (*Lippia graveolens* H. B. K.) en Mapimí, Durango. *Revista Chapingo, Serie Ciencias Forestales y del Ambiente*, 19(2), pp.305–321. Available at: [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S2007-40182013000200011](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-40182013000200011).
- Hernández, T. et al., 2009. Composition and antibacterial activity of essential oil of *Lippia graveolens* H.B.K. (Verbenaceae). *Boletín Latinoamericano y del Caribe de Plantas Medicinales y Aromaticas*, 8(4), pp.295–300. Available at: <http://www.redalyc.org/articulo.oa?id=85611265010>.
- Lima, A.S. et al., 2013. Genomic characterization, molecular cloning and expression analysis of two terpene synthases from *Thymus caespititius* (Lamiaceae). *Planta*, 238(1), pp.191–204. Available at: <https://link.springer.com/article/10.1007/s00425-013-1884-2>.
- Majdi, M. et al., 2017. Tissue-specific gene-expression patterns of genes associated with thymol/carvacrol biosynthesis in thyme (*Thymus vulgaris* L.) and their differential changes upon treatment with abiotic elicitors. *Plant Physiology and Biochemistry*, 115, pp.152–162.
- Pascual, M.E. et al., 2001. *Lippia*: traditional uses, chemistry and pharmacology: a review. *Journal of Ethnopharmacology*, 76(3), pp.201–214. Available at: <https://www.sciencedirect.com/science/article/pii/S0378874101002343>.
- Pilau, M.R. et al., 2011. Antiviral activity of the *Lippia graveolens* (Mexican oregano) essential oil and its main compound carvacrol against human and animal viruses. *Brazilian Journal of Microbiology*, 42(4), pp.1616–1624. Available at: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1517-83822011000400049](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1517-83822011000400049).
- Salgueiro, L.R. et al., 2003. Antimicrobial activity and chemical composition of the essential oil of *Lippia graveolens* from Guatemala. *Planta Medica*, 69(1), pp.80–83. Available at: <https://www.thieme-connect.com/DOI/DOI?10.1055/s-2003-37032>.
- Sánchez, Ó. et al., 2007. Evaluación del riesgo de extinción de *Lippia graveolens* de acuerdo al numeral 5.7 de la NOM-059-SEMARNAT-2001. In *Método de evaluación de riesgo de extinción de las especies silvestres de México (MER)*. México, D.F.: Secretaría del Medio Ambiente y Recursos Naturales, Instituto de Ecología UNAM, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, pp. 91–110.
- Tholl, D., 2006. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Current Opinion in Plant Biology*, 9(3), pp.297–304. Available at: <https://www.sciencedirect.com/science/article/pii/S1369526606000537>.
- Zhou, C., Yang, Y. & Jong, A., 1990. Mini-prep in ten minutes. *Biotechniques*, 8(2), pp.172–173.