



Excelencia que trasciende

DEL VALLE
GRUPO EDUCATIVO

Análisis filogenético de *Begomovirus*
identificados en diferentes localidades y cultivos
de Guatemala

Trabajo de Investigación presentado por Edgar Roberto Rosales
Gálvez para optar al grado de Licenciado en Biología

Guatemala

2020



Excelencia que trasciende

DEL VALLE
GRUPO EDUCATIVO

Análisis filogenético de *Begomovirus*
identificados en diferentes localidades y cultivos
de Guatemala

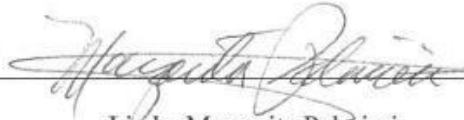
Trabajo de Investigación presentado por Edgar Roberto Rosales
Gálvez para optar al grado de Licenciado en Biología

Guatemala

2020

Vo. Bo. :

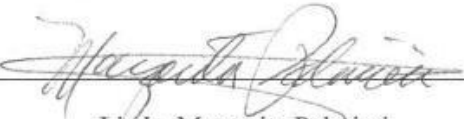
(f)



Licda. Margarita Palmieri

Tribunal Examinador:

(f)



Licda. Margarita Palmieri

(f)



Lic. Jose Miguel Morales

(f)



Licda. Ana Lucía Dubón

Fecha de aprobación:

Guatemala, 11 de noviembre de 2020

Índice

Lista de figuras.....	I
Resumen.....	II
Abstract.....	III
I. Introducción.....	1
II. Objetivos.....	2
III. Justificación.....	2
IV. Marco teórico.....	3
A. <i>Begomovirus</i>	3
B. Mosca blanca.....	6
C. Importancia económica de los cultivos.....	7
D. DNASTAR.....	8
V. Metodología.....	9
VI. Resultados.....	11
VII. Discusión de resultados.....	19
VIII. Conclusiones.....	21
IX. Recomendaciones.....	22
X. Literatura citada.....	22
XI. Anexos.....	25

Lista de Figuras

1. Diagrama de la organización genómica de los begomovirus del nuevo y viejo mundo, tanto monopartitos como bipartitos.	4
2. Diagrama de la replicación de ADN en el modo RCR	5
3. Gráfico de barras demostrando la cantidad de publicaciones utilizando DNASTAR	8
4. Captura de pantalla de un alineamiento y análisis filogenético de muestras de prueba en la aplicación MegAlignPro.....	10
5. Capturas de pantalla del análisis realizado en JModelTest para AIC, AICc y BIC respectivamente	11
6. Árbol filogenético generado utilizando el análisis de “Neighbor-joining”	12
7. Árbol filogenético de “Maximum Likelihood” generado utilizando el modelo de sustitución de nucleótidos de HKY + I (BIC).....	16

Resumen

Se realizó el análisis filogenético de diferentes muestras de *Begomovirus* (*Geminiviridae*) con datos de secuencias almacenadas provenientes de colectas llevadas a cabo en diferentes cultivos en Guatemala en diferentes localidades, por miembros del Laboratorio de Protección Vegetal en el Centro de Estudios Agrícolas y Alimentarios de la Universidad del Valle de Guatemala. Estas secuencias se analizaron utilizando el programa DNASTar con un segmento del componente A de las muestras de *Begomovirus* que incluye la región común (RC) y parte del gen de la proteína de la cápside (CP) del virus. Se elaboró un árbol filogenético para establecer las relaciones existentes entre los diferentes *Begomovirus* encontrados en los diferentes cultivos y localidades, identificando la presencia de “Tomato Mosaic Havana Virus” (TMHV) y “Pepper Huasteco Yellow Vein Virus” (PHYVV) dentro de las muestras analizadas, también se encontró una alta probabilidad de que exista la presencia de “Tomato Leaf Curl Sinaloa Virus” (TLCSV), “Bean Golden Mosaic Virus” (BGMV) y Melon chlorotic leaf curl virus” (MCLF). Además, se generó una guía sobre el uso de este programa para su futura utilización en los laboratorios de Protección Vegetal. Esta herramienta será muy útil para estudios de diversidad y para diagnóstico e identificación de diferentes entidades de interés y sobre todo para su diferenciación de otras especies muy similares o en casos en que se sospeche la presencia de alguna especies o biotipo nuevo para el país.

Abstract

A phylogenetic analysis of different samples of *Begomovirus* (*Geminiviridae*) was done with sequences that were obtained from previous recollections done in different plants in Guatemala by members of the Plant Protection laboratory in the Center on Agricultural and Alimentary Studies of the Universidad del Valle de Guatemala. These sequences of the A component segment of *Begomovirus* samples were analyzed using the DNASTAR program. That segment included the Common region (CR) and part of the gen that conforms the capsid protein (CP) of the virus. A phylogenetical tree with the sequences was generated to establish the existing relationships between the different *Begomovirus* found in the different plants and localities, identifying the presence of Tomato Mosaic Havana Virus (TMHV) and Pepper Huasteco Yellow Vein Virus (PHYVV) between the analyzed samples as well as the high probability of presence of Tomato Leaf Curl Sinaloa Virus (TLCSV), Bean Golden Mosaic Virus (BGMV) and Melon Chlorotic Leaf Curl Virus (MCLF). Finally, a guide for use of the MegAlign Pro application and the DNASTAR program was created for its future use in the Plant Protection Laboratory.

I. Introducción

A través de la historia el ser humano se ha dedicado a agricultura como uno de los pilares que sostienen la economía mundial. Hoy en día uno de los grandes problemas que muchos de los agricultores tienen que controlar son los distintos tipos de virus que pueden ser distribuidos en los cultivos debido a distintos vectores, especialmente insectos.

El cultivo en invernadero ha sido uno de los principales métodos que están utilizando en la actualidad para la producción de algunos cultivos para exportación, gracias al control que permite sobre las condiciones ambientales que rodean a dicho cultivo y representa una barrera relativa para la entrada de vectores de virus y enemigos naturales de éstos. Pero, aún, así, no es un método perfecto ya que uno de los principales desafíos que existen para el cultivo en invernadero es el establecimiento de métodos adecuados para el manejo del cultivo y las normativas que se deben establecer para evitar que plagas ingresen al invernadero. Si un vector de virus, en este caso, logra entrar al invernadero, el cultivo correría gran peligro porque es un ambiente cerrado. Por eso es tan importante tener programas de control de plagas y programas de sanitación adecuados. Uno de los principales problemas que se han tenido que afrontar, es la propagación de un insecto de la familia *Aleyrodidae*, orden Hemiptera, la mosca blanca (*Bemisia tabaci* y *B. argentifolii*) (Mehta et al., 1994). Generalmente se considera a la mosca blanca como uno de los mayores problemas no sólo por el hecho de que el insecto se puede alimentar de la planta y puede producir daños directos a las plantas como la producción de una mielecita que atrae la propagación de un hongo llamado fumagina. Este cubre la superficie de la hoja y evita que pueda fotosintetizar el alimento. También es importante porque es un vector de distintos virus que causan enfermedades en los cultivos y porque ingresa fácilmente a un invernadero por diferentes medios, como por plantas nuevas, por el descuido de dejar una puerta mal cerrada, por la vestidura de un trabajador o aún por algún agujero que quede sin reparar en la malla del invernadero. Las moscas blancas se reproducen rápidamente y como son eficientes vectores de virus varios virus a un amplio rango de hospederos, pueden transmitirlos a varios cultivos rápidamente, causando grandes daños en algunos casos.

Muchas enfermedades causadas por virus, en especial que afectan a solanáceas, son causadas por distintos virus pertenecientes a la familia *Geminiviridae* y al género *Begomovirus* (Fauquet, C:M: 2008) infectan una gran variedad de plantas dicotiledóneas. Estas enfermedades reciben su nombre según la planta que infectan, como el caso del “Tomato yellow leaf curl virus” (TYLCV) en el tomate (*Solanum lycopersicum*) o “Pepper leaf curl virus” (PepLCV) en el chile pimiento (*Capsicum annuum*). El principal método que se utiliza para el control de plagas y enfermedades causadas por la mosca blanca está orientado hacia el control del vector utilizando distintos controles biológicos o químicos, programas de sanidad vegetal estrictos, el uso de semillas certificadas, y en años recientes se ha tratado de utilizar controles genéticos directamente en la planta para evitar que ésta pueda infectarse con el virus (Teni, 2004).

En los años de 2006 y 2007 en la Universidad del Valle de Guatemala se realizó una investigación en conjunto con Popoyán, Fasagua y la Universidad de California, Davis, Estados Unidos, en el cual se logró identificar la presencia de *Begomovirus* en plantaciones de tomate en áreas del El Progreso y Asunción Mita, en el departamento de Jutiapa. Esta investigación llegó a determinar la presencia de varios *Begomovirus* en ambas regiones pero los más prevalentes fueron: Pepper Huasteco Yellow Vein Virus (PHYVV), Pepper Golden Mosaic Virus (PGMV), Tomato Mosaic Havana Virus (ToMHV), Tomato Severe Leaf Curl Virus (TSLCV) y finalmente encontrado a partir

del 2006, Tomato Yellow Leaf Curl Virus (TYLCV), un *Begomovirus* con un solo anillo y que viene del viejo mundo, esto fue en la región nororiental de Guatemala (Tesis Mónica Morales y Estuardo Solares).

En 2005, Nakhla et al. reportaron ocho grupos principales de *Begomovirus* en cultivos de tomate en Guatemala, Honduras y Costa Rica. De éstos los más comunes en Guatemala son *Pepper golden mosaic virus* (PepGMV), *Pepper huasteco yellow vein virus* (PHYVV), *Tomato leaf curl Sinaloa Virus* (ToLCSinV) y *Tomato Mosaic Havana virus* (ToMHV).

II. Objetivos

A. General

- Elaboración de una guía para la utilización del programa DNASTar para la comparación de secuencias y elaboración de árboles filogenéticos.

B. Específicos

- Realizar un análisis filogenético de las secuencias proveídas por los laboratorios de la Universidad el Valle de Guatemala utilizando el programa DNASTar.
- Utilizar el árbol filogenético generado por el análisis para establecer las relaciones existentes entre los diferentes *Begomovirus* en los diferentes cultivos y localidades para determinar si pertenecen a *Begomovirus* conocidos para reportar su presencia en Guatemala o son *Begomovirus* nuevos.

III. Justificación

Con la información en las secciones anteriores se resalta la importancia de poder controlar e identificar a tiempo una infección de *Begomovirus* en los cultivos, pero la posibilidad de infecciones mixtas, así como una sintomatología muy similar entre las diferentes variedades de virus, complican la identificación de éste a simple vista. Debido a esto es necesario utilizar métodos de diagnóstico adecuados para poder identificar al virus (Por ejemplo: ELISA, PCR, RT-PCR, DNA barcoding y secuenciación de genomas) y poder tomar las medidas necesarias para controlarlo. La ventaja de usar métodos genéticos sobre los serológicos es que, con los últimos, la muestra se puede mandar a secuenciar y por comparaciones con secuencias existentes en el GenBank se puede llegar a identificar el virus hasta especie o cepa, esto no lo puede hacer con ELISA a no ser que se haya desarrollado un anticuerpo específico contra ese virus y no exista entrecruzamiento.

Si no se sabe cuál es el virus que afecta a un cultivo es muy difícil que se pueda realizar un control adecuado del mismo y en este problema es donde los métodos bioinformáticos pueden ser utilizados para mejorar el proceso y evitar que las personas tomen decisiones incorrectas que causen mayores pérdidas en los cultivos. Cuando existe una gran cantidad de muestras de este tipo que analizar es cuando programas de bioinformática (Por ejemplo: Mega X, DNASTar, bioedit, Etc.) pueden demostrar de lo que son capaces al poder analizar los datos y ayudarnos a determinar las relaciones entre los mismo y así llegar a un punto claro de identificación del posible virus.

En Guatemala, el campo biotecnológico no se encuentra muy desarrollado en comparación a otros países, por lo que el aprender a utilizar estos programas y poder pasar el conocimiento a otras personas es una manera efectiva de fomentar el uso de los mismos y explotar el potencial que poseen. Debido a esto, los primeros pasos necesarios para poder normalizar el uso de los mismos requieren que se realicen estudios con éstos para dejar la documentación necesaria y poder replicar los procedimientos y aplicarlos a diferentes situaciones.

IV. Marco teórico

A. Antecedentes

a. *Begomovirus*

Begomovirus se refiere a un género de virus pertenecientes a la familia *Geminiviridae* con más de 100 especies descritas, esta familia se clasifica según la organización del genoma, tipo de insecto vector y rango de hospederos; Los *Begomovirus* son comúnmente conocidos por infectar una gran variedad de plantas dicotiledóneas. Éstos se caracterizan por tener un genoma circular de una cadena simple de ADN, siendo tanto monopartitos como bipartitos, dicha cadena se encuentra rodeada por una cápside icosaédrica, formando estructuras de partículas gemelas (Lazarowitz & Shepherd, 1992). Son transmitidos únicamente por especies de la mosca blanca, en Guatemala por *Bemisia tabaci* en todo el país, tanto biotipo A y biotipo B. Este último se encuentra principalmente en campos de cucurbitáceas en la zona oriental del país (comunicación personal, Palmieri, 2020 palmieri@uvg.edu.gt).

Los *Begomovirus* son generalmente clasificados en los del nuevo mundo y los del viejo mundo (Europa, Asia y África), siendo los del viejo mundo de una sola partícula mientras que los del nuevo mundo de dos partículas (Gafni & Epel, 2002). Con el intercambio comercial de la actualidad, esto ya no es tan simple porque han sido transportados virus de una partícula al nuevo mundo y viceversa. Es por esto que es importante saber con qué *Begomovirus* estamos afrontándonos en nuestro cultivo. En cuanto a los *Begomovirus* de dos partículas, éstos presentan dos componentes designados como A y B, cada uno de ellos consistiendo en alrededor de 2,600 nt. Los genes del componente A son los que se encuentran involucrados en la replicación y formación de la cápside mientras que los del componente B determinan el movimiento del virus dentro de la planta, los síntomas que éstas presentan y los posibles hospederos (Ghanim et al., 1998).

1. Genoma de *begomovirus*

Los *begomovirus* bipartitos están caracterizados por la presencia de 2.7 kb de ADN-A y 2.6 kb de ADN-B, mientras que los monopartitos presentan 2.7 kb de ADN-A. Ambos componentes genómicos presentan marcos de lectura parcialmente traslapados (ORFs) de manera bidireccional, el ADN-A posee seis ORFs: AV1/V1, AV2/V2, AC2/C2, AC3/C3 Y AC4/C4. También se ha reportado la presencia de AC5/C5 en algunos *begomovirus* bipartitos. El ADN-B posee dos ORFs: BV1 y BC1. Los ORFs virales se encuentran separados por una región intergenérica (IR) y poseen una región común (CR) que consiste en nucleótidos conservados entre los componentes A y B (Roshan, Kulshreshtha & Hallan, 2017).

Gafni y Epel en 2002 mencionan que las proteínas codificadas por el genoma de los *Begomovirus* son:

CP (Proteína de la cápside): V1, es esencial para la propagación del virus y como el nombre indica la generación de la cápside. V2 es esencial en el movimiento del virus dentro del hospedero, dentro y fuera del núcleo celular.

REP (Proteína de la replicación): C1, es la proteína que se encarga de la replicación del virus.

TraP (Proteína de la activación transcripcional): C2, está proteína se encarga de iniciar la transcripción por medio de un promotor de proteína de cápside. Se encuentra en el núcleo.

Ren (proteína promotora de la replicación): C3, promueve la acumulación de ADN viral e interactúa con la proteína C1

C4: se encarga del desarrollo de los síntomas en el hospedero y la división celular.

La figura 1 muestra la estructura de los genomas de los *Begomovirus*, tanto bipartitos como monopartitos. Se puede visualizar las diferentes regiones del genoma que codifican para las proteínas y la dirección de su síntesis.

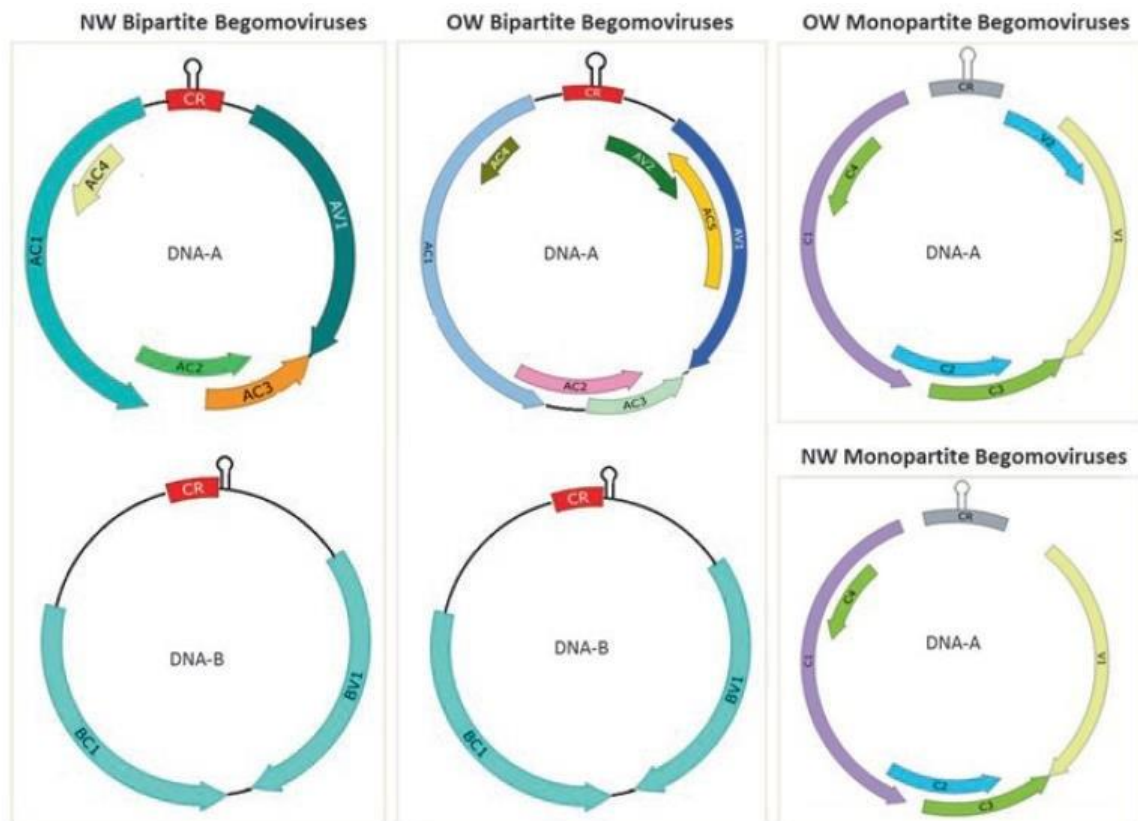


Figura 1. Diagrama de la organización genómica de los begomovirus del nuevo y viejo mundo, tanto monopartitos como bipartitos. Fuente: Roshan, Kulshreshtha & Hallan, 2017.

2. Replicación del virus

El modo de replicación predominante en los *Begomovirus* es el mecanismo de círculo rodante (RCR). Aunque ambos componentes del virus se replican de esta manera, el componente A es el único que se puede replicar de manera independiente, es decir sin la presencia del componente B, mientras que el componente B no es capaz de replicarse sin la presencia del componente A (Yadava, Suyal & Mukherjee, 2010).

Una vez que el genoma viral es introducido al núcleo, este inicia un proceso de replicación de 3 etapas: En la primera etapa (Iniciación), en conjunto con únicamente los factores celulares del hospedero, el ssADN es convertido en intermediarios o formas replicativas (RF) de dsADN supercoloide. El dsADN resultante es transcripcionalmente activo y lleva a la transcripción de varios factores virales. La segunda etapa (Elongación) inicia con el dsADN actuando como un molde para la síntesis del ssADN usando tantos factores virales como del hospedero a través del modo RCR. El ssADN sintetizado puede volver a entrar al ciclo de replicación de ADN. Este ssDNA provee un fundamento para el ensamblaje viral y la transmisión en la siguiente etapa. En la última etapa (Terminal) la proteína Rep corta el nuevo ssADN sintetizado para así eventualmente generar muchas copias de éste y replicar el ADN viral (Yadava, Suyal & Mukherjee, 2010). En la Figura 2 se presenta un diagrama de la replicación del ADN de los *Begomovirus*.

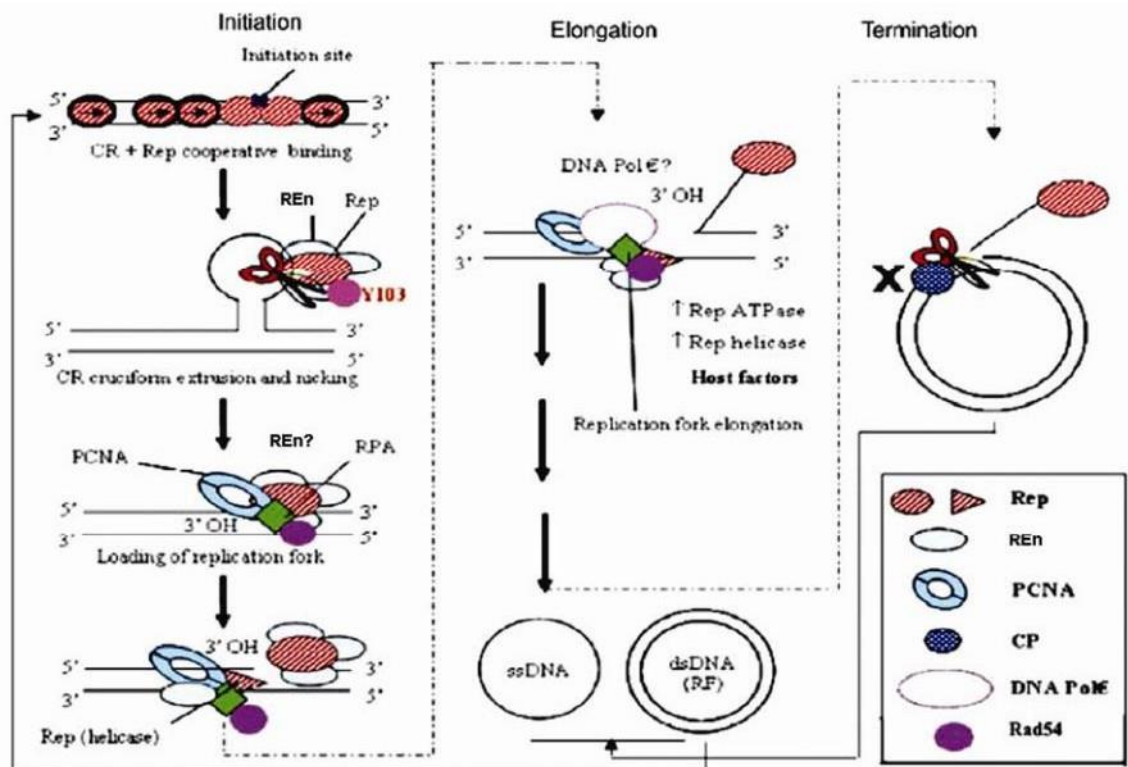


Figura 2. Diagrama de la replicación de ADN en el modo RCR. Fuente: Yadava, Suyal & Mukherjee, 2010.

3. Posibilidad de infecciones mixtas

Especialmente en las regiones tropicales y subtropicales se ha visto una tendencia a que las infecciones mixtas se produzcan. El problema con estas infecciones mixtas es que en ciertos casos se puede generar una interacción de co-infección en la cual dos o más virus provocan síntomas mucho más severos de lo normal en la planta con una infección causada por un sólo virus o en otros casos uno de los virus puede afectar el ciclo de replicación de uno de los otros virus que se encuentran hospedados en la planta (Harrison & Robinson, 1999).

Que uno de los virus afecte el ciclo de replicación de otro implica un problema mayor, ya que cuando esto sucede se pueden producir eventos de recombinación o complementación que pueden llevar a la evolución del virus, especialmente en *Begomovirus* donde se ha visto que la recombinación es demasiado frecuente llegando al punto donde puede ocurrir entre cepas, especies, géneros e incluso familias (Harrison, 1985).

4. Hospederos del virus

Se han realizado pruebas de inoculación en laboratorio para determinar el rango de hospederos de *Begomovirus* así como muestreos en campo que incluyen a 6 familias principales de plantas: *Asclepiadiaceae*, *Asteraceae*, *Fabaceae*, *Malvaceae*, *Solanaceae*, *Gentianaceae*, *Cleomaceae*, *Cucurbitaceae* y *Apiaceae* (Nakhla, & Maxwell, 1998). También se han reportado algunas malezas y plantas que no presentan síntomas como posibles hospederos.

5. Síntomas en la planta

Los síntomas de infección por *Begomovirus* varían dependiendo del virus, la planta en la cual se encuentre, la edad de la planta, las condiciones del ambiente en el cual se encuentren, la cepa del virus, pero generalmente se pueden mencionar: Mosaico amarillo, clorosis moteada, márgenes de las hojas con clorosis, hojas rizadas y deformadas, pliegues en las hojas, hojas más pequeñas de lo normal, plantas de menor tamaño y abscisión de la flor (Lapidot et al., 2001). Esta variedad de síntomas y la similitud que tienen algunos con síntomas producidos por otros virus u otros patógenos y/o por otros factores como desnutrición o falta de algún elemento importante, es lo que ha llevado a que sea necesario realizar una identificación por medio de PCR para poder realizar un diagnóstico exacto.

b. Mosca blanca

Este virus es exclusivamente transmitido por un vector insecto, la mosca blanca, que pertenece al orden Hemiptera, suborden Sternorrhyncha y familia Aleyrodidae. Se conocen más de 1200 especies de mosca blanca, pero las especies más comunes son las del género *Bemisia*, el cual se cree que se originó en el sur este de Asia o África, de este género es importante mencionar las dos especies más conocidas: *Bemisia tabaci* (biotipo A y biotipo B) y *T. vaporariorum* (Brown, 2000), siendo la primera el vector de *Begomovirus*. La mosca permite que la transmisión del virus ocurra rápida y eficientemente gracias a la gran variedad de plantas que este insecto consume en su dieta. La transmisión del virus ocurre de forma persistente, circulatoria y no propagativa y ocurre cuando la mosca se alimenta de los jugos del floema de la planta.

Se ha visto que el virus no se puede transmitir por inoculación mecánica o por semilla, pero, se ha logrado una transmisión exitosa del virus por medio de injertos. Para que una mosca pueda transmitir el virus es necesario que ésta la adquiera ya sea durante su estado de larva o de adulto, de una planta previamente infectada con el virus. Las larvas no son capaces de transmitir el virus, pero sí de contraerlo. Mientras que los adultos son tanto capaces de transmitir el virus como de contraerlo, pero no pueden pasarlo a su descendencia, lo cual conlleva a una gran capacidad de transmisión en los cultivos que presentan una infestación de mosca blanca, ya que conserva el virus durante toda su vida. Por lo tanto, para que haya una infección de *Begomovirus* en un cultivo es necesaria la presencia de al menos una planta infectada y el insecto vector (Serra et al., 2006), o que el viento lleve al campo moscas infectadas de campos aledaños.

c. Importancia económica de los cultivos

1. Tomate (Solanum lycopersicum):

Guatemala es el principal productor de tomate en Centro América, el cultivo de tomate se encuentra distribuido mayormente en los departamentos de Jutiapa, Baja Verapaz y Chiquimula con una producción de 7,277,590 quintales de tomate producidos en 2015 a partir de la cual se generaron 18,297,146.00 USD en exportación de tomate (MAGA, 2015); De acuerdo Scotta, Sánchez & Arregui, 2015, existe una reducción del 60% a 80% de la producción en el tomate debido a las infecciones producidas por *Geminivirus (Begomovirus)*, transmitido por la mosca blanca. Con esto se puede ver que, de no ser controlada, una infección en las plantaciones podría llegar a producir una pérdida en casos extremos de 10,978,287.60 USD hasta 14,637,716.80 USD.

También, en el caso de pérdidas extremas del cultivo hay que tomar en cuenta el desempleo que esto generaría, en el 2015 se estima según datos de BANGUAT Y MAGA que el cultivo de tomate genera 2,788,569 jornales por año, que son equivalentes a 9,959 empleos permanentes.

2. Chile pimiento (Capsicum annum):

Los principales productores de Chile pimiento en Guatemala se encuentran en los departamentos de: Jutiapa, Baja Verapaz, Guatemala, Chiquimula y Alta Verapaz. Para el año 2015 se produjeron alrededor de 1,237,060 quintales de Chile pimiento que es equivalente a 11,522,167.00 USD en exportación. También se generan alrededor de 827,450 jornales por año que son equivalentes a 2,955 empleos permanentes (MAGA, 2015). Se estima que el Chile pimiento tiene una reducción del 50% a 60% en cuanto a producción por planta al estar infectada con *Begomovirus* por lo que, si asumimos el caso planteado anteriormente, una infección de *Begomovirus* en este cultivo podría llegar a generar pérdidas de alrededor de 5,761,083.50 USD hasta 6,913,300.20 USD.

3. Frijol (Phaseolus vulgaris):

La producción de frijol en Guatemala se encuentra distribuida a lo largo de todo el territorio nacional, con los departamentos que cuentan con la mayor producción de éste siendo: Petén, Jutiapa, Chiquimula, Santa Rosa, Jalapa, Quiché, Alta Verapaz, Huehuetenango, Guatemala y Chimaltenango. La producción se encuentra dividida en frijol rojo y frijol blanco, pero la suma de ambos para el año 2015 generó 1,585,967.00 USD y alrededor de 15,944,350 jornales por año que es un equivalente a

56,944 empleos permanentes (MAGA, 2015). En el frijol se han reportado pérdidas de alrededor del 60% a 75% (Inoue-Nagata, Lima & Gilbertson, 2016) por lo que nuevamente asumiendo el caso anterior podría generar pérdidas desde 951,580.20 USD hasta 1,189,475.25 USD.

4. Melón (*Cucumis melo*):

El cultivo de melón se localiza principalmente en Zacapa con un 86% de la producción nacional proveniente de dicho departamento. La exportación de Melón en el año 2016 generó 160,636,759.00 USD, Específicamente de melón Cantaloupe (MAGA, 2016). Debido a infecciones por begomovirus se han reportado perdidas desde el 80% hasta 90% (Barboza et al., 2019) por lo que en este caso podríamos asumir que una infección de begomovirus sería podría causar una pérdida de alrededor de 128,509,167.2 USD hasta 144,572,813.1 USD.

d. DNASTar

DNASTar ha sido un software de bioinformática muy utilizado en el campo de la investigación durante los últimos 35 años, solamente en el año 2019 se reportaron 3,370 publicaciones que han utilizado el programa en diferentes tipos de estudios (DNASTar, 2020).

Estudios utilizando el programa incluyen: El análisis del genoma del pez de cueva mexicano *Astyanax mexicanus* (Berning et al. 2019) y la investigación del virus Zika en mosquitos en el estado de Florida en los Estados Unidos (Mutebi, 2016). Gracias a las diferentes aplicaciones proveídas dentro del programa muchas investigaciones pueden ser realizadas en temas de: Alineamiento de secuencias múltiples, Diseño de primers para PCR, Predicción de estructura de proteínas, Análisis de genoma completo, ensamble de genoma *De novo*, etc (DNASTar, 2020).

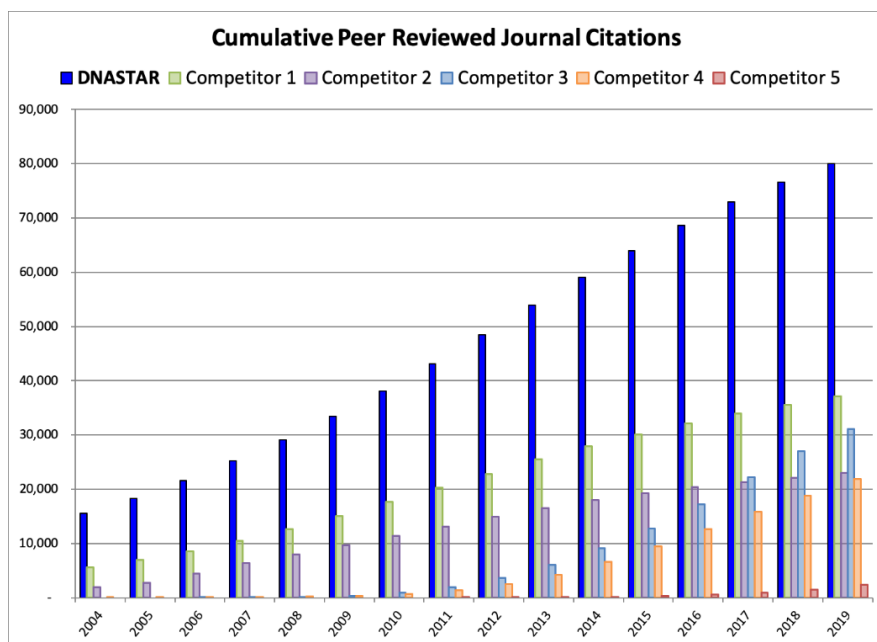


Figura 3. Gráfico de barras demostrando la cantidad de publicaciones en revistas científicas utilizando DNASTar vs los programas de competencia en el mercado a través de los años (Fuente: DNASTar, 2020).

V. Metodología

En este caso se realizó una comparación de diferentes secuencias de *Begomovirus* para determinar si estas pertenecen a *Begomovirus* conocidos cuya presencia haya sido previamente reportada en Guatemala y la relación entre los diferentes virus provenientes de distintos cultivos para poder obtener una aproximación de la identificación de estos en el caso de no pertenecer a una variedad previamente reportada.

A. Recolección de secuencias de *Begomovirus* en Guatemala.

Todas las secuencias se obtuvieron de un banco de secuencias existente en los Laboratorios de Protección Vegetal de la Universidad del Valle de Guatemala, obtenidas de estudios previamente realizados desde los años 2000 hasta 2018. Dichas secuencias provienen de diferentes cultivos, localidades y épocas. Las secuencias fueron analizadas tal y como fueron reportadas por Macrogen en Corea, La compañía encargada de la secuenciación.

Los fragmentos fueron secuenciados por medio de los primers generales 514 y 1048. Dichos primers amplifican un fragmento de aproximadamente 555 pares de bases en el anillo A que incluye parte del gen de la cápside (AV1).

B. Muestras de Genbank

Se realizó una búsqueda de secuencias del componente A de *Begomovirus* que incluye la región común (RC) y parte del gen de la proteína de la cápside (CP) del virus para utilizar como comparación dentro del análisis filogenético y poder identificar las muestras recolectadas en Guatemala. Adicionalmente se seleccionaron 2 muestras de *T. vaporariorum* con el tamaño adecuado (Aproximadamente 600 pares de bases) para poder ser utilizadas como “outgroup” en el análisis filogenético.

C. Análisis filogenético

Para el análisis filogenético se utilizó el programa DNASTar con la aplicación MegAlignPro versión 17.2.0.153. Los datos obtenidos del genbank y las muestras recolectadas en Guatemala se cargaron al programa en formato FASTA, así como también una muestra general de *T. vaporariorum* para poder ser utilizada como “Outgroup” dentro del árbol filogenético.

D. Determinación del modelo de sustitución de nucleótidos

Un modelo de sustitución de nucleótidos permite un cálculo de la probabilidad de intercambio entre nucleótidos a lo largo de las ramas de un árbol filogenético (Lemmon & Moriarty, 2004). Para determinar el modelo necesario para el análisis de estas secuencias se utilizó el programa Jmodeltest-2, que permite la selección estadística de un modelo de sustitución de nucleótidos en base a “Phyml” (un algoritmo simple, rápido y preciso para estimar grandes filogenias por medio de máxima verosimilitud). utilizando el análisis de “Likelihood scores” se determinó el modelo de sustitución de nucleótidos óptimo para el análisis de los datos utilizados utilizando tanto el modelo de “Criterio de información de Akaike” (AIC), La “corrección del criterio de información de Akaike” (AICc) y el

“Criterio de información Bayesiana” (BIC), siendo estos los métodos u estrategias de selección más aceptados (Posada, 2008).

E. Alineamiento de secuencias y árbol filogenético

Las secuencias se alinearon utilizando Clustal W, el cual es un proceso de alineamiento progresivo que alinea las secuencias más similares de primero en orden hasta llegar a las de menor similitud en base a un algoritmo. Después se realizó el análisis de “Neighbor-Joining”, Este análisis permite realizar un árbol filogenético de manera rápida para obtener una idea de la distribución de nuestros datos a lo largo del árbol. Y finalmente el análisis de” Maximum likelihood” o máxima verosimilitud que nos permite utilizar el modelo de sustitución de nucleótidos seleccionado por Jmodeltest para obtener el árbol filogenético deseado generado por medio de un cálculo más extensivo y un método bayesiano (Kumar et al., 2018). A partir de este árbol se pudo determinar la relación entre las diferentes muestras, así como la identificación de éstas.

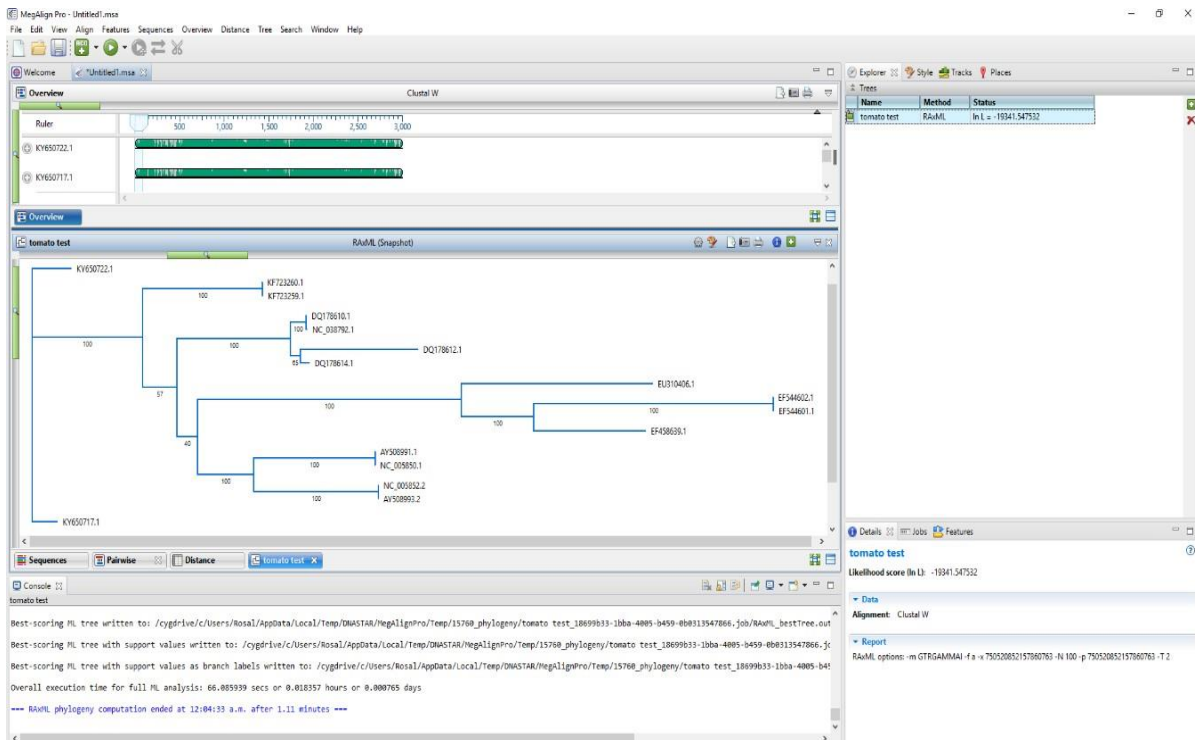


Figura 4. Captura de pantalla de un alineamiento y análisis filogenético de muestras de prueba en la aplicación MegAlignPro.

F. Redacción de la guía de uso del programa

Después de haber obtenido los resultados del análisis filogenético se redactará un documento describiendo la utilización del programa paso por paso, incluyendo gráficas y capturas de pantalla, para el futuro uso de éste dentro de los laboratorios de la Universidad del Valle de Guatemala.

VI. Resultados

A. JModelTest

Los modelos seleccionados por Jmodeltest fueron TPM2uf+ I (AICc Y AIC) y HKY + I (BIC). De acuerdo con Burnham & Anderson, 2004 AICc siempre debe ser preferido sobre AIC ya que según lo establecido los cálculos de AICc tiende a converger con AIC siempre y cuando contemos con un numero de muestra lo suficientemente grande. Por otro lado, El modelo preferido BIC es diferente y de acuerdo con Posada, 2008 la selección del modelo debe ser tomada según preferencia, por lo que se decidió utilizar el modelo seleccionado por el análisis BIC.

ID	Name	Partition	-lnL	p	AIC	deltaAIC	weight	cumWeight	uDelta
38	TPM2uf+I	010212	3368.2771	86	6908.5543	0.0	0.2611	0.2611	-
40	TPM2uf+I+G	010212	3367.5805	87	6909.161	0.6067	0.1928	0.454	-
62	TIM2+I	010232	3368.2573	87	6910.5145	1.9602	0.098	0.552	-
64	TIM2+I+G	010232	3367.5805	88	6911.161	2.6067	0.0709	0.623	-
14	HKY+I	010010	3370.6957	85	6911.3915	2.8372	0.0632	0.686	-
22	TrN+I	010020	3369.7487	86	6911.4974	2.9431	0.06	0.746	-
78	TVM+I	012314	3367.9622	88	6911.9245	3.3702	0.0484	0.794	-
16	HKY+I+G	010010	3370.0045	86	6912.009	3.4547	0.0464	0.841	-
86	GTR+I	012345	3367.0908	89	6912.1816	3.6273	0.0426	0.883	-
80	TVM+I+G	012314	3367.2681	89	6912.5363	3.982	0.0357	0.919	-
88	GTR+I+G	012345	3366.3376	90	6912.6752	4.1209	0.0333	0.952	-
39	TPM2uf+G	010212	3371.1609	86	6914.3217	5.7674	0.0146	0.967	-
24	TrN+I+G	010020	3370.7989	87	6915.5978	7.0435	0.0077	0.975	-
63	TIM2+G	010232	3371.1707	87	6916.3414	7.7871	0.0053	0.98	-
23	TrN+G	010020	3372.5526	86	6917.1052	8.5509	0.0036	0.984	-
15	HKY+G	010010	3373.6019	85	6917.2038	8.6495	0.0035	0.987	-
87	GTR+G	012345	3369.785	89	6917.5701	9.0158	0.0029	0.99	-
79	TVM+G	012314	3370.852	88	6917.704	9.1497	0.0027	0.993	-

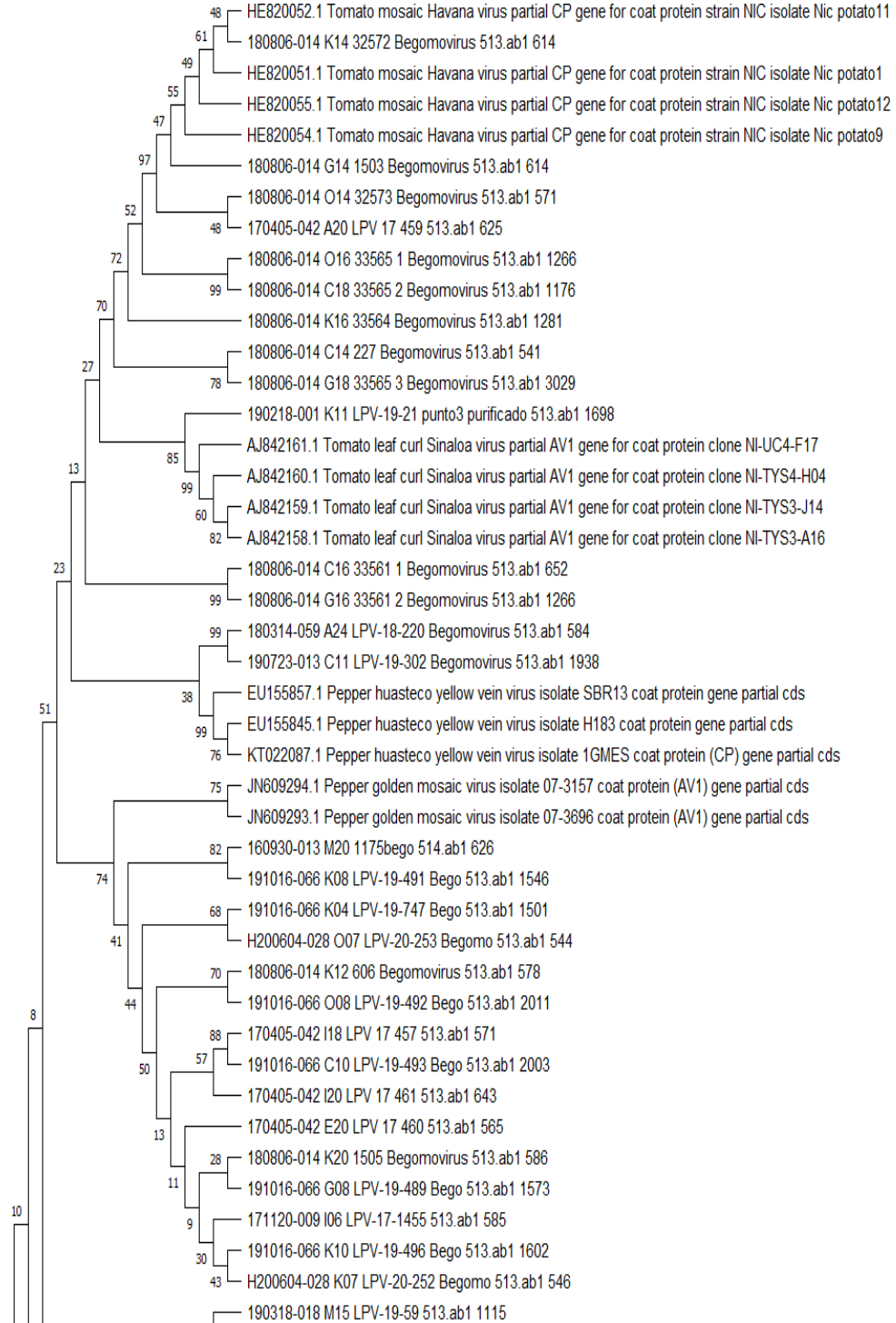
ID	Name	Partition	-lnL	p	AICc	deltaAICc	weight	cumWeight	uDelta
38	TPM3uf+I	010212	3368.2771	86	6917.6015	0.0	0.2846	0.285	-
40	TPM2uf+I+G	010212	3367.5805	87	6918.4241	0.8227	0.1886	0.473	-
62	TIM2+I	010232	3368.2573	87	6919.7777	2.1762	0.0959	0.569	-
14	HKY+I	010010	3370.6957	85	6920.2253	2.6238	0.0767	0.646	-
22	TrN+I	010020	3369.7487	86	6920.5446	2.9431	0.0653	0.711	-
64	TIM2+I+G	010232	3367.5805	88	6920.6428	3.0413	0.0622	0.773	-
16	HKY+I+G	010010	3370.0045	86	6921.0562	3.4547	0.0506	0.824	-
78	TVM+I	012314	3367.9622	88	6921.4063	3.8049	0.0425	0.866	-
86	GTR+I	012345	3367.0908	89	6921.8848	4.2834	0.0334	0.9	-
80	TVM+I+G	012314	3367.2681	89	6922.2395	4.638	0.028	0.928	-
88	GTR+I+G	012345	3366.3376	90	6922.6024	5.001	0.0234	0.951	-
39	TPM2uf+G	010212	3371.1609	86	6923.3689	5.7674	0.0159	0.967	-
24	TrN+I+G	010020	3370.7989	87	6924.8609	7.2595	0.0075	0.975	-
63	TIM2+G	010232	3371.1707	87	6925.6046	8.0031	0.0052	0.98	-
15	HKY+G	010010	3373.6019	85	6926.0377	8.4362	0.0042	0.984	-
23	TrN+G	010020	3372.5526	86	6926.1524	8.5509	0.004	0.988	-
79	TVM+G	012314	3370.852	88	6927.1859	9.5844	0.0024	0.99	-
87	GTR+G	012345	3369.785	89	6927.2733	9.6719	0.0023	0.993	-

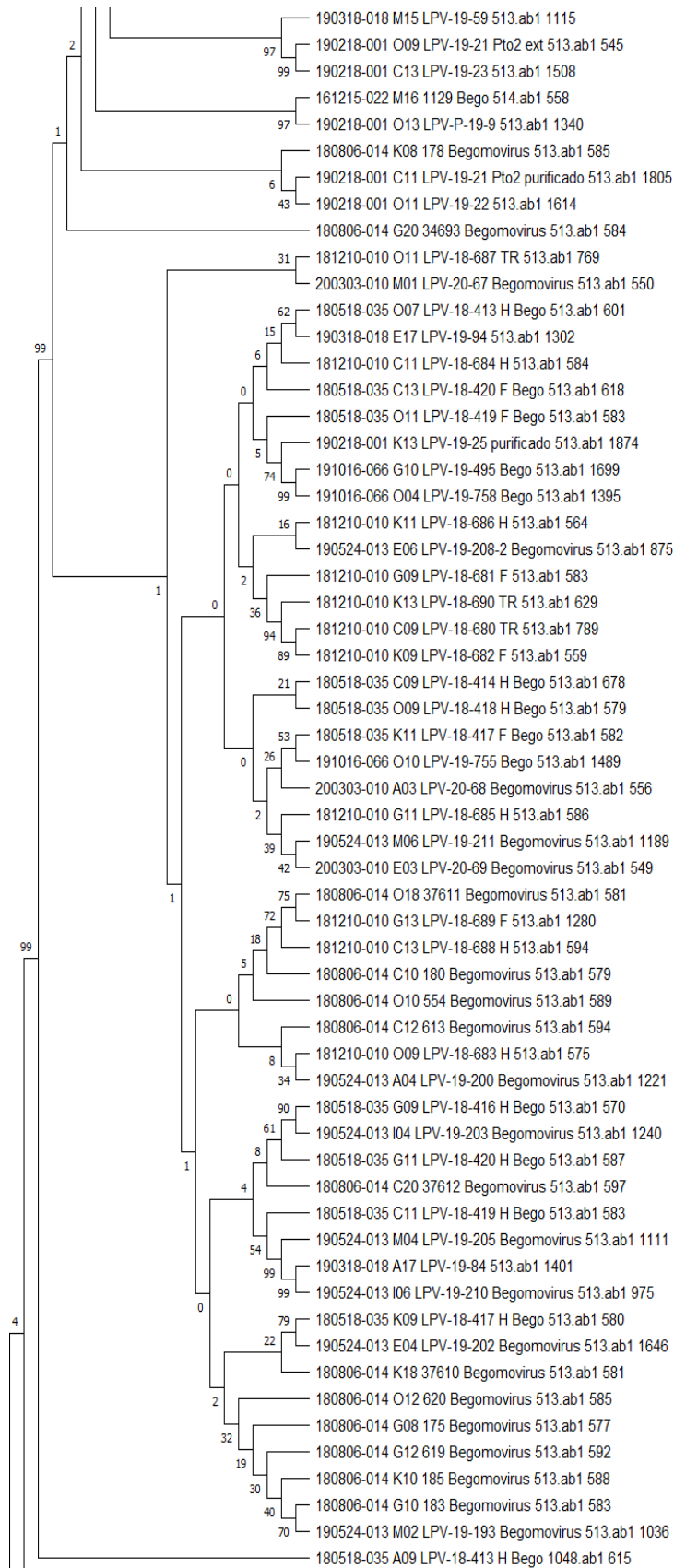
ID	Name	Partition	-lnL	p	BIC	deltaBIC	weight	cumWeight	uDelta
14	HKY+I	010010	3370.6957	85	7375.6797	0.0	0.6773	0.677	-
38	TPM2uf+I	010212	3368.2771	86	7378.3048	2.6251	0.1823	0.86	-
22	TrN+I	010020	3369.7487	86	7381.2479	5.5682	0.0418	0.901	-
15	HKY+G	010010	3373.6019	85	7381.4921	5.8124	0.037	0.938	-
16	HKY+I+G	010010	3370.0045	86	7381.7595	6.0798	0.0324	0.971	-
39	TPM2uf+G	010212	3371.1609	86	7384.0722	8.3925	0.0102	0.981	-
40	TPM2uf+I+G	010212	3367.5805	87	7384.3737	8.6939	0.0088	0.99	-
62	TIM2+I	010232	3368.2573	87	7385.7272	10.0475	0.0045	0.994	-
23	TrN+G	010020	3372.5526	86	7386.8557	11.176	0.0025	0.997	-
30	TPM1uf+I	012210	3373.1555	86	7388.0615	12.3818	0.0014	0.998	-
46	TPM3uf+I	012012	3373.9867	86	7389.7239	14.0442	6.0E-4	0.999	-
24	TrN+I+G	010020	3370.7989	87	7390.8105	15.1307	4.0E-4	0.999	-
63	TIM2+G	010232	3371.1707	87	7391.5541	15.8744	2.0E-4	0.999	-
64	TIM2+I+G	010232	3367.5805	88	7391.8359	16.1561	2.0E-4	1	-
78	TVM+I	012314	3367.9622	88	7392.5994	16.9197	1.0E-4	1	-
32	TPM1uf+I+G	012210	3372.4088	87	7394.0303	18.3506	1.0E-4	1	-
31	TPM1uf+G	012210	3376.1673	86	7394.0851	18.4054	1.0E-4	1	-
48	TPM3uf+I+G	012012	3373.0901	87	7395.3928	19.7131	0.0	1	-

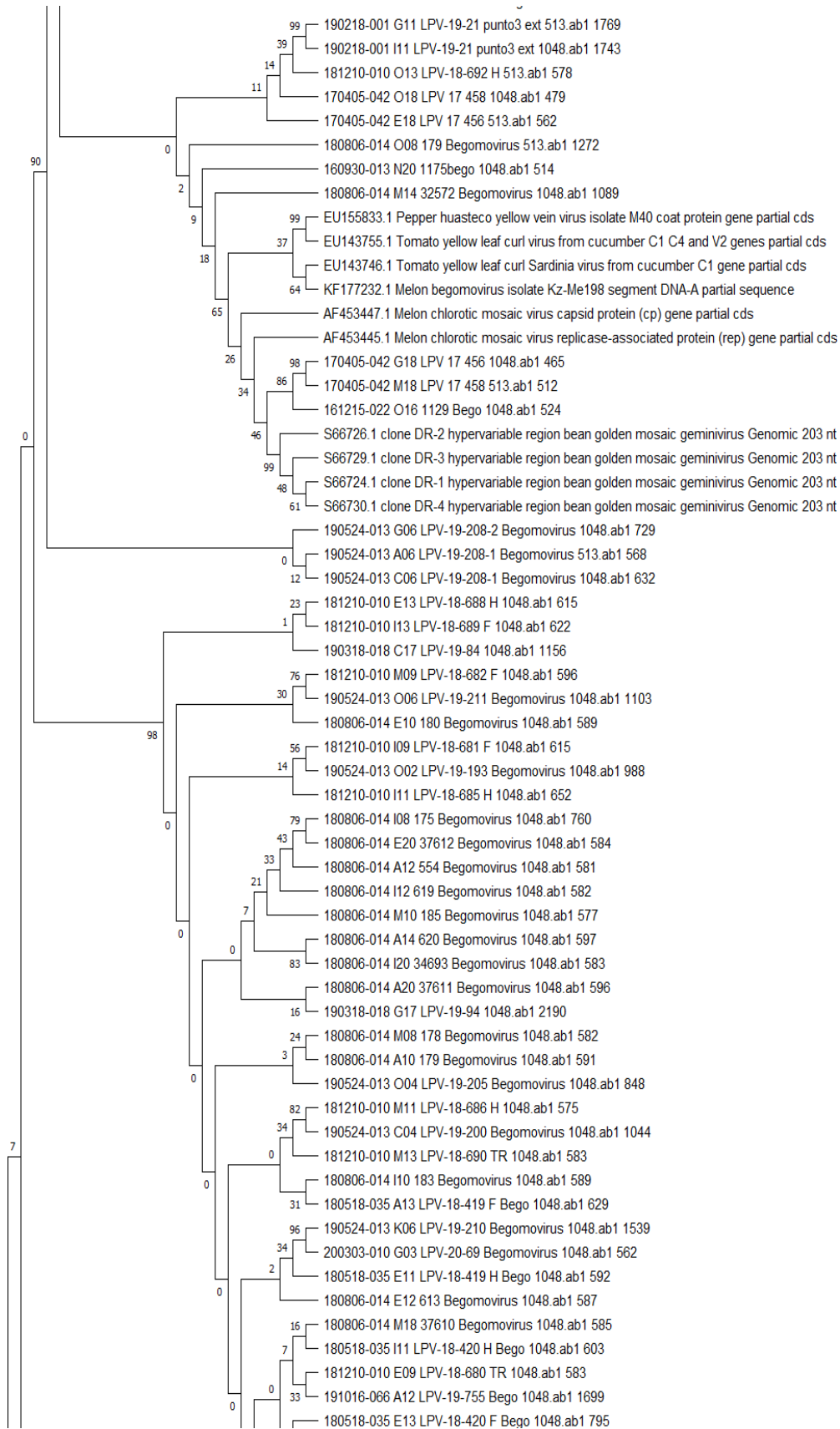
Figura 5. Capturas de pantalla del análisis realizado en JModelTest para AIC, AICc y BIC respectivamente

B. *Árbol de Neighbor-Joining*

El siguiente es el primer árbol generado y que como se mencionó con anterioridad no es el más exacto, pero es el que nos permite tener una idea de la relación entre las diferentes muestras. En las ramas se observa el valor de “Bootstrap” de las diferentes subdivisiones en el árbol, este valor nos indica la cantidad de veces que la rama que le sigue apareció dentro de uno de los árboles generados dentro del análisis (Valores de 0 a 100).







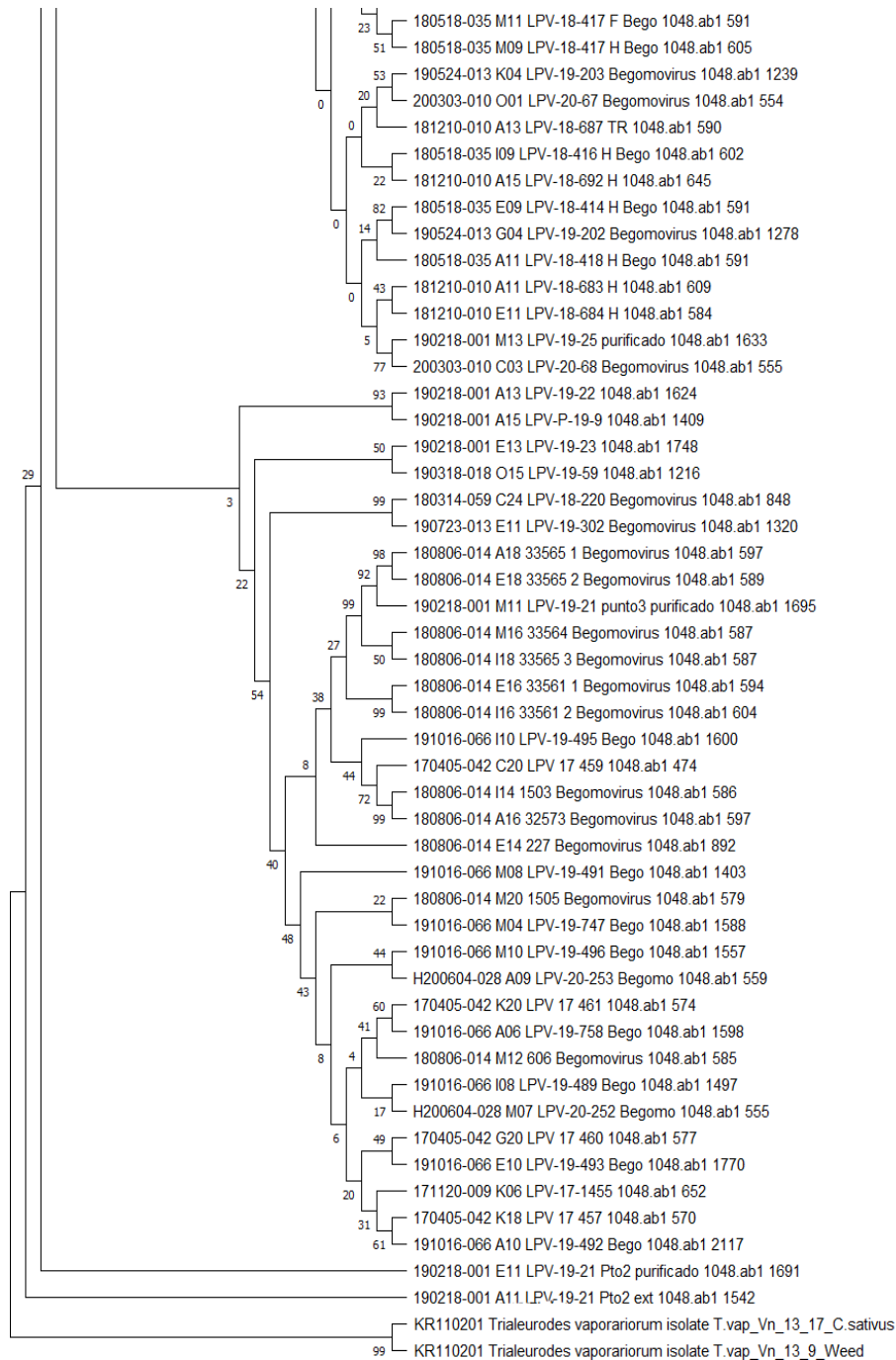
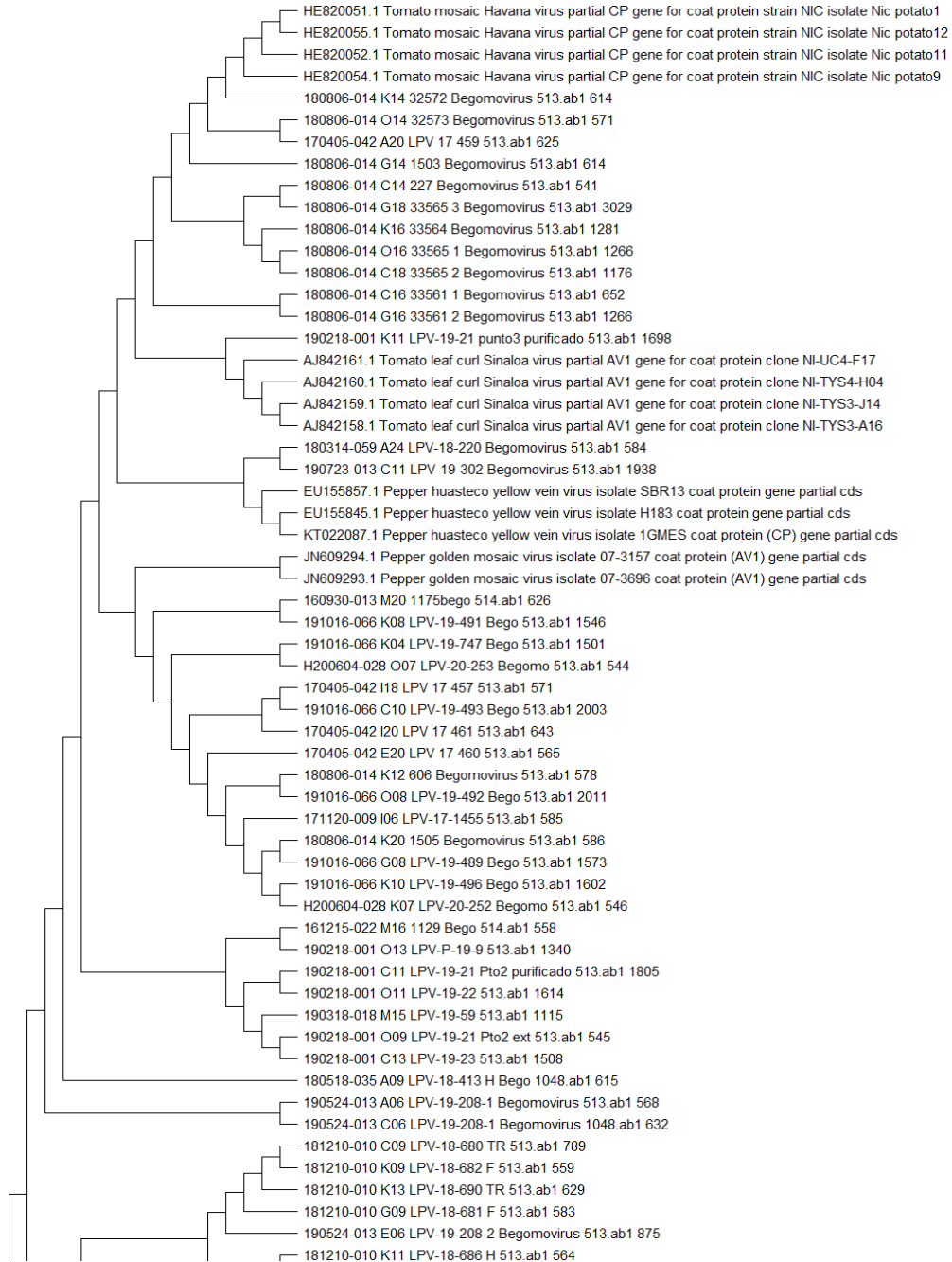


Figura 6. Árbol filogenético generado utilizando el análisis de “Neighbor-joining” con los valores de Bootstrap en las ramas (Valores de 0 a 100) Indicando el número de veces que dicha rama estuvo presente en uno de los árboles generados (1000 repeticiones); Posteriormente editado para mejor presentación.

C. Árbol de *Maximum Likelihood* HKY + I (BIC)

Esté árbol es el resultado de 1000 repeticiones utilizando el modelo HKY + I (Hasegawa-Kishino-Yano) seleccionado por el análisis BIC en Jmodeltest. Este árbol es más en comparación al árbol generado utilizando el método de “Neighbor-Joining”, pero aun así los resultados siempre tienden a ser bastante similares entre ambos. En este caso debido a la utilización de un modelo de sustitución de nucleótidos es fácil asumir las relaciones entre las muestras según su posición en las ramas del árbol, En otras palabras, Mientras más cercana se encuentre una rama de otra mayor será su relación entre sí.



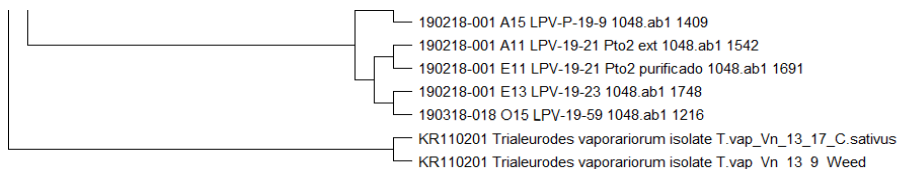


Figura 7. Árbol filogenético de “Maximum Likelihood” generado utilizando el modelo de sustitución de nucleótidos de HKY + I (BIC); posteriormente editado para mejor presentación.

VII. Discusión de resultados

Se decidió utilizar muestras de secuencias de *T. vaporariorum* como un outgroup debido a dos razones principales: 1. Las secuencias cuentan con un tamaño adecuado para su uso en este análisis (Aproximadamente 600 bp) y esto es importante para estandarizar que todas las secuencias utilizadas tuvieran un tamaño similar, 2. Las secuencias fueron consideradas como una referencia con la suficiente distancia genética como para poder diferenciarse claramente dentro de los árboles filogenéticos.

En la base de los árboles se pueden observar las secuencias de *T. vaporariorum* utilizadas lo cual no es nada fuera de lo normal, pero vale la pena ser mencionado ya que ninguna de las muestras de referencia o las muestras analizadas presentaron una fuerte relación a estas.

Es necesario resaltar que a pesar de que los virus lleven un cultivo dentro del nombre de su variedad, como lo es en el “Tomato Yellow Leaf Curl Virus” (TYLCV), no significa que el virus infecte exclusivamente a dicha planta. Como se mencionó con anterioridad, existe una gran posibilidad de la aparición de infecciones mixtas (Harrison & Robinson, 1999), así como a través de los años en diferentes estudios se ha demostrado que los *Begomovirus* tienden a pasar por efectos de recombinación muy a menudo debido a su sistema de replicación en ciclo (Padidam et al. 1999; Pita, 2001; Martin, 2015) y esto debe ser tomado en cuenta al analizar los resultados presentes en los árboles filogenéticos.

De los resultados de los árboles filogenéticos se puede observar que la muestra K14 32572 se encuentra muy relacionada a las muestras de “Tomato Mosaic Havana Virus” (ToMHV), por lo que se puede decir que está muestra pertenece a un *Begomovirus* ToMHV. Considerando que este tipo de virus ha sido reportado previamente en Guatemala, este resultado ayuda a confirmar nuevamente lo previamente descrito por Nakhla en 2005 y la identificación de esta muestra.

Se puede observar que existe un bloque agrupado que inicia en la secuencia G14 1503 4 y termina en la muestra K11 LPV-19-21, este bloque se encuentra entre dos secuencias de referencia: “Tomato Leaf Curl Sinaloa Virus” (TLCSV) y TMHV. A partir de esto es posible asumir que las muestras se encuentran más relacionadas a TLCSV que a TMHV, a pesar de ver esta relación no es posible asumir que la secuencia pertenece a un virus TLCSV (valor de Bootstrap de 18), pero existe una alta probabilidad de que el virus perteneciente a esta muestra sea una variante relacionada a virus de tomate debido a un valor de bootstrap de 72 justo en la división entre TLCSV Y ToMHV.

En ambos árboles las muestras C16 33561 1 y G16 33561 2 se encuentran relacionadas a muestras de referencia de *Begomovirus* de Tomate y Chile pimiento, pero están más relacionados a la muestra de TLCSV que a “Pepper Huasteco Yellow Vein Virus” (PHYVV). Debido a esto se puede

asumir que existe una mayor posibilidad de que estos virus sean más una variedad de Tomate que de Chile pimiento, pero no es posible determinar exactamente su identificación.

Por otro lado, las muestras A24 LPV-18-220 y C11 LPV-19-302 aparecen altamente relacionadas a las muestras de referencias de PHYVV por lo que se puede asumir que estas muestras definitivamente son de PHYVV. Ambas muestras fueron provenientes de plantas de ejote francés provenientes del municipio de Salamá en el departamento de Baja Verapaz. Además de poder realizar la identificación exacta de estas muestras se encontró que existe un pequeño bloque desde M20 1175 hasta A09 LPV-18-413 que también está relacionado a PHYVV, pero no lo suficiente para determinar que dichas muestras son definitivamente de ese tipo de virus. Es posible asumir que a partir de la relación existente las muestras pertenezcan a una variedad de *Begomovirus* de Chile pimiento. También es importante mencionar que la muestra A09 LPV-18-413 proviene de las hojas de una planta de melón en el departamento de El Progreso

En la figura 5 se puede observar que desde la muestra G11 LPV-19-21 (sandía) hasta la muestra M14 32572 (falta procedencia) se encuentran bastante relacionados a los *Begomovirus* de cultivos de pepino y melón por lo que se puede asumir que existe una alta probabilidad de que estas sean una variedad o familiar cercano a virus como “Melon chlorotic leaf curl virus” (MCLCuV). También es importante mencionar que la muestra G11 LPC-1921 que proviene de diferentes moscas blanca encontradas en cultivos de sandía en el departamento de Santa Rosa, fue positiva para este virus. Pero el valor de Bootstrap (Valor de Bootstrap de 26) presente en el árbol no indica que no es una relación lo suficientemente cercana para determinar si es este virus, Considerando que Brown et al. en 2001 reportan un 80% de incidencia de MCLCuV en Guatemala, específicamente en Zacapa, de las muestras analizadas muy pocas se encuentran relacionadas a *Begomovirus* de melón o alguna otra cucurbitácea.

Las muestras G18 LPV 17 456 (Chile pimiento, proveniente de El progreso), M18 LPV 17 458 (Macuy, Provniente del El Progreso) y O16 1129 (Ayote, proveniente de Jutiapa) se encuentran muy relacionadas a “Bean Golden Mosaic Virus” (BGMV) por lo que existe una alta probabilidad de que sean un virus de este tipo o sean otra variedad de *Begomovirus* de frijol. La poca presencia de este virus dentro de las muestras es sorprendente considerando que existen reportes de la presencia de este virus dentro del territorio nacional incluso antes de los años 90 principalmente en los departamentos de Jalapa, Jutiapa y Escuintla (Faria et al., 1994), mientras que en este caso se encontró presencia de este tipo de virus en El progreso en dos cultivos diferentes.

En ambos árboles (Figura 5 y 6) se puede observar que existe un bloque de muestras que no se encuentran relacionadas a ninguna de las muestras de referencia utilizadas, debido a esto se puede asumir que estas muestras podrían pertenecer a *Begomovirus* que no han sido previamente reportados en Guatemala.

Finalmente, en el caso de los *Begomovirus* que no pudieron ser identificados es recomendable tomar las muestras y realizar un análisis separado en cada uno de los grupos con más variedades de virus que puedan estar relacionadas o no se hayan reportado previamente en Guatemala ya que éstos podrían probar la presencia de nuevas variantes en el país

En cuanto al uso del programa DNASTar y la aplicación MegAlign Pro, ambos demostraron tener una facilidad de uso para el usuario, así como menús y opciones intuitivas que dirigen al usuario durante el proceso sin necesidad de que éste posea amplio conocimiento previo del uso de otros programas de este tipo. El editor de árboles filogenéticos es una herramienta bastante útil que no todos los programas del mismo tipo poseen (Por ejemplo: Mega X o Mr.bayes; Kumar, 2018), las

posibilidades de edición del árbol son bastante diversas y permiten que los árboles puedan tener un diseño bastante profesional así como también uno bastante llamativo e interesante que puede ser utilizado para objetivos diferentes a publicaciones científicas donde es necesario atraer al lector con figuras y no con texto simple.

La interfaz puede ser un poco abrumadora al principio ya que la ventana principal se encuentra dividida en 3 o hasta 4 secciones diferentes (Figura 3) dependiendo de qué proceso se esté realizando. Se ha estudiado el uso de diseño minimalista, optimizado y centrado alrededor del usuario ya que esto tiende a que la información pueda fluir de un punto a otro de manera orgánica sin tener que indicarle explícitamente al usuario como hacerlo (Collen, 2002), el programa se beneficiaría bastante de poder utilizar este tipo de diseño en futuras versiones para evitar este tipo de problemas.

La instalación del programa probó ser un poco difícil al mismo tiempo, ya que el sitio no es muy amigable con el usuario y no provee opciones directas para descargar el programa sin comprar una licencia previamente y descargar los archivos de instalación directamente de la página del usuario. Esto tiende a complicar que personas se interesen en el programa ya que no pueden ni siquiera acceder a los programas previo a la obtención de una licencia. Este problema se evita en la guía de uso redactada (Ver anexo I) por medio de la instalación de una versión de prueba que permita acceder a todas las funciones al utilizar una licencia válida.

Problemas como estos son los que deberían de ser mencionados en guías de usuario ya que al no contar con la información muchas personas podrían evitar usar programas de bioinformática. Éstos suelen tener una imagen pública de ser “Complicados” o “Sólo para expertos”, cuando en realidad con un diseño más intuitivo y amigable hacia el usuario se podría extender el público objetivo hacia el cual estos programas están enfocados.

VIII. Conclusiones

Se logró identificar la presencia de “Tomato Mosaic Havana Virus” (TMHV) y “Pepper Huasteco Yellow Vein Virus” (PHYVV) dentro de las muestras analizadas, también se encontró una alta probabilidad de que exista la presencia de Tomato Leaf Curl Sinaloa Virus” (TLCSV), “Bean Golden Mosaic Virus” (BGMV) y Melon chlorotic leaf curl virus” (MCLF).

El hecho de que varias de las muestras no pudieron ser identificadas o relacionadas a alguna de las muestras de referencia significa que existe una gran posibilidad de que éstas sean nuevas variantes o cepas o especies de *Begomovirus* que no han sido reportados previamente en el territorio nacional, por lo que realizar un análisis más profundo de las mismas podría llevar a una mejora del manejo de las enfermedades causadas por *Begomovirus* en las plantaciones del país así como también una mejor base de datos para poder identificar los diferentes virus que afecten los cultivos de las personas.

Finalmente se creó una guía detallada para el uso del programa DNASTAR y la aplicación MegAlign Pro para su uso en los laboratorios de la Universidad de Valle de Guatemala que se puede encontrar en la sección de Anexos de este documento.

IX. Recomendaciones

Se recomienda realizar un análisis de haplotipos con las diferentes muestras, así como datos geográficos para determinar si existe una conexión entre las diferentes muestras y el efecto que la biogeografía del país tiene sobre la distribución de los mismo y la *Bemisia tabaci*.

También, con los resultados encontrados es posible que se puedan encontrar variedades nuevas de *Begomovirus* que no se encuentran reportadas en el país al compáralas con secuencias de variedades que tengan una alta posibilidad de entrar fácilmente al territorio nacional a través de importaciones o las aduanas.

Razlizar una evaluación de este manual como herramienta para usuarios de este programa, tanto principiantes como algunos avanzados.

X. Literatura citada

Barboza, N. M., E. Hernández, A. K. Inoue-Nagata, E. Moriones, & L. Hilje (2019). Achievements in the epidemiology of begomoviruses and their vector *Bemisia tabaci* in Costa Rica. *Revista de Biología Tropical*, 67(3).

Berning, D., H. Adams, H. Luc, & J. B. Gross (2019). In-Frame Indel Mutations in the Genome of the Blind Mexican Cavefish, *Astyanax mexicanus*. *Genome Biology and Evolution*, 11(9), 2563–2573. <https://doi.org/10.1093/gbe/evz180>

Brown, J. K. (2000). Molecular markers for the identification and global tracking of whitefly vector–Begomovirus complexes. *Virus Research*, 71(1), 233–260. [https://doi.org/10.1016/S0168-1702\(00\)00221-5](https://doi.org/10.1016/S0168-1702(00)00221-5)

Brown, J. K., A. M. Idris, D. Rogan, M. H. Hussein, & M. Palmieri (2001). Melon chlorotic leaf curl virus, a New Begomovirus Associated with *Bemisia tabaci* Infestations in Guatemala. *Plant disease*, 85(9), 1027. <https://doi.org/10.1094/PDIS.2001.85.9.1027C>

Burnham, K. P., & D. R., Anderson. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research*, 33(2), 261-304.

Collen, Mackenzie. (2002). The Need for a Design Lexicon: Examining Minimalist, Performance-centered, and User-centered Design. *Technical Communication*, Volume 49, Number 4, November 2002, pp. 405-410(6)

DNASTAR®. Version 12.0. DNASTAR. Madison, WI.

DNASTAR. (2020). 2019 Publication Review. Accedido en 26 de noviembre de 2020

Faria, J. C. de, R. Gilbertson, S. F. Hanson, F. Morales, P. Ahlquist, A. Charkowski & D. Maxwell (1994). Bean Golden Mosaic Geminivirus Type II Isolates from the Dominican Republic and Guatemala: Nucleotide Sequences, Infectious Pseudorecombinants, and Phylogenetic Relationships. *Phytopathology*, 84, 321–329. <https://doi.org/10.1094/Phyto-84-321>

- Fauquet, C. M. (2008). Taxonomy, Classification and Nomenclature of Viruses. In B. W. J. Mahy & M. H. V. B. T.-E. of V. (Third E. Van Regenmortel (Eds.) (pp. 9–23). Oxford: Academic Press. <https://doi.org/https://doi.org/10.1016/B978-012374410-4.00509-4>
- Gafni, & B. L. Epel. (2002). The role of host and viral proteins in intra- and inter-cellular trafficking of geminiviruses. *Physiological and Molecular Plant Pathology*, 60(5), 231–241. <https://doi.org/https://doi.org/10.1006/pmpp.2002.0402>
- Ghanim, M., S. Morin, M. Zeidan, & H. Czosnek, (1998). Evidence for Transovarial Transmission of Tomato Yellow Leaf Curl Virus by Its Vector, the Whitefly Bemisia tabaci. *Virology*, 240(2), 295–303. <https://doi.org/https://doi.org/10.1006/viro.1997.8937>
- Harrison, B. (1985). Advances in Geminivirus Research. *Annual Review of Phytopathology*, 23(1), 55–82. <https://doi.org/10.1146/annurev.py.23.090185.000415>
- Harrison, B., & D. Robinson, (1999). Natural genomic and antigenic variation in whitefly-transmitted geminiviruses (begomoviruses). *Annual Review of Phytopathology*, 37(1), 369–398. <https://doi.org/10.1146/annurev.phyto.37.1.369>
- Inoue-Nagata, A., M. Lima, & R. Gilbertson, (2016). A review of geminivirus diseases in vegetables and other crops in Brazil: current status and approaches for management. *Horticultura Brasileira*. scielo.
- Kumar, S., G. Stecher, M. Li, C. Knyaz, & K. Tamura (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, 35(6), 1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Lapidot, M., M. Friedmann, M. Pilowsky, R. Ben-Joseph & S. Cohen (2001). Effect of Host Plant Resistance to Tomato yellow leaf curl virus (TYLCV) on Virus Acquisition and Transmission by Its Whitefly Vector. *Phytopathology*®, 91(12), 1209–1213. <https://doi.org/10.1094/PHYTO.2001.91.12.1209>
- Lazarowitz, S., & R. Shepherd (1992). Geminiviruses: Genome structure and gene function. *Critical Reviews in Plant Sciences*, 11(4), 327–349. <https://doi.org/10.1080/07352689209382350>
- Lemmon, A. R., & E. C. Moriarty (2004). The Importance of Proper Model Assumption in Bayesian Phylogenetics. *Systematic Biology*, 53(2), 278–298. <https://doi.org/10.1080/10635150490423520>
- Martin D. P. (2015) ‘RDP4: Detection and Analysis of Recombination Patterns in Virus Genomes’, *Virus Evolution*, 1: vev003
- MAGA (Ministerio de Agricultura, Ganadería y Alimentación: MAGA Guatemala, Centro América). 2016. El Agro en cifras. Disponible en: <https://precios.maga.gob.gt/archivos/agro-en-cifras/individuales/Melon%20Agro%20en%20Cifras%202016.pdf>
- MAGA (Ministerio de Agricultura, Ganadería y Alimentación: MAGA Guatemala, Centro América). 2015. Melón Agro en cifras. Disponible en: https://www.mineco.gob.gt/sites/default/files/Integracion%20y%20comercio%20exterior/el_agro_en_cifras_2015.pdf
- MegAlign Pro®. Version 12.0. DNASTAR. Madison, WI.

- Mehta, P., J. A. Wyman, M. Nakhla, & D. Maxwell (1994). Transmission of Tomato Yellow Leaf Curl Geminiviruses by *Bemisia tabaci* (Homoptera: Aleyrodidae). *Journal of Economic Entomology*, 87(5), 1291–1297. <https://doi.org/10.1093/jee/87.5.1291>
- Mutebi, J., H. R. Hughes, K. L. Burkhalter, L. Kothera, C. Vasquez, & J. L. Kenney (2018). Zika Virus MB16-23 in Mosquitoes, Miami-Dade County, Florida, USA, 2016. *Emerging Infectious Diseases*, 24(4), 808-810. <https://dx.doi.org/10.3201/eid2404.171919>.
- Nakhla, M., 2005. Molecular Characterization of Tomato-infecting begomoviruses in Central America and development of DNA-based detection methods. *Acta Hort.* 695:227-288.
- Nakhla, M. & D. Maxwell (1998). Epidemiology and management of tomato yellow leaf curl disease. In *Plant Virus Disease Control*. 565-583
- Padidam M., S. Sawyer & C. M. Fauquet (1999) ‘Possible Emergence of New Geminiviruses by Frequent Recombination’, *Virology*, 265: 218–24
- Pita J. S. (2001) ‘Recombination, Pseudorecombination and Synergism of Geminiviruses are Determinant Keys to the Epidemic of Severe Cassava Mosaic Disease in Uganda’, *Journal of General Virology*, 82: 655–65
- Posada, D. (2008). jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*, 25, 1253–1256. <https://doi.org/10.1093/molbev/msn083>
- Roshan, P., A. Kulshreshtha & V. Hallan (2017) Genome Organization of Begomoviruses. In: Saxena S., Tiwari A. (eds) *Begomoviruses: Occurrence and Management in Asia and Africa*. Springer, Singapore. https://doi.org/10.1007/978-981-10-5984-1_2
- Teni, Rudy. (2004). Experiencias en la producción de híbridos de tomate (*Lycopersicon esculentum* Mill) tolerantes a virosis transmitida por mosca blanca. Universidad de San Carlos de Guatemala. Recuperado de http://biblioteca.usac.edu.gt/tesis/01/01_2068.pdf
- Serra. 2006. Desarrollo epidemiológico de la virosis en la Comunidad Valenciana: el caso del “virus de la cuchara” del tomate. *Conselleria de Medio Ambiente, agricultura y agua de la Region de Murcia*. Servicio de Sanidad Vegetal. España. P 49-53.
- Scotta, R., D. Sánchez, & C. Arregui (2015). Determinación de las pérdidas causadas por la mosca blanca de los invernaderos (*Trialeurodes vaporariorum*) en cultivos de tomate bajo invernadero. *FAVE Sección Ciencias Agrarias*, 13. <https://doi.org/10.14409/fa.v13i1/2.4962>
- Yadava, P., G. Suyal, & S. Mukherjee (2010). Begomovirus DNA replication and pathogenicity. *Current Science*, 98, 360–368.

XI. ANEXOS

ANEXO I. Manual para el uso del software DNASTar y la aplicación MegAlign Pro



Excelencia que trasciende

DELVALLE
GRUPO EDUCATIVO

**Guía para generar árboles filogenéticos
utilizando el programa DNASTAR y la aplicación
MegAlignPro**

INDICE

Instalación del programa.....	1
Cargar las secuencias en el programa	8
Alineamiento de secuencias	11
Creación de árboles filogenéticos.....	12
Referencias.....	16

Introducción

DNASTar ha sido un software de bioinformática muy utilizado en el campo de la investigación durante los últimos 35 años, solamente en el año 2019 se reportaron 3,370 publicaciones que han utilizado el programa en diferentes tipos de estudios (DNASTar, 2020).

Crear árboles filogenéticos con este programa es bastante intuitivo y fácil de entender, los mayores problemas que podría encontrar al tratar de seguir la guía son durante la instalación o sentirse un poco perdido al momento de observar la página principal. Pero, si sigue los pasos en orden y de manera correcta no debería de tener ninguna complicación utilizándolo y se familiarizará con el programa rápidamente, idealmente siendo capaz de realizar el proceso sin necesidad de esta guía en el futuro.

¿Cómo instalar el programa?

Pasos:

1. Diríjase a la página utilizando su explorador de preferencia, esto debe llevarlo a la prima principal de DNASTAR.

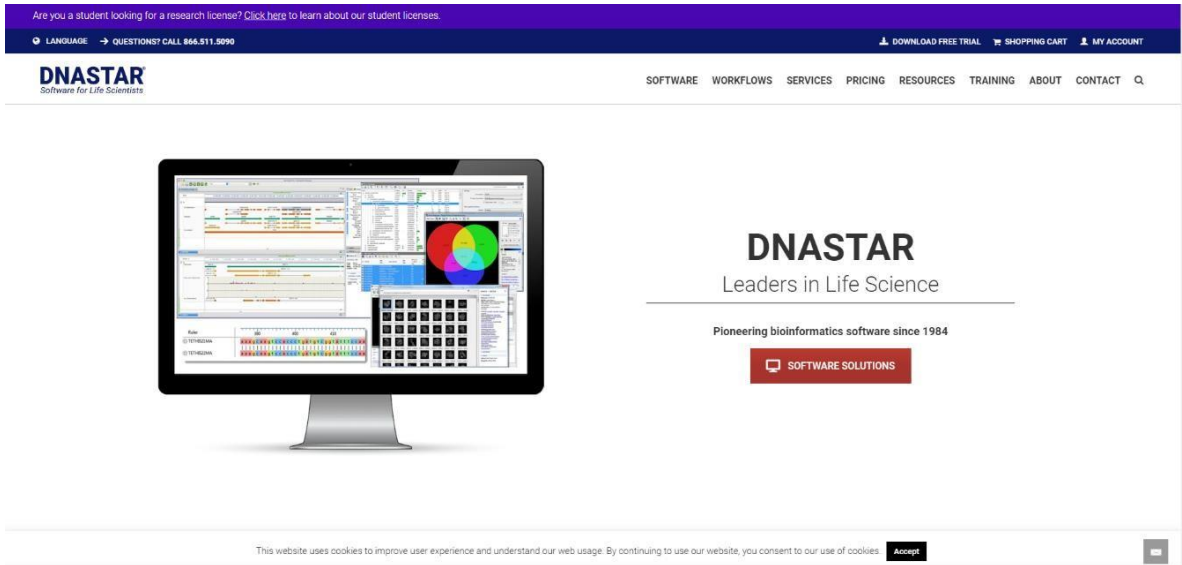


Figura 1. Página principal del sitio web de DNASTAR

2. Seleccione la opción de Software y haga click sobre “DNASTAR Lasergene”

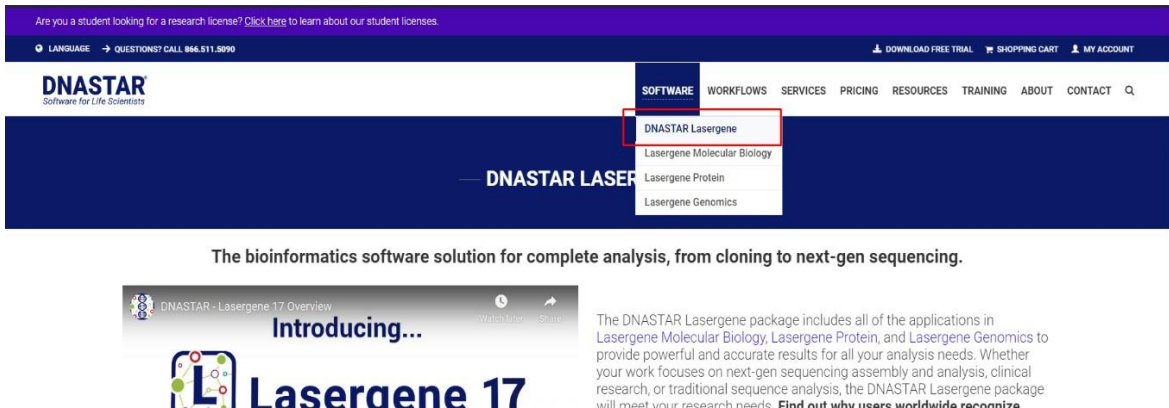


Figura 2. Ubicación de la opción “DNASTAR Lasergene” bajo el menú de Software

3. Al final de la página debe encontrar una opción de “14 day free trial”, haga click en el botón de “Free trial download”.

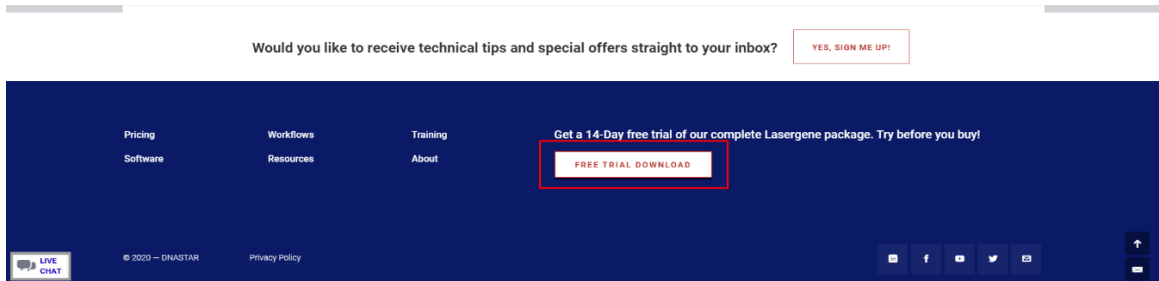


Figura 3. Ubicación de la opción de “Free trial” al final de la página

Nota: La opción de prueba gratuita es la manera más fácil de descargar la aplicación sin tener que preocuparse de descargar todas las diferentes opciones por separado, cuando instale el programa utilizaremos la licencia completa de la universidad que dejará que use el programa completo sin preocuparse de los 14 días de prueba.

4. Después de hacer click en el botón debería de llegar a una página que le solicita llenar un formulario. Coloque la información necesaria, seleccione el sistema operativo de la computadora en la cual piensa utilizar el programa y haga click sobre “acces my free trial”

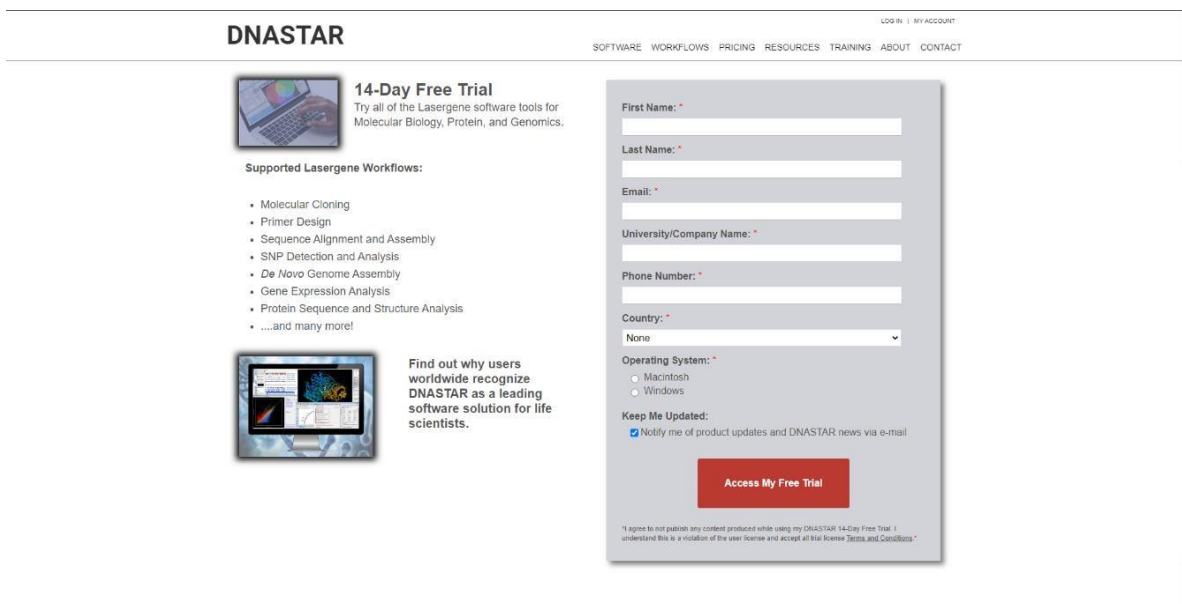


Figura 4. Formulario para la descarga del programa

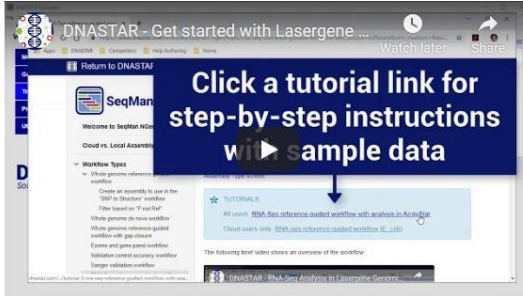
5. Después de llenar el formulario llegará a una nueva página donde encontrará un botón para descargar el programa, haga click sobre él.

Thank You!

Thank you for your request for a free trial. To start downloading, click the link below. A product key for authorizing your software has been sent from DNASTAR to the email address you provided. (You may need to check your spam filter to ensure you receive the message).

[Download for Windows](#)

Watch the video below to see how to get started with your DNASTAR Lasergene trial using DNASTAR Navigator!



If you experience slow download speeds, you may want to try downloading the installer from one of our global host sites. Use this option to select another site: Brazil

Figura 5. Página con la opción de descarga

- Después de hacer click iniciará la descarga de un archivo ejecutable (.exe)



Figura 6. Archivo de descarga

- Una vez terminada la descarga haga click sobre el archivo para iniciar el instalador. De ser necesario haga click en “Si” en la ventana que se abrirá si necesita permisos de administrador.
- Al abrir la ventana del instalador haga click en la opción “Next”

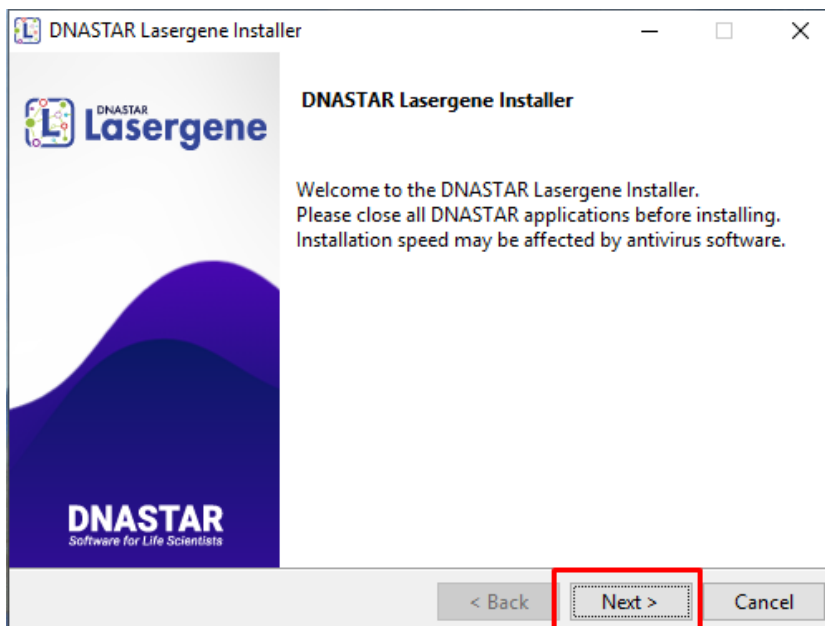


Figura 7. Venta inicial del instalador

- Seleccione la opción para aceptar los acuerdos de la licencia y haga click en “Next”

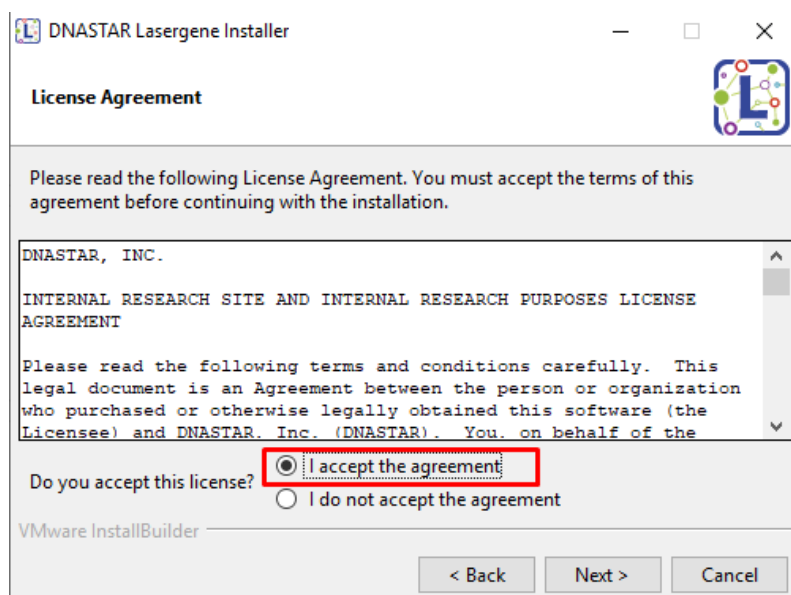


Figura 8. Acuerdo de uso de Licencia

10. En la siguiente ventana tendrá la opción de cambiar la ubicación del directorio donde desea instalar el programa, si no desea cambiar el directorio puede dejar la opción predeterminada. De lo contrario haga click sobre el folder y seleccione el directorio donde desea instalar el programa.

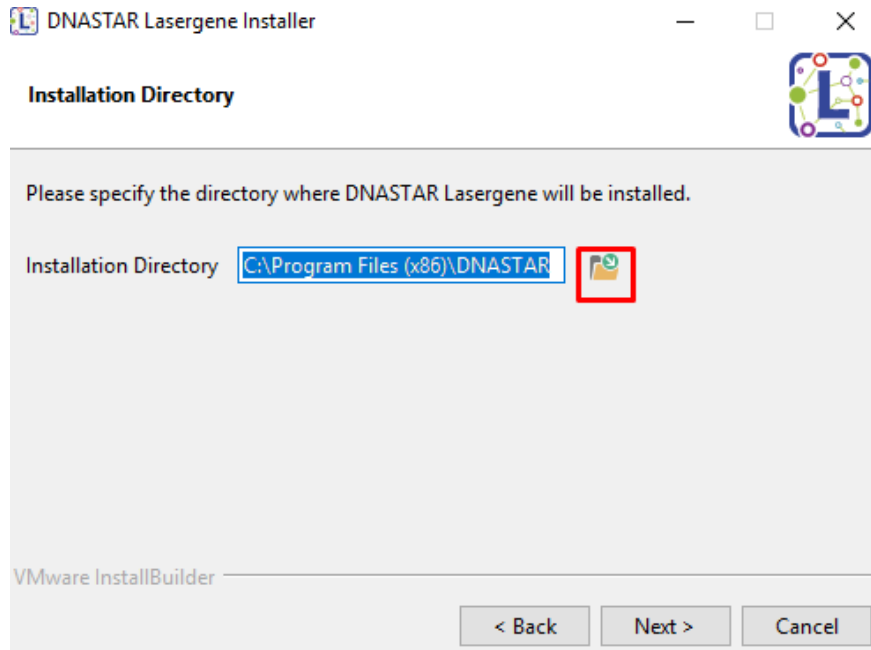


Figura 9. Ubicación del botón para cambiar directorio de instalación

11. Al hacer click en “Next” se abrirá una nueva ventana indicándole que todo está listo para la instalación, Haga click en “Next “nuevamente y el programa se instalará en su equipo.

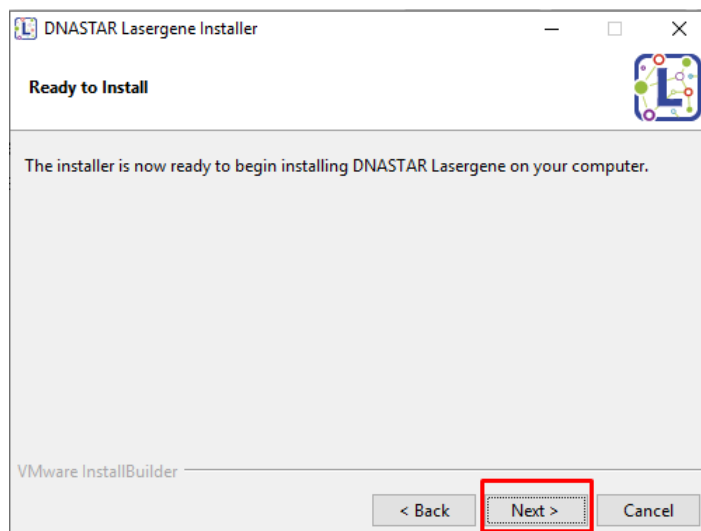


Figura 10. Ventana de instalación

12. Mientras se instala el programa aparecerá una nueva ventana solicitando la clave de la licencia, escriba en la nueva ventana la licencia proveída por la universidad

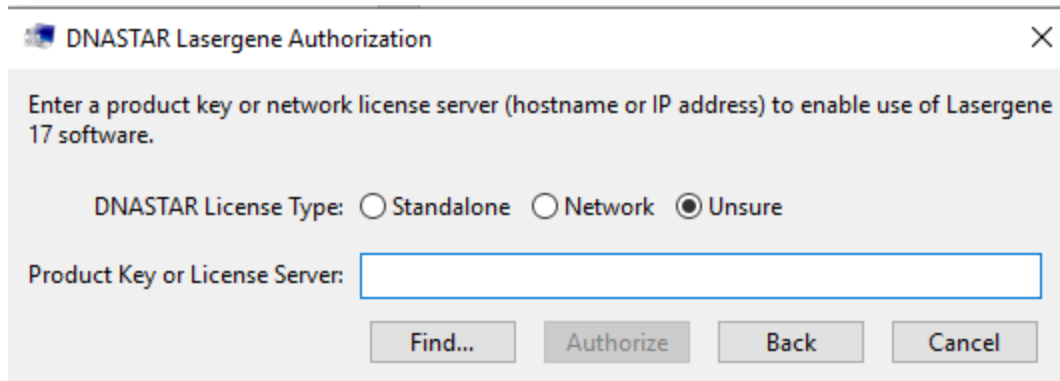


Figura 11. Nueva ventana solicitando la licencia

13. Después de Escribir la licencia Haga click en "Authorize" y aparecerá una nueva ventana indicándole que su licencia se está autorizando. Al autorizarse otra nueva ventana aparecerá indicando que la autorización fue correcta.

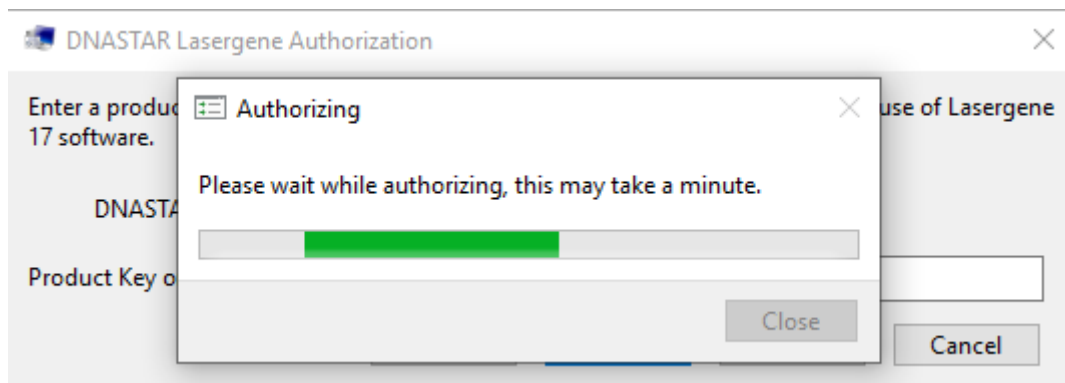


Figura 12. Nueva ventana de autorización

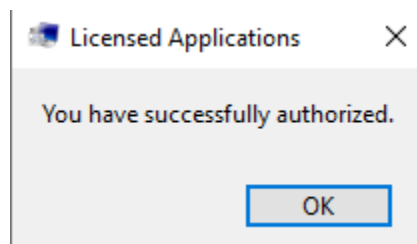


Figura 13. Autorización exitosa

14. Después de hacer click en “OK” el programa debería de haber terminado de instalarse, de lo contrario espere a que la instalación termine.

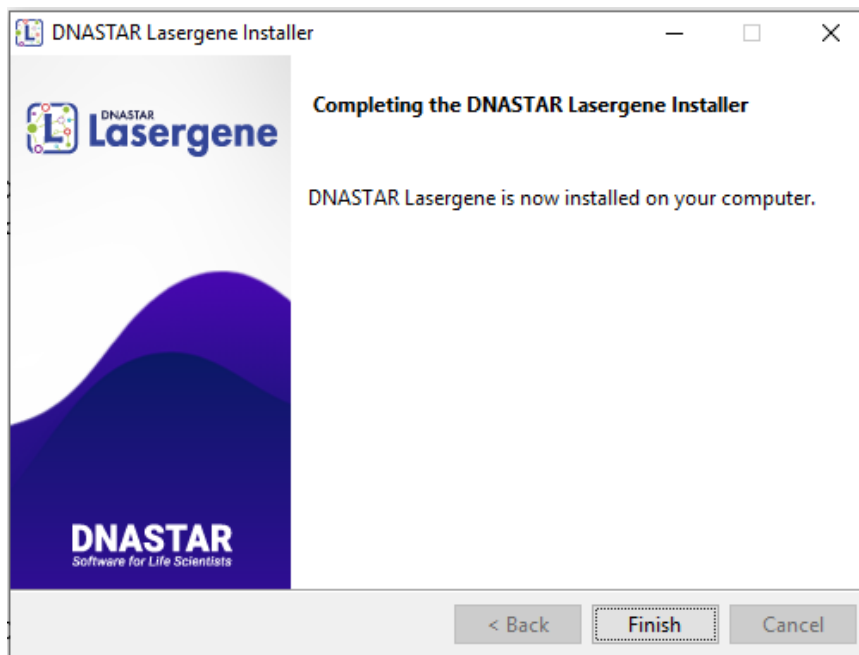


Figura 14. Instalación completada

15. Al hacer click en “Finish” el programa se abrirá en una nueva ventana y estará listo para ser utilizado

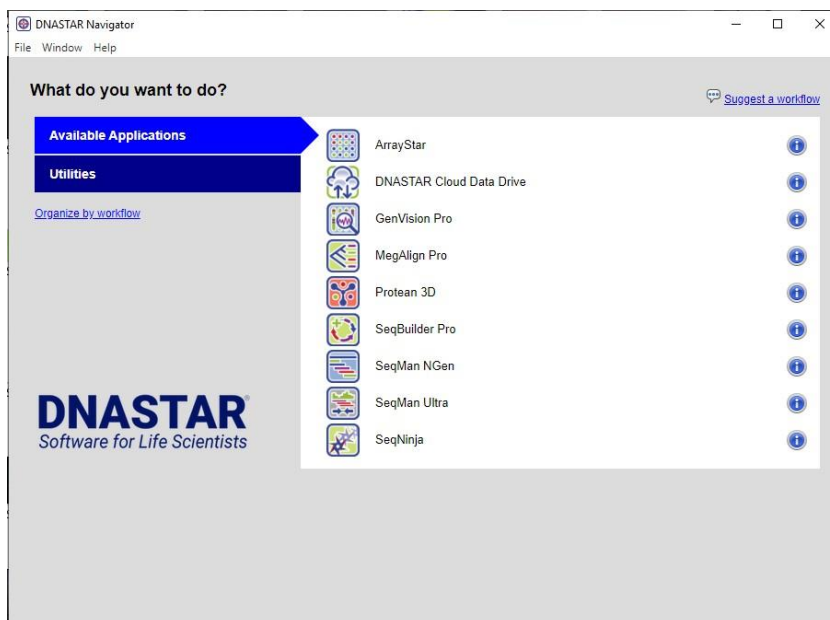


Figura 15. Nueva ventana del programa

¿Cómo utilizar DNASTAR para la creación de árboles filogenéticos?

Secuencias:

Para utilizar el programa es necesario que las secuencias a utilizar se encuentren en formato **FASTA**. Si sus secuencias están en un formato diferente por favor conviértalas a formato **FASTA** antes de utilizarlas.

Una secuencia regularmente se encontrará en el siguiente formato si la obtiene directamente de una compañía de secuenciación o genbank:

```
1 aagctgggag tatcgattaa gaacttttat ttacatttat cagttttgct attcattctt
61 atgtttcttt atttcttgca aaattgtagt atcactacc ttaatccta aaatatatgt
121 ctactttaa gattttaatt gtttgggtgc gaaattttag ctgtgaaaat gagccgcact
181 tatgaattcg gtataaattg gtttcatcct ggactttcaa gggagactgc tgaaaaattg
241 ttgttaaatg aatgcactga ggatggatct ttttaatac gggaatgcaa ctcggcagta
301 ggcgattttg ttcttacct cttacaccag aataacgtgt gccattatca gatccgaagg
361 catggtgggg atgccttttt tcaatagat gataaagtaa agattttaca tggcctagaa
```

Muchos genbanks ofrecen la opción de descargar un archivo en formato fasta con la secuencia por lo que debería de ser fácil de obtener si utiliza este tipo de secuencias. Un archivo en formato fasta tiene el siguiente formato:

```
>GAMC01006777.1:1-420 TSA: Ceratitis capitata comp60376_c0_seq1 mRNA sequence
AAGCTGGGAGTATCGATTAAGAACTTTTATTTACATTTATCAGTTTTGCTATTCTTATGTTTCT
TT
ATTTCTTGCAAATTGTAGTATCACTACCCTTAATCCTTAAAATATATGTCTTACTTTAAGATTTTAA
TT
GTTTGGTGTCGAAATTTTAGCTGTGAAAATGAGCCGCACTTATGAATTCGGTATAAATTGGTTTCA
TCCT
GGACTTTCAAGGGAGACTGCTGAAAATTGTTGTTAAATGAATGCACTGAGGATGGATCTTTTTTA
ATAC
GGGAATGCAACTCGGCAGTAGGCGATTTTGTCTTACTCTTACACCAGAATAACGTGTGCCATT
ATCA
GATCCGAAGGCATGGTGGGGATGCCTTTTTTTCAATAGATGATAAAGTAAAGATTTTACATGGCCT
AGAA
```

De ser necesario puede utilizar programas como Mega X, Clustal w o Bioedit para poder transformar sus archivos simplemente guardándolos en formato fasta después de abrirlos en el programa.

1. En la ventana principal de DNASTAR seleccione la aplicación “MegAlign Pro”

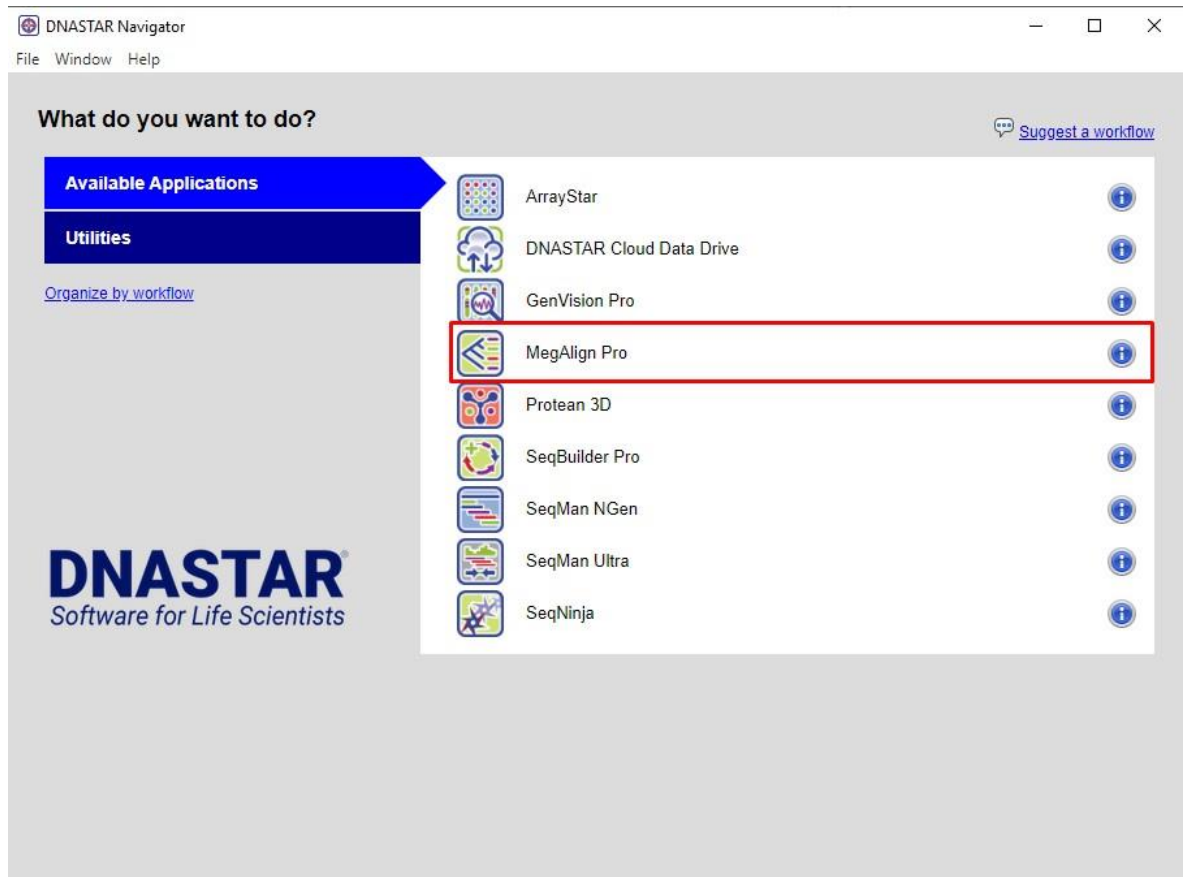


Figura 16. Ubicación de la aplicación MegAlign Pro en la ventana principal de DNASTAR

Nota: como podrá ver en la ventana principal hay una gran variedad de aplicaciones que pueden ser utilizadas para diversos fines como: Diseño de primers para PCR, Predicción de estructura de proteínas, Análisis de genoma completo, ensamble de genoma *De novo*, etc. En este caso solo nos interesa utilizar la aplicación MegAlign para alinear secuencias y generar árboles filogenéticos, pero si lo desea puede explorar el resto de las aplicaciones y ver la utilidad que este programa posee para su uso en la investigación.

2. Después de hacer click se iniciará la aplicación en una nueva ventana

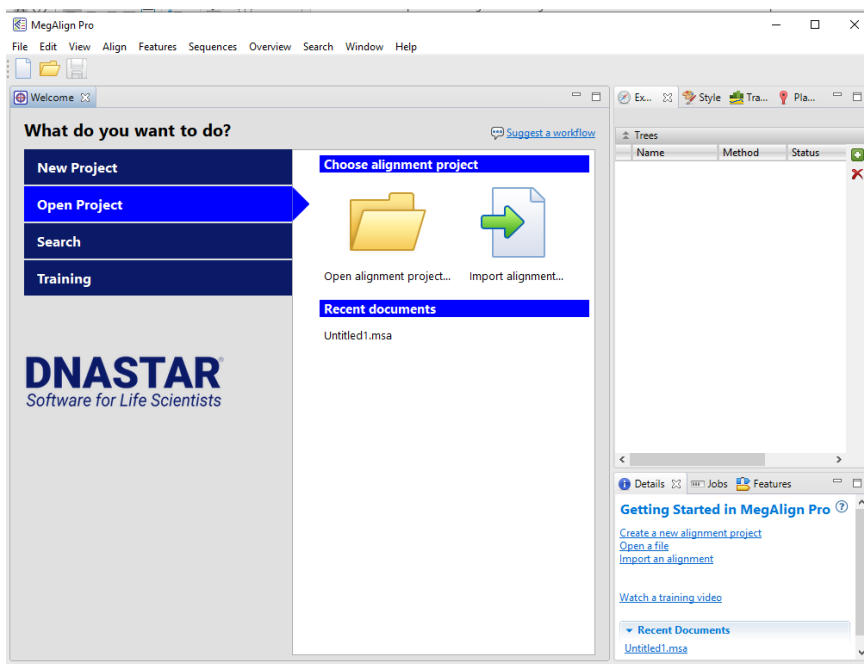


Figura 17. Ventana principal de MegAlign Pro

3. Haga click en la opción "New project" y seleccione "New blank alignment project"

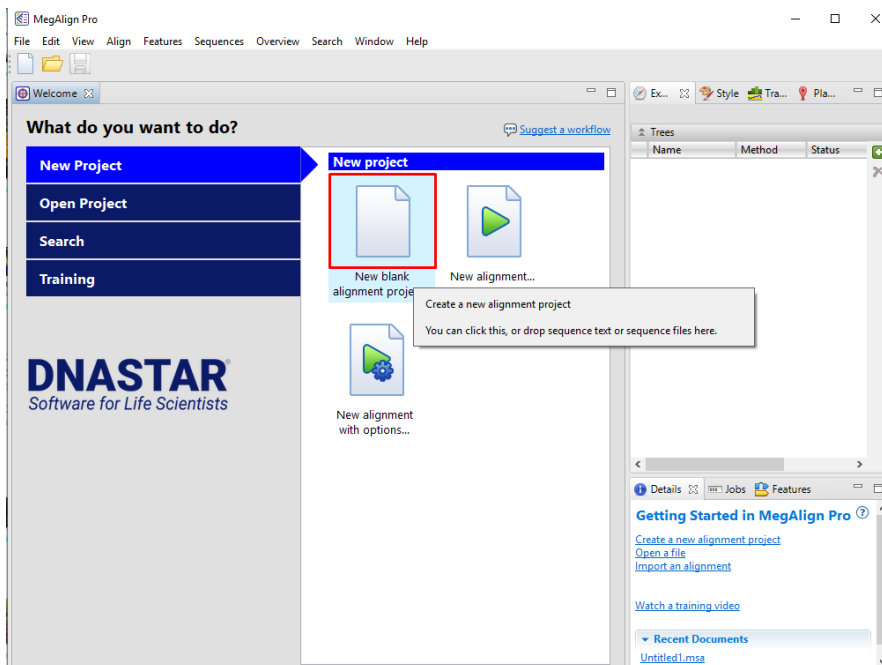


Figura 18. Crear un nuevo proyecto

- Esto creará un nuevo proyecto, al abrirse puede hacer click en el botón verde en la barra superior para agregar sus secuencias al proyecto. Después de hacer click sobre el botón se abrirá una nueva ventana donde puede navegar hasta la ubicación del archivo con sus secuencias.

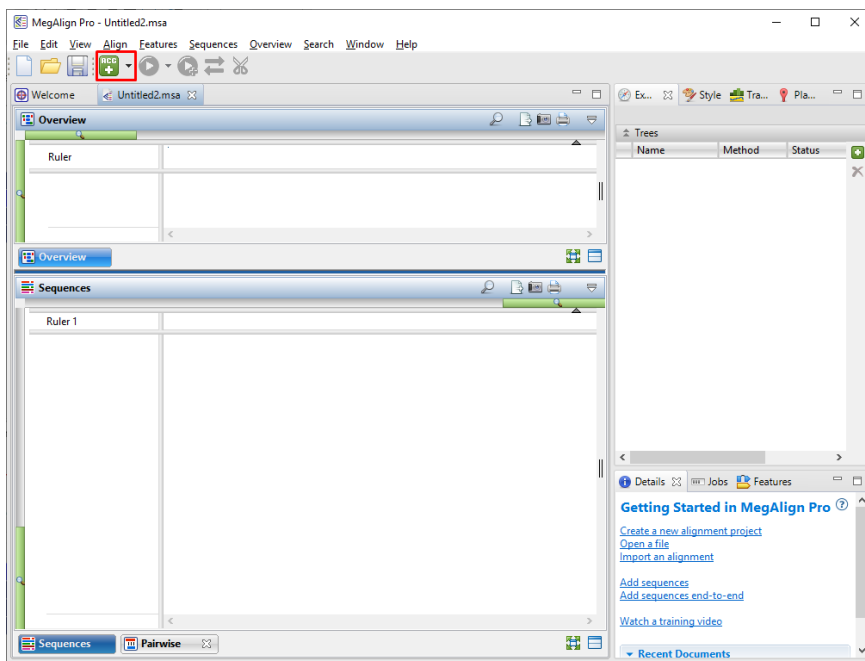


Figura 19. Ubicación del botón para añadir sus secuencias

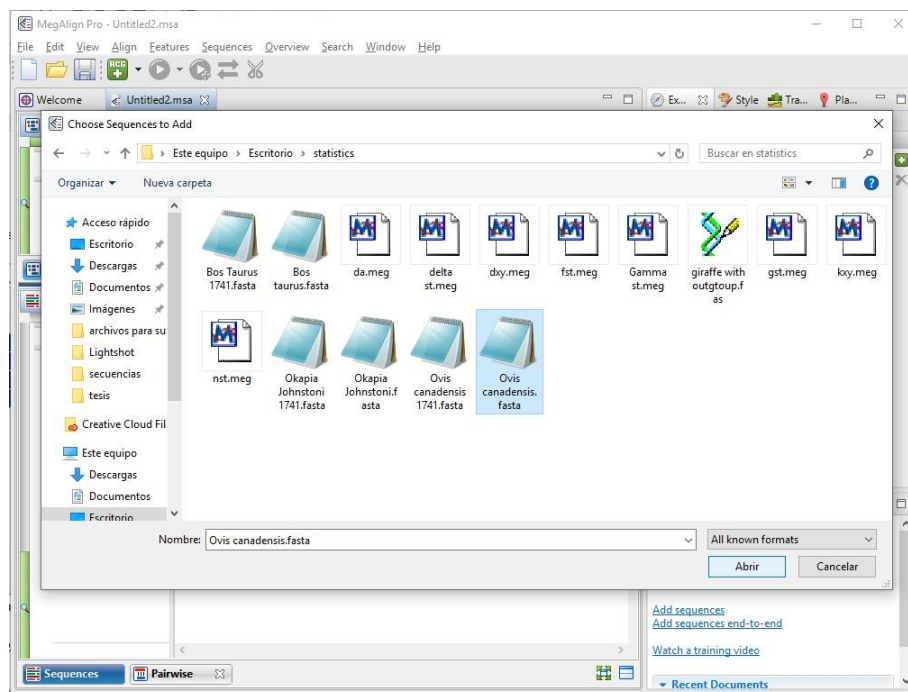


Figura 20. Ventana para seleccionar el archivo con sus secuencias

5. Se abrirá un cuadro de dialogo donde puede seleccionar si sus secuencias son ADN, ARN o Proteínas. Elija la opción que corresponda a sus secuencias

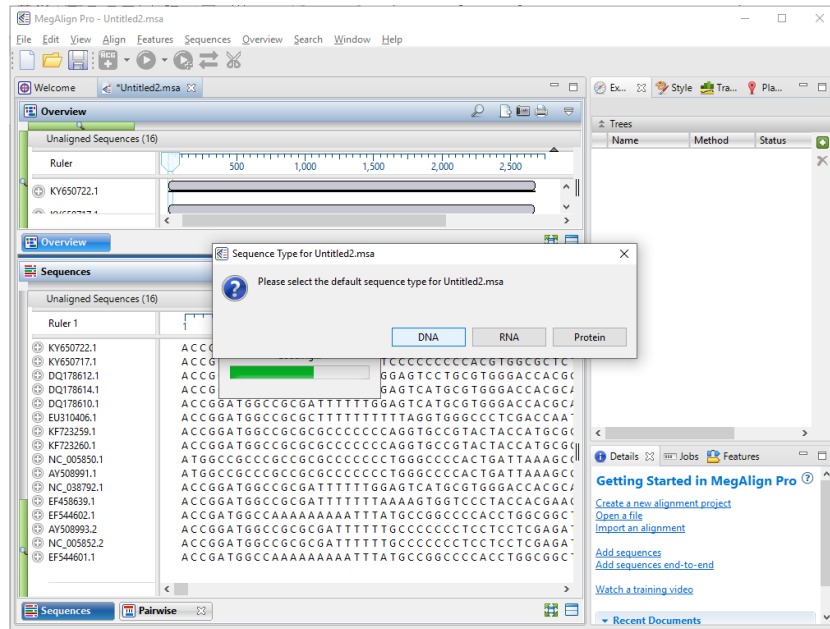


Figura 21. Cuadro de dialogo

6. Después de cargar sus secuencias puede hacer click en el nuevo botón verde que aparecerá en la barra superior para realizar el alineamiento que desee utilizar.

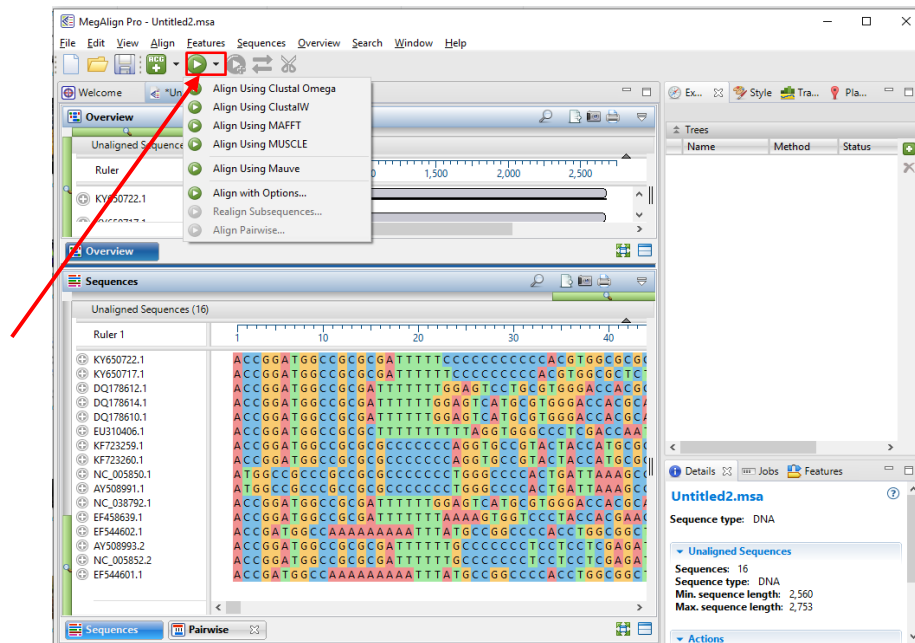


Figura 22. Boton para alinear secuencias.

7. Después de alinear sus secuencias, en la barra inferior aparecerá una nueva opción con el nombre "Tree" al seleccionarla puede elegir si realizar in árbol de "Nieghbor-joining" o "Maximun Likelihood".

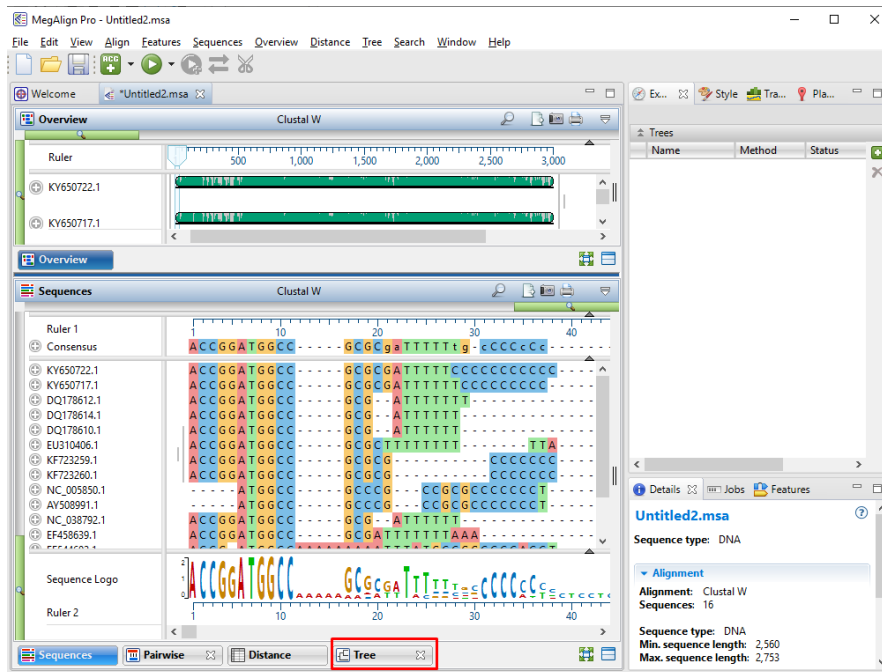


Figura 23. Ubicación de la opción "Tree" en la barra inferior

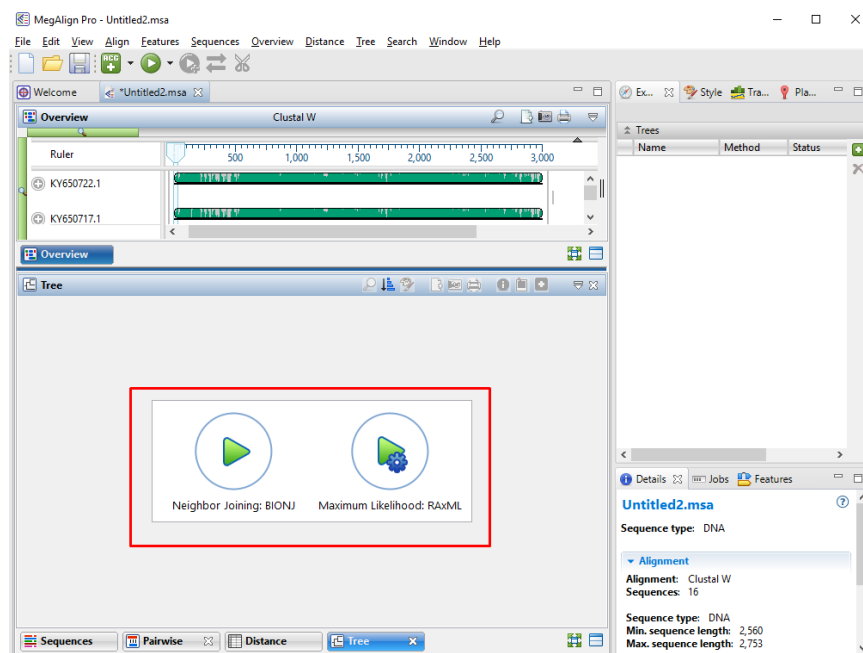


Figura 24. Opciones para crear un árbol

Árbol de “Neighbor-joining”

8. Seleccione la opción de “Neighbor Joining”

8.1. Se abrirá un cuadro de dialogo donde puede colocar un nombre para el árbol que se va a generar. Después de nombrar su árbol haga click en “OK”

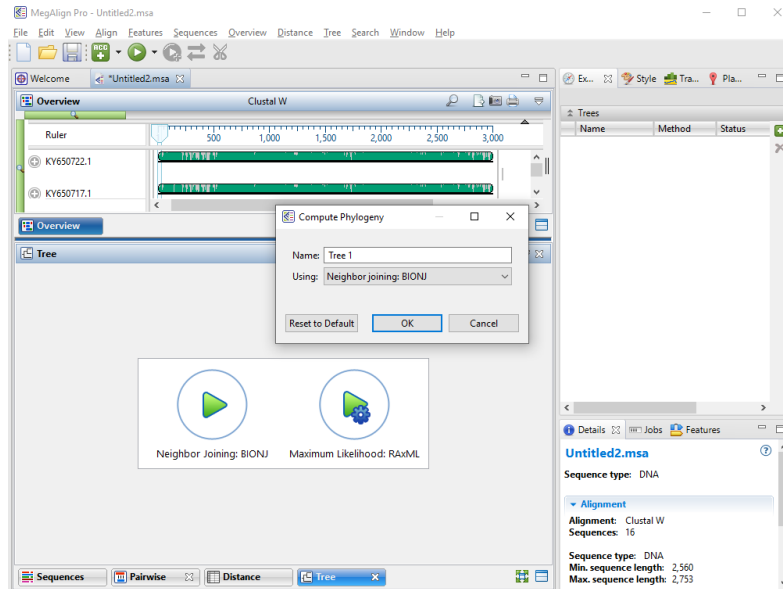


Figura 25. Cuadro de dialogo para nombrar su árbol

8.2. Después de hacer click en “OK” el programa realizará el análisis para generar su árbol y después de terminar obtendrá su árbol en la ventana “Tree”.

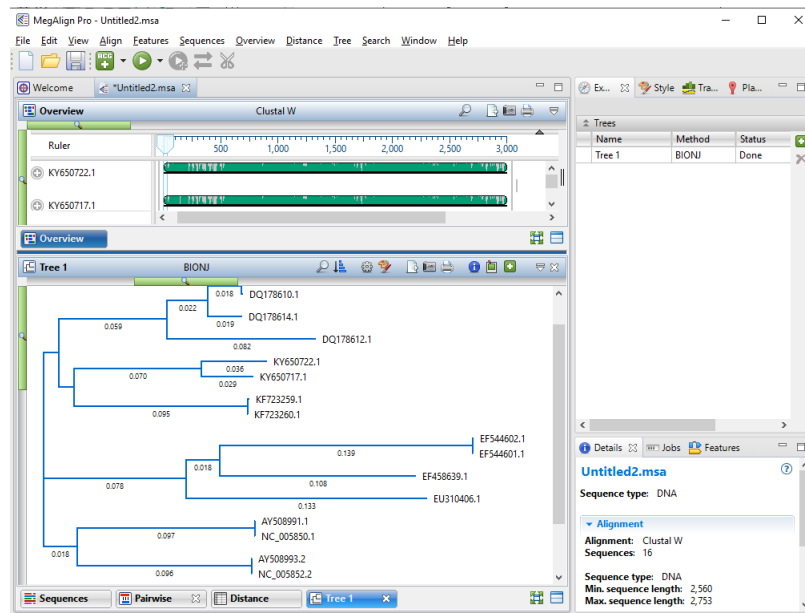


Figura 26. Árbol generado

- 8.3. Puede hacer click en la opción “Style” en la ventana derecha para modificar su árbol o puede hacer click en el icono de cámara en la barra sobre su árbol para guardarlo como una imagen

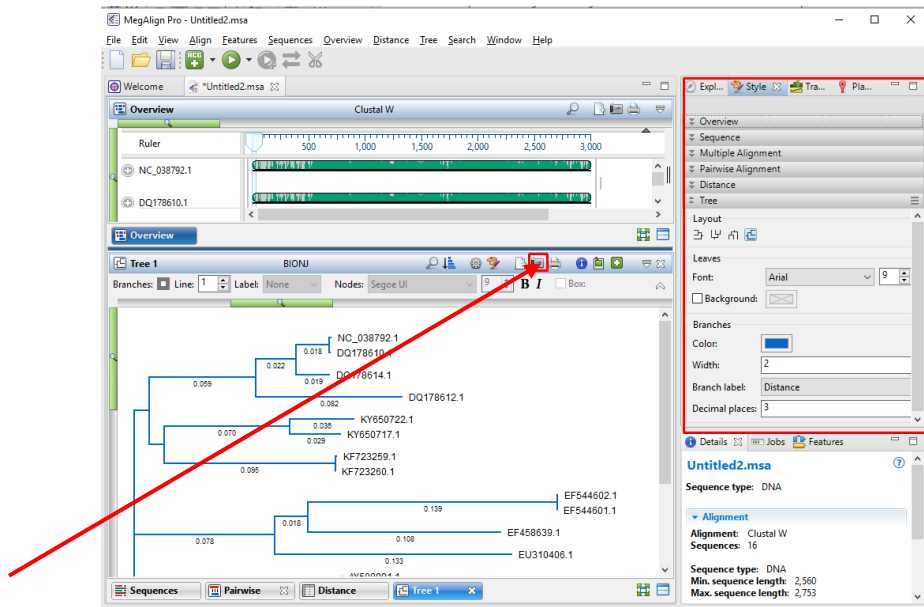


Figura 25. Opciones de estilo y opción para guardad su árbol.

Árbol de Maximun likelihood

9. Seleccione la opción “Maximum likelihood”
9.1. Se abrirá un nuevo cuadro de dialogo donde puede elegir el nombre de su árbol, así como ajustar las opciones de su análisis.

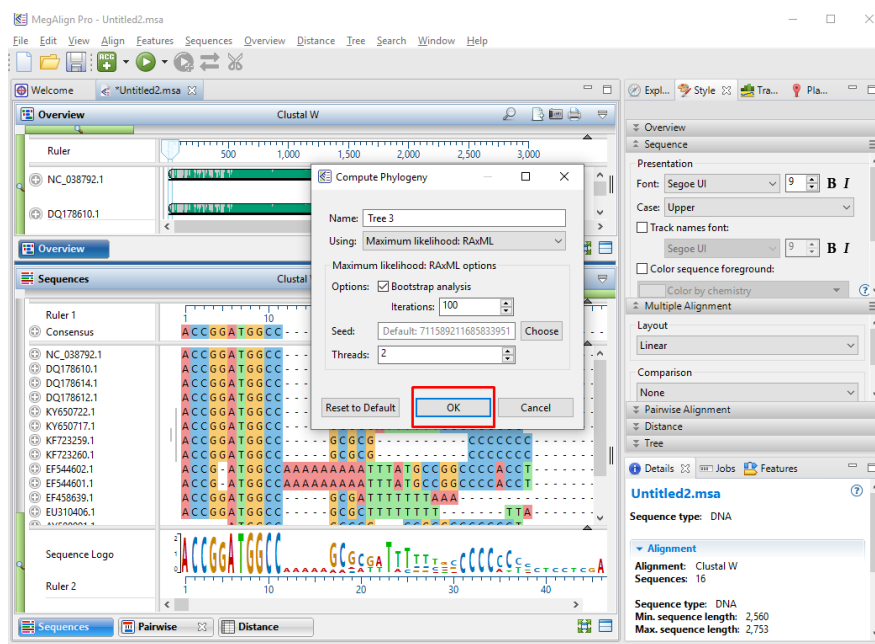


Figura 26. Cuadro de dialogo con las opciones para el análisis de máxima verosimilitud.

Nota: La opción “Threads” significa la cantidad de recursos de su computadora que se utilizarán para el análisis. Si no sabe la cantidad optima que su computadora permite no cambie el valor predeterminado.

- 9.2. Después de hacer click en “OK” el programa realizará el análisis y generará el árbol. Una vez terminado el análisis puede utilizar la opción de “Style” para editar su árbol o guardarlo utilizando el botón de cámara en la barra sobre su árbol.

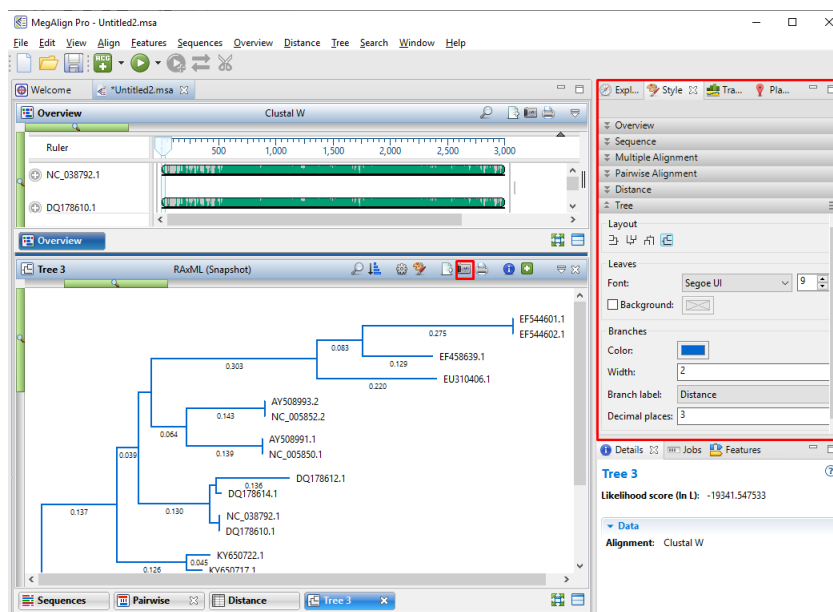


Figura 27. opciones de estilo y para guardar du árbol.

Nota: Este análisis tiende a ser más tardado y el tiempo que tomará depende de las especificaciones de su computadora.

Referencias

DNASTAR. (2020). 2019 Publication Review. Accedido en 26 de noviembre de 2020

DNASTAR®. Version 12.0. DNASTAR. Madison, WI.

MegAlign Pro®. Version 12.0. DNASTAR. Madison, WI.