



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ciencias Biológicas

Escuela Profesional de Ciencias Biológicas

**Obtención del genoma nuclear de *Eugenia uniflora* L.
(Myrtaceae) para la identificación de terpeno sintasas**

TESIS

Para optar el Título Profesional de Biólogo con mención en
Biología Celular y Genética

AUTOR

Frank Lino GUZMAN ESCUDERO

ASESORES

Mabel ALVARADO GUTIERREZ

Rogelio MARGIS (Coasesor)

Lima, Perú

2022

Referencia bibliográfica

Guzman, F. (2022). *Obtención del genoma nuclear de Eugenia uniflora L. (Myrtaceae) para la identificación de terpeno sintasas*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Biológicas, Escuela Profesional de Ciencias Biológicas]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Frank Lino Guzman Escudero
Tipo de documento de identidad	DNI
Número de documento de identidad	40959413
URL de ORCID	https://orcid.org/0000-0002-5048-4213
Datos de asesor	
Nombres y apellidos	Mabel Alvarado Gutiérrez
Tipo de documento de identidad	DNI
Número de documento de identidad	41353553
URL de ORCID	https://orcid.org/0000-0001-8135-9223
Datos de coasesor	
Nombres y apellidos	Rogelio Margis
Tipo de documento de identidad	Cédula de identidad
Número de documento de identidad	BR / 43103839049
URL de ORCID	https://orcid.org/0000-0002-2871-4473
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Monica Arakaki Makishi
Tipo de documento	DNI
Número de documento de identidad	08136445
Miembro del jurado 1	
Nombres y apellidos	Gustavo Sandoval Peña

Tipo de documento	DNI
Número de documento de identidad	41020762
Miembro del jurado 2	
Nombres y apellidos	Claudia Barletta Carrillo
Tipo de documento	DNI
Número de documento de identidad	40846770
Datos de investigación	
Línea de investigación	A.1.1.3. Diversidad Genética
Grupo de investigación	Laboratorio de Genómica de Plantas y Poblaciones - LGPP
Agencia de financiamiento	Brasil. Coordinación de la formación del personal de nivel superior (CAPES).
Ubicación geográfica de la investigación	Edificio: Centro de Biotecnología País: Brasil Estado: Rio Grande del Sur Ciudad: Porto Alegre Barrio: Agronomia
Año o rango de años en que se realizó la investigación	Enero 2020 – enero 2022
URL de disciplinas OCDE	Bioquímica, Biología molecular https://purl.org/pe-repo/ocde/ford#1.06.03 Genética, Herencia https://purl.org/pe-repo/ocde/ford#1.06.07 Biotecnología agrícola, Biotecnología alimentaria https://purl.org/pe-repo/ocde/ford#4.04.01



Universidad Nacional Mayor de San Marcos
(Universidad del Perú, Decana de América)

FACULTAD DE CIENCIAS BIOLÓGICAS

**ACTA DE SESIÓN PARA OPTAR AL TÍTULO PROFESIONAL DE
BIÓLOGO CON MENCIÓN EN BIOLOGÍA CELULAR Y GENÉTICA
(MODALIDAD: SUSTENTACIÓN VIRTUAL DE TESIS)**

Siendo las 10:02 horas del 21 de abril de 2022, en el Salón de Grados Virtual de la Facultad de Ciencias Biológicas cuya dirección electrónica fue <https://zoom.us/j/92523453080?pwd=SXZhSlJoNEtyRDhWMmpBMFcvNGM2QT09>, y en presencia del Jurado formado por los profesores que suscriben, se inició la sesión para optar al **Título Profesional de Biólogo con mención en Biología Celular y Genética** de **FRANK LINO GUZMAN ESCUDERO**.

Luego de dar lectura y conformidad al expediente N°, el titulado expuso su tesis: **“OBTENCIÓN DEL GENOMA NUCLEAR DE *Eugenia uniflora* L. (MYRTACEAE) PARA LA IDENTIFICACIÓN DE TERPENOS SINTASAS”**, y el Jurado efectuó las preguntas del caso calificando la exposición con la nota **19**, calificativo: **Aprobado con máximos honores**.

Finalmente, el expediente será enviado a la Escuela Profesional de Ciencias Biológicas y al Consejo de Facultad para que se apruebe otorgar el **Título Profesional de Biólogo con mención en Biología Celular y Genética** a **FRANK LINO GUZMAN ESCUDERO** y se eleve lo actuado al Rectorado para conferir el respectivo título, conforme a ley.

Siendo las 11:50 horas se levantó la sesión.

Ciudad Universitaria, 21 de abril de 2022.

Dra. MONICA ARAKAKI MAKISHI
(PRESIDENTA)

Dra. MABEL ALVARADO GUTIERREZ
(ASESORA)

Mg. GUSTAVO SANDOVAL PEÑA
(MIEMBRO)

Mg. CLAUDIA BARLETTA CARRILLO
(MIEMBRO)

RESUMEN

Pitanga (*Eugenia uniflora* L.) es un miembro de la familia Myrtaceae que se distribuye en la Mata Atlántica y se caracteriza por la síntesis de metabolitos secundarios con importantes actividades biológicas, así como por tener un amplio espectro ecológico en ambientes que son climáticamente contrastantes. Sin embargo, la disponibilidad de recursos genómicos que permitan estudiar esta especie aún es limitada y escasa en comparación con otras especies de plantas. Por ese motivo, el presente trabajo de investigación tiene como objetivo la obtención de un genoma de referencia para *E. uniflora* que permita identificar los miembros de la familia de las terpeno sintasas (TPS) que son responsables de la síntesis de terpenos de esta especie. El ensamblaje *de novo* del genoma, a partir de lecturas secuenciadas utilizando la plataforma Illumina, permitió obtener un tamaño total de 385.1 Mbp con un valor de N50 igual a 26 199 bp. Asimismo, la predicción génica utilizando como referencia proteínas de Viridiplantae permitió obtener 30 663 genes codificadores de proteínas de pitanga. Un análisis de genómica comparativa con otras especies de plantas permitió identificar 2 219 genes específicos de *E. uniflora*, de los cuales el 40% fueron anotados funcionalmente con 1 772 términos ontológicos. Por otro lado, utilizando la información de la anotación de los genes predichos, se identificaron 20 genes de TPS que representan a las subfamilias TPS-b, TPS-c, TPS-f y TPS-g. Los datos generados en este trabajo permitirán comprender y conocer los mecanismos moleculares involucrados en la adaptabilidad a ambientes naturales y en la síntesis de compuestos con importancia farmacológica en pitanga.

Palabras clave: Myrtaceae, secuenciamiento de DNA, ensamblaje de DNA, predicción génica, genómica comparativa, terpeno sintasas.

ABSTRACT

Pitanga (*Eugenia uniflora* L.), a member of the Myrtaceae family, is distributed in the Atlantic Forest and is characterized by the synthesis of secondary metabolites with important biological activities, as well as by having a wide ecological spectrum in environments that are climatically contrasting. However, the availability of genomic resources to study this species is still limited and scarce compared to other plant species. For this reason, this research aims to obtain a reference genome for *E. uniflora* and identify the members of the terpene synthase (TPS) family that are responsible for the synthesis of terpenes. *De novo* assembly of the genome, from reads sequenced using the Illumina platform, allowed to obtain a total size of 385.1 Mbp with an N50 value equal to 26 199 bp. Furthermore, using Viridiplantae proteins as a reference, was possible to obtain 30,663 pitanga protein coding genes. A comparative genomics analysis with other plant species identified 2 219 genes specific to *E. uniflora*, 40% of which were functionally annotated with 1 772 ontological terms. On the other hand, using the predicted gene annotation information, 20 TPS genes representing the TPS-b, TPS-c, TPS-f, and TPS-g subfamilies were identified. The data generated in this work may allow us to understand and know the molecular mechanisms involved in the adaptability to natural environments and in the synthesis of compounds with pharmacological importance in pitanga.

Keywords: Myrtaceae, DNA sequencing, DNA assembly, gene prediction, comparative genomics, terpene synthases.