

# Maestría en Bioinformática y Biología Computacional



**Ensamblaje del genoma de *Ganoderma australe* e  
identificación de Proteínas Fúngicas  
Inmunomoduladoras (FIPs) putativas y enzimas  
asociadas a la ruta del ácido ganodérico**

**DANIEL AGUDELO VALENCIA**

**Director: Jhon Fredy Betancur PhD.**

**Codirector: Germán López Gartner PhD.**

**Tutora: Andrea González Muñoz MSc.**

**UNIVERSIDAD DE MANIZALES  
MAESTRÍA EN BIOINFOMÁTICA Y BIOLOGÍA  
COMPUTACIONAL  
MANIZALES**

**2017**

# Tabla de contenido

Agradecimientos.....	6
Resumen .....	7
Abstract .....	8
Lista de tablas.....	9
Lista de figuras.....	10
Lista de anexos .....	11
Lista de abreviaturas.....	13
1. Introducción .....	14
1.1. Campo temático.....	14
1.2. Planteamiento del problema .....	15
1.3. Justificación.....	16
1.4. Objetivos.....	17
1.4.1. Objetivo general.....	17
1.4.2. Objetivos específicos .....	17
2. Referente teórico y antecedentes.....	18
2.1. Importancia biológica de los hongos .....	18
2.1.1. Propiedades biológicas de los hongos.....	18
2.2. Género <i>Ganoderma</i> .....	20
2.3. Ácidos Ganodéricos.....	21
2.4. Proteínas Fúngicas Inmunomoduladoras (FIPs).....	23
2.4.1. Estructura Molecular.....	24
2.4.2. Aplicaciones y prospectos .....	25
3. Materiales y métodos .....	27
3.1. Material Biológico.....	27
3.1.1. Cultivo del Hongo.....	27
3.1.2. Extracción y cuantificación de ADN genómico .....	27
3.1.3. Identificación molecular de <i>G. australe</i> .....	27
3.1.4. Secuenciación genómica.....	28
3.2. Ensamblaje y anotación del genoma .....	28
3.3. Anotación de proteínas FIPs.....	30

4. Resultados y Discusiones .....	32
4.1. Obtención del ADN y secuenciación del genoma .....	32
4.1.1. Identificación molecular de <i>G. australe</i> .....	33
4.1.2. Ensamblaje del genoma .....	34
4.2. Anotación del Ensamblaje .....	38
4.2.1. Anotación del ensamblaje con Maker2.....	38
4.2.2. Evaluación de la ruta metabólica del ácido ganodérico.....	40
4.3. Análisis de las proteínas FIPs .....	43
5. Conclusiones generales .....	50
5.1. Contribuciones de la Tesis.....	50
5.2. Impactos Potenciales de la Tesis .....	50
5.3. Recomendaciones y trabajos futuros .....	51
6. Referencias bibliográficas .....	52
ANEXOS.....	67
Anexo I.....	67
A. Resultado de la secuenciación del genoma de <i>G. australe</i> por la tecnología de secuenciación Sequel de PacBio. ....	67
B. Comandos utilizados para el ensamblaje del genoma de <i>G. australe</i> utilizando la herramienta Canu. ....	68
Anexo II. Análisis del ensamblaje del genoma de <i>G. australe</i> .....	69
A. Comparación genómica utilizando la herramienta Mummer de los genomas de <i>G. australe</i> vs <i>G. lucidum</i> .....	69
B. Comparación genómica utilizando la herramienta Mummer de los genomas de <i>G. australe</i> vs <i>G. boninense</i> .....	70
C. Filogenia generada con la herramienta Mega7 con el método de máxima parsimonia	71
D. Filogenia generada con la herramienta Mega7 con el método de Neighbour- Joining	72
E. Filogenia generada con la herramienta Mega7 con el método de UPGMA ....	73
F. Filogenia generada con la herramienta Mega7 con el método del máximo probable (Maximum-likelihood).....	74
Anexo III. Configuración para la anotación del ensamblaje del genoma de <i>G. australe</i> con la herramienta Maker. ....	75

A.	Archivo de configuración para el entrenamiento del modelo con el genoma de <i>G. lucidum</i> .	75
B.	Archivo de configuración para la anotación del genoma de <i>G. australe</i> .....	78
C.	Anotación de las secuencias predichas por Maker2 en el servidor de KAAS.	81
D.	Tabla consenso de las regiones repetitivas identificadas con la herramienta RepeatMasker en el genoma de <i>G. australe</i> .....	82
	Anexo IV. Estadísticas obtenidas del aislamiento de la secuencia FIP putativa del genoma ensamblado de <i>G. australe</i> .....	84
A.	Secuencias utilizadas en Blast y Hmmer para la búsqueda de FIPs en el genoma de <i>G. australe</i>	84
B.	Búsqueda de Hmmer contra las FIPs de Swissprot ( <i>G. lucidum</i> y <i>F. velutipes</i> )	86
C.	Predicción de la estructura del gen putativo de la FIP de <i>G. australe</i> .....	89
D.	Alineamiento de las secuencias FIP de UniProt con la secuencia putativa de <i>G. australe</i>	92
E.	Propiedades de la secuencia putativa FIP encontrada en el genoma ensamblado de <i>G. australe</i> .	94
F.	Búsqueda de homología de la secuencia putativa FIP dentro del repositorio de Pfam	96
G.	Resultados predicción secuencia putativa de <i>G. australe</i> en I-Tasser .....	97
H.	Resultados predicción de las secuencias FIPs de <i>G. applanatum</i> y <i>C. rutilus</i> en I-Tasser	99

# Agradecimientos

## Doctores e investigadores

- Al Dr. Germán López Gartner, codirector de la tesis, quien ha venido, a través de los años, como un excelente director y, sobre todo, mentor en mi proceso de formación, y al que le tengo mucha gratitud por sus ideas y consejos.
- Al Dr. Jhon Fredy Betancur, director de la tesis, por su acogimiento, dirección y amabilidad al guiar este proceso de investigación.
- A Andrea González quién fue mi tutora desde la parte de BIOS. Gracias por todos los buenos consejos, ideas y constancia para sacar este proyecto adelante.
- A Paula Tatiana Uribe, estudiante del doctorado en Ciencias Biomédicas. Gracias por ser mi compañera y aliada en este proceso. Sé que no hubiéramos sacado esto de no ser por ti y por tus grandes aportes y confianza en que todo iba a salir bien aun cuando no lo parecía. ¡Gracias, gracias!

## Otras personas

- A las personas del laboratorio de Biología Molecular de la Universidad de Manizales por su amabilidad y disposición para colaborarnos.
- A las personas de BIOS que estuvieron implicados en este proceso (Andrea Garavito y Catalina Álvarez) gracias por su ayuda.

## Agradecimientos personales

- A Dios por iluminarme y acompañarme en cada etapa de mi vida.
- A mi mamá que siempre ha sido el motor de nuestras vidas y quién siempre me ha acompañado en los logros que he venido cosechando.
- A mis hermanos Camilo y Marcelo que me acompañan e impulsan a lograr las metas que me propongo.
- A mi abuelo, que no está conmigo, pero que fue el ejemplo de vida y a quién le dedico todos mis logros.

## Resumen

Los hongos poseen un gran número de metabolitos con potencial biotecnológico, especialmente, en el área de la salud. Uno de los componentes promisorios son las Proteínas Fúngicas Inmunomoduladoras (FIPs) y los ácidos ganodéricos, las cuales tienen la capacidad de estimular el sistema inmune, y, por lo tanto, pueden ser usados como coadyuvantes en el tratamiento de enfermedades como el cáncer. Las FIPs se han identificado en varias especies asociadas al género *Ganoderma*. Debido al potencial farmacológico que presentan estas proteínas, el objetivo de este trabajo fue identificar las secuencias putativas de proteínas FIP en el genoma de *Ganoderma australe*, una especie para la cual aún no se han descrito estos compuestos. Para dar cumplimiento a esto, se realizó el cultivo del micelio del hongo en medio sólido y líquido y extraer el ADN genómico. Posteriormente, se secuenció el genoma del hongo por medio de la plataforma Sequel de la tecnología PacBio. El genoma secuenciado se ensambló utilizando el programa de ensamblaje *de novo* CANU. El genoma fue caracterizado estructuralmente utilizando el pipeline MAKER2, y la anotación funcional se realizó mediante la herramienta de INTERPROSCAN y la herramienta de anotación automática de KEGG (KAAS), donde se reconstruyó la ruta metabólica asociada a los ácidos ganodéricos. Para la identificación de secuencias putativas de proteínas FIPs, se realizaron búsquedas por homología con BLAST y HMMER contra perfiles de proteínas FIPs reportadas en repositorios del NCBI y UNIPROT. Se obtuvo la predicción de genes de 10.713 genes en el genoma ensamblado de *G. australe*, de los cuales más de 1.000 genes son conservados en el linaje Basidiomycota. Entre las secuencias codificantes predichas, se identificaron cuatro secuencias de enzimas claves para la ruta del ácido ganodérico, asociadas su transformación. Además, se identificó una secuencia génica putativa de la FIP de 681 nucleótidos, que posee un exón de 342 nucleótidos en su sitio de transcripción y que codifica para 113 aminoácidos. Con este estudio, se reporta el primer ensamblaje del genoma de *G. australe*, y la primera secuencia putativa de FIPs para esta especie, lo cual contribuye a la aplicación biotecnológica de este hongo y sus componentes y que ayuden a generar nuevos productos que sean eficientes en el tratamiento adyuvante de enfermedades de gran impacto a nivel global.

**Palabras clave:** Ácido ganodérico, Bioinformática, FIPs, *Ganoderma*, Genoma, Proteínas.

## Abstract

Fungi possess a great number of components with biotechnological potential, mostly in biomedical fields. Among these, Fungal Immunomodulatory Proteins (FIPs) are proteins with immunomodulatory actions against diseases like cancer or diabetes. This group of proteins have been found and characterized in several species within *Ganoderma* genus. Given this, we aimed to identify putative FIPs sequences in the genome of *Ganoderma australe*. To accomplish this, we obtained genomic DNA extracted from fungal mycelium from laboratory culture. Genomic DNA was sequenced using PacBio Sequel platform and assembled with Canu *de novo* assembly software. The genome of *G. australe* was structurally characterized using Maker2 pipeline, based on *ab initio* gene predictors and protein evidences from *Ganoderma* databases. Functional annotation was made with INTERPROSCAN and KEGG automatic annotation tool (KAAS), for the reconstruction of metabolic pathway for ganoderic acids. Blast and Hmmer tools were used to search the genome against protein profiles from NCBI and UNIPROT databases, in order to identify putative FIPs sequences. In total, 10,713 predicted genes were obtained from *G. australe* genome assembly, of which more than a 1,000 are conserved within the Basidiomycota lineage. Among the predicted sequences, four sequences were identified from the ganoderic acids pathway, which are key in metabolite production in this pathway. As well, a 681-nucleotide putative FIP sequences was recognize, containing 342 nucleotides in the transcription region that translate into 113 aminoacids. Therefore, in this study, we report the first genome sequence for *G. australe* and the first putative FIP gene sequence for this species. This contributes to the biotechnological application of this fungus and its components. Besides, the search for promising medical compounds in *Ganoderma* genus is necessary to provide new products that can be effective in coadjuvant treatments for high-risk diseases around the world.

**Keywords:** Bioinformatics, FIPs, *Ganoderma*, Ganoderic acid, Genome, Proteins.



## Lista de tablas

Tabla 1. Proteínas Fúngicas Inmunomoduladoras (FIPs) reportadas para el género *Ganoderma*.

Tabla 2. Reporte de FIPs para los Filos Basidiomycota y Ascomycota.

Tabla 3. Concentración de ADNg cuantificados con el NanoDrop y Glomax/Qubit

Tabla 4. Reporte con QUAST de estadísticas del ensamblaje del genoma de *G. australe*.

Tabla 5. Resultado de BUSCO para el ensamblaje de *G. australe*.

Tabla 6. Enzimas evaluadas para la ruta del ácido ganodérico dentro del genoma ensamblado de *G. australe*.

Tabla 7. Resultado de homología con BlastP y Hmmer para las proteínas FIPs dentro del genoma de *G. australe*.

## Lista de figuras

Figura 1. Descripción de la ruta metabólica de los ácidos ganodéricos en *G. lucidum*. Tomado de Liu *et al.* (2012)

Figura 2. Estructura primaria y secundaria de la FIP-gts en *G. tsugae*. Tomado de Lin *et al.* (1997).

Figura 3. Metodología descrita para el cultivo de *G. australe* y secuenciación del ADNg del hongo. Figura 4.

Figura 4. Metodología descrita para el ensamblaje y anotación del *genoma de G. australe*.

Figura 5. Electroforesis del ADNg extraído del micelio de *G. australe*. Pozos 1-3: ADN lambda a 15 ng, 30 ng y 45 ng; pozos 4-6: ADN lambda a 15 ng, 30 ng y 45 ng con HindIII; pozos 7-25: ADNg de *G. australe*.

Figura 6. Identificación molecular de *G. australe* usando las regiones ITS 4 y 5 del ADN ribosomal.

Figura 7. a. Resultado de la anotación de las secuencias putativas del genoma de *G. australe* con KEGG. b. Reconstrucción de las secuencias putativas dentro de la ruta de biosíntesis de triterpenos.

Figura 8. Estructura del gen putativo de la FIPs encontrado en el genoma de *G. australe* basado en la predicción con el programa Fgenesh.

Figura 9. Predicción de dominios proteicos para la secuencia putativa de la FIP de *G. australe* con la herramienta InterProScan.

Figura 10. Modelo de estructura terciaria para la secuencia proteica putativa de la FIP de *G. australe*, con base en la predicción obtenida con la herramienta I-Tasser.

## Lista de anexos

- Anexo I. A. Resultado de la secuenciación del genoma de *G. australe* por la tecnología de secuenciación Sequel de PacBio. 64
- B. Comandos utilizados para el ensamblaje del genoma de *G. australe* utilizando la herramienta Canu. 65
- Anexo II. Análisis del ensamblaje del genoma de *G. australe* 66
- A. Comparación genómica utilizando la herramienta Mummer de los genomas de *G. australe* vs *G. lucidum* 66
- B. Comparación genómica utilizando la herramienta Mummer de los genomas de *G. australe* vs *G. boninense* 67
- C. Filogenia generada con la herramienta Mega7 con el método de máxima parsimonia 68
- D. Filogenia generada con la herramienta Mega7 con el método de Neighbour-Joining 69
- E. Filogenia generada con la herramienta Mega7 con el método de UPGMA 70
- F. Filogenia generada con la herramienta Mega7 con el método del máximo probable (Maximum-likelihood) 71
- Anexo III. Configuración para la anotación del ensamblaje del genoma de *G. australe* con la herramienta Maker2. 72
- A. Archivo de configuración para el entrenamiento del modelo con el genoma de *G. lucidum*. 72
- B. Archivo de configuración para la anotación del genoma de *G. australe*. 75
- C. Anotación de las secuencias predichas por Maker2 en el servidor de KAAS. 78
- D. Tabla consenso de las regiones repetitivas identificadas con la herramienta RepeatMasker en el genoma de *G. australe* 79
- Anexo IV. Estadísticas obtenidas del aislamiento de la secuencia FIP putativa del genoma ensamblado de *G. australe*. 81
- A. Secuencias utilizadas en Blast y Hmmer para la búsqueda de FIPs en el genoma de *G. australe* 81
- B. Búsqueda de Hmmer contra las FIPs de Swissprot (*G. lucidum* y *F. velutipes*) 83
- C. Predicción de la estructura del gen putativo de la FIP de *G. australe* 86
- D. Alineamiento de las secuencias FIP de UniProt con la secuencia putativa de *G. australe* 89
- E. Propiedades de la secuencia putativa FIP encontrada en el genoma ensamblado de *G. australe*. 91
- F. Búsqueda de homología de la secuencia putativa FIP dentro del repositorio de Pfam 93

G. Resultados predicción secuencia putativa de *G. australe* en I-Tasser 94

H. Resultados predicción de las secuencias FIPs de *G. applanatum* y *C. rutilus* en I-Tasser

96

## Lista de abreviaturas

ADNg: ADN genómico

BUSCO: Benchmarking Universal Single-Copy Orthologs.

CYP450: Citocromo P450.

FDS: Farnesyl Difosfato Sintasa.

FIPs: Fungal Immunomodulatory Proteins.

GO: Gene Ontology.

HPC: Computación de alto rendimiento.

ITS: Internal Transcribed Sequence.

LS: Lanosterol Sintasa.

KAAS: KEGG Automatic Annotation Server.

KEGG: Kyoto Encyclopedia of Genes and Genomes.

NCBI: National Center for Biotechnology Information.

PDB: Protein Data Bank.

QUAST: Quality Assessment Tool for Genome Assemblies.

SQM: Squaleno Monooxigenasa.

SS: Squaleno Sintasa.

UGT: Uridindifosfato Glucosil Transferasa

# 1. Introducción

## 1.1. Campo temático

En el año 2014 la Organización Mundial de la Salud (OMS) reportó que las enfermedades no transmisibles (enfermedades cardiovasculares, autoinmunes, cáncer, enfermedades respiratorias crónicas y diabetes) constituyeron la principal causa de muerte a nivel mundial (OMS 2014), ya que varían en tasas desde el 1,3% hasta el 33% de incidencia, mostrando un aumento considerable. Por esta razón, la OMS plantea que: “reducir la carga mundial de enfermedades no transmisibles es una prioridad absoluta y una condición necesaria para un desarrollo sostenible”; como consecuencia, ha propuesto elaborar estrategias enfocadas en la modificación de los factores de riesgo asociados a este tipo de enfermedades, con las cuales se disminuya la mortalidad en un 25% para el año 2025 (OMS 2014). No obstante, aunque los gobiernos desarrollen estrategias de promoción y prevención para este tipo de enfermedades, es una realidad que cada vez un número más elevado de la población mundial llegará a padecer de alguna de ellas, debido a estilos de vida y/o por factores genéticos predisponentes.

En la actualidad, se cuenta con un abanico de tratamientos farmacológicos que permiten controlar y prevenir complicaciones posteriores de estas patologías, que en muchos casos desencadenan reacciones adversas y efectos secundarios en los pacientes agravando aún más su estado de salud (Escobar *et al.* 2000). Además, estas patologías crónicas representan un aumento significativo en el costo para los sistemas de salud.

Debido a esto, en las últimas décadas se han intensificado las investigaciones enfocadas en la búsqueda de compuestos de naturaleza biológica, que favorezcan la generación de nuevas alternativas para el tratamiento y control de estas patologías, y que aporten a la reducción de las reacciones adversas y efectos secundarios en los pacientes, con el fin de aumentar las expectativas de vida de quienes padecen este tipo de enfermedades de alta carga.

Gracias al desarrollo de la biotecnología, la industria farmacéutica ha desarrollado medicamentos cada vez más eficientes, a partir de avances en la búsqueda de fuentes de compuestos bioactivos para la obtención de nuevos fármacos. Es así como la exploración de metabolitos secundarios y otros compuestos biológicos se han convertido en un área activa de investigación. Con respecto a ello, desde hace más de medio siglo se han venido desarrollando estudios en varias especies de hongos micromicetos, orientadas al conocimiento de las actividades biológicas generadas por sus metabolitos secundarios y otras clases de productos celulares, y su gran potencial terapéutico sobre la salud humana como: antiviral, antimicótico, antibacterial, hipoglucemiante, antitumoral, antihiperlipidémico, antialérgico, antiinflamatorio e inmunomodulador, entre otros (Suárez Arango & Nieto 2013).

En la búsqueda y obtención de productos naturales a partir de hongos las estrategias químicas tradicionalmente empleadas para la obtención de este tipo de moléculas pueden no ser suficientes, ya que, como se ha observado en algunos estudios, se requiere un porcentaje elevado de biomasa para obtener una concentración muy baja de estos compuestos. Es en este punto donde cobra importancia el avance investigativo en ciencias ómicas en macromicetos con potencial terapéutico, como es el caso de las especies asociadas al género *Ganoderma*, ya que los análisis genómicos empleando herramientas bioinformáticas y de biología computacional permiten explorar el potencial genético y metabólico de estos hongos. De esta manera, se puede lograr dilucidar los mecanismos moleculares para determinar los efectos medicinales de los macromicetos y enfocar nuevas investigaciones para la obtención de este tipo de moléculas empleando las herramientas de la tecnología del ADN recombinante (Wasser 2014).

## **1.2. Planteamiento del problema**

Las propiedades de las Proteínas Fúngicas Inmunomoduladoras (FIPs) han sido ampliamente estudiadas por su potencial terapéutico en el tratamiento de enfermedades cancerígenas y autoinmunes como el VIH y la diabetes, como agentes inmunomoduladores (Ng 2004, Moradali, Mostafavi, Ghods & Hedjaroude 2007, Xu *et al* 2011a). Sin embargo, su aprovechamiento se dificulta fundamentalmente por dos razones: una es que la producción de *Ganoderma* spp. no ha permitido un nivel de sostenibilidad y escalamiento comercial que satisfaga la demanda del mismo por la difícil producción que presenta en el cultivo en laboratorio (Lin *et al.* 2009). El segundo aspecto es que la expresión de la proteína en material seco del hongo es muy baja, se encuentra en cantidades de 2-5 mg/L en el tejido vivo del hongo, cantidad que no es comparable con la escala industrial, a la cual se produce alrededor de 700 mg/L de una proteína de interés (Lin *et al.* 2009).

Además, en algunas formas de preparación del hongo, las propiedades terapéuticas de las moléculas se reducen sustancialmente. Todo ello hace imperioso el desarrollo de estrategias biotecnológicas para la producción en biorreactor del metabolito y/o proteína de interés, utilizando un organismo modelo que favorezca la expresión, exigiendo la clonación de los genes involucrados en dicho metabolismo o de la proteína, partiendo necesariamente de su identificación y caracterización a partir del recurso genético original. En esta parte se introducen las herramientas bioinformáticas como facilitadoras en la identificación y caracterización de los elementos que se requieren evaluar, y así, poder obtener resultados confiables que permitan su implementación en el laboratorio.

### 1.3. Justificación

*Ganoderma australe* es una especie prominente en el área bioinformática por el poco estudio que posee en el ámbito genómico, haciéndola un objeto de estudio de gran interés a fin de descubrir el potencial genético que pueda aplicarse en áreas como la medicina, la biotecnología y la industria en general. Esta especie ha mostrado grandes cualidades como biorremediador y descomponedor de componentes vegetales en especies maderables (Rigas *et al.* 2007), sirviendo como primer referente y aporte para áreas ecológicas (Mendoca *et al.* 2008). Por su parte, el género *Ganoderma* posee una especie ampliamente estudiada, *Ganoderma lucidum*, de la cual se dispone un genoma completo (Chen *et al.* 2012) y de múltiples estudios de caracterización bioquímica (Wang *et al.* 2014, Bishop *et al.* 2015, Mota *et al.* 2015, Saltarelli *et al.* 2015, Xu *et al.* 2017), siendo un referente cercano y significativo a la hora de realizar estudios dirigidos al descubrimiento de genes de interés contenidos en el genoma de *G. australe*. Además, se cuenta con investigaciones enfocadas en la identificación de genes que codifican para la proteína fúngica inmunomoduladora en otras especies del género *Ganoderma* como *G. atrum*, *G. sinense*, *G. japonicum*, *G. applanatum*, *G. microsporum*, entre otros, y en Basidiomicetos del orden Polyporales como *Flammulina velutipes*, *Trametes versicolor* y *Volvariella volvacea*. Esto lleva a *G. australe* a ser una especie de gran interés para entender su funcionamiento desde su genoma y su metabolismo, con un énfasis en los genes que codifiquen para las FIPs y las enzimas que están asociadas a la producción de los ácidos ganodéricos.

Igualmente, el continuo avance de la tecnología y la masiva cantidad de información generada a partir de los sistemas biológicos desde el punto de vista volumétrico en las áreas genómicas y transcriptómicas, permiten una exploración importante de esta especie en cuanto su contenido génico y metabólico para contribuir en el campo biotecnológico en áreas medicinales, nutracéuticas e industriales y, por medio de ellos, dar vía a desarrollos promisorios en campos aún no explorados por la ciencia en nuestro país.



## **1.4. Objetivos**

### ***1.4.1. Objetivo general***

Identificar secuencias de Proteínas Fúngicas Inmunomoduladoras (FIPs) putativas y las enzimas claves de la ruta del ácido ganodérico en el genoma de *Ganoderma australe*.

### ***1.4.2. Objetivos específicos***

- Obtener la secuencia del genoma de *Ganoderma australe* por tecnologías de secuenciación de nueva generación (NGS).
- Caracterizar el genoma de *Ganoderma australe* para la búsqueda de las enzimas claves asociadas a la ruta del ácido ganodérico mediante herramientas bioinformáticas.
- Definir las secuencias putativas de Proteínas Fúngicas Inmunomoduladoras (FIPs) dentro del genoma de *Ganoderma australe*.

## 2. Referente teórico y antecedentes

### 2.1. Importancia biológica de los hongos

El reino Fungi es considerado uno de los mayores recursos de biodiversidad presente en casi todos los nichos ecológicos (Dighton 2016). Los hongos cumplen roles importantes en los ecosistemas y contienen un significativo número de metabolitos con distintas propiedades en su estructura química y bioactividad, haciéndolos un recurso abundante y promisorio para el descubrimiento de nuevas aplicaciones en diversos ámbitos biotecnológicos (Dighton & White 2017). Las culturas humanas, en especial la cultura China, han utilizado los hongos para diferentes usos, desde la industria del blanqueamiento de pulpa de papel hasta la industria farmacéutica y la producción de alimentos con propiedades nutraceuticas (McMeekin 2004)

Mueller y Schmith (2007) sugieren que la estimación tradicional del número de especies de hongos, en alrededor de 1,5 millones de especies, es todavía muy modesta y no incluye por ejemplo especies provenientes de muestras ambientales complejas, en muchos casos sin describir completamente. Además, del número de especies estimadas, solo se han descrito alrededor del 7%. Aunque, estimaciones más recientes sugieren que pueden existir cerca de 5 millones de especies de hongos en total, superando la biodiversidad vegetal conocida hasta la fecha (Blackwell 2011).

Un aspecto muy relevante de las aproximaciones biotecnológicas de los hongos, sobre todo en el campo de las ciencias biomédicas, procede del hecho que muchos de los modelos para el desarrollo e implementación de medicamentos a partir de productos naturales corresponden a sustancias bioactivas provenientes de éstos o de metabolitos secundarios de diversos microorganismos (Strobel 2003). En particular, impresiona el hecho de que la mayoría de los metabolitos secundarios descubiertos en los hongos poseen estructuras químicas muy similares a las producidas en los medicamentos (Strobel 2003, Cooper 2004, König *et al.* 2006, Zhong & Xiao 2009, Zhao *et al.* 2010). Desafortunadamente, únicamente del 5 al 10% de las especies de hongos de uso humano pueden ser cultivados eficientemente (Manoharachary *et al.* 2005), insinuando la necesidad de implementar estudios en macromicetos que contrarresten y den respuesta a la problemática que existe en el momento en cuanto a una mayor incidencia de un grupo de enfermedades infecciosas y no infecciosas, relacionadas con estados de respuesta inmune inapropiada, y que aún no poseen un tratamiento adecuado o que emplean medicamentos de origen químico con efectos secundarios (Cuéllar & Rodríguez 2003, Nova *et al.* 2004).

#### 2.1.1. Propiedades biológicas de los hongos

Los hongos macromicetos presentan compuestos con propiedades medicinales y biotecnológicas importantes, entre los que destacan especies de los como *Grifola frondosa*, *Pleurotus*

sp. y *Agaricus* sp. (Illana-Estaban 2008, Smiderle *et al.* 2012, Vincent *et al.* 2013). Un grupo de compuestos de importancia son las **lectinas**, que constituyen un diverso conjunto de proteínas de unión a carbohidratos, aisladas de diferentes organismos como humanos, plantas y hongos (Zhang *et al.* 2010). Estas proteínas poseen la capacidad de especificidad inmunomoduladora, actividad antitumoral y antiproliferativa (Sze 2004, Zhang *et al.* 2010, Yau *et al.* 2015, Ng *et al.* 2016). Su actividad biológica se ejerce a través de la unión y la agrupación de receptores glicosilados en la superficie celular, conduciendo a la activación de mecanismos de traducción de señales, en la matriz extracelular, en el citoplasma y en glicoproteínas nucleares (Pohleven *et al.* 2009). De esta manera, las lectinas están involucradas en procesos celulares como la adhesión celular, la migración, la diferenciación, la apoptosis y la proliferación, y, por lo tanto, pueden poseer propiedades inmunomoduladoras y tener un rol potencial en tratamiento del cáncer (Pohleven *et al.* 2009). Así mismo, las lectinas han sido aisladas en especies de hongos como *Volvariella volvacea*, *Tricholoma mongolicum*, *Ganoderma capense*, entre otras (Jung-Yaw & Tze-Bin 1984, Kawagishi *et al.* 1990, Wang *et al.* 1996, Ngai & Ng 2004).

De igual manera, se han presentado reportes en especies de *Agaricus*, *Lentinus*, *Podospora*, *Rhizoctonia*, *Polyporus*, *Pholiota*, *Coriolus*, *Neurospora* y *Phlebia* de la presencia de **lacasas**, que son glicoproteínas que poseen átomos de cobre como centros catalíticos de acción (Cañas & Camarero 2010). Estas proteínas catalizan la oxidación de los compuestos fenólicos y aminas aromáticas con el oxígeno, jugando un rol importante en la degradación de Lignina (Wang & Ng 2006). En el ámbito biotecnológico, las lacasas pueden considerarse como biosensores, como agentes para descontaminar aguas y como colorantes para la industria textil (Wang & Ng 2006, Li *et al.* 2010). Un estudio realizado por Xu, Wang & Ng (2012) donde demostraron el potencial inhibitorio de la lacasas para combatir enfermedades del Virus de Inmunodeficiencia Humana (VIH) con un hongo del género *Lentinus*, lo cual comprueba que este tipo de proteínas poseen aplicación en biotecnología ambiental y en el ámbito de la salud.

Además de las lacasas, los hongos poseen otros compuestos como los **terpenos**. Estos compuestos son un diverso grupo orgánico, de cinco carbonos, que se unen formando bloques, y que adquieren sus nombres de acuerdo al número de unidades repetidas de isoprenos, como los monoterpenos (10 carbonos), sesquiterpenos (15 carbonos), diterpenos (20 carbonos), triterpenos (30 carbonos), tetraterpenos (40 carbonos) (Elsayed *et al.* 2014). Estos componentes existen ampliamente en las plantas como principales componentes de resinas y aceites esenciales (Lull, Wichers & Savelkoul 2005, Elsayed *et al.* 2014). En los macromicetos están presentes en formas modificadas (isoprenoides y terpenoides), con gran potencial medicinal ya que presentan actividades antifúngicas,

antimicrobianas, antitumorales, inmunomoduladoras entre otras (Pillai, John & Thomas 2011, Tang *et al.* 2011, Patlolla & Rao 2012, Li *et al.* 2013).

El género *Ganoderma* presenta una variedad de terpenos como los ácidos ganodéricos, ácidos ganodérmicos, ganoderoles, ganoderales, entre otros, que han sido aislados y estudiados de especies como *G. lucidum* y *G. applanatum* (Lui *et al.* 2007, Jeong *et al.* 2008, Xu, Zhao & Zhong 2010, López-Gartner *et al.* 2015), los cuales ejercen funciones hepatoprotectoras, antifúngicas, antibacteriales, entre otras. Estos componentes pueden presentar efectos en enfermedades cancerosas como inhibición de la proliferación y del comportamiento invasivo de las células cancerosas de mama y próstata a través de la disminución regulada de la expresión de ciclina D1 y la supresión de la secreción del activador urocinasa del plasminógeno (uPA); inhibición del crecimiento e inducción de apoptosis de cáncer de próstata y mama a través del aumento regulado de la expresión de p21 y Bax; inhibición del crecimiento de hepatoma a través de la supresión de protein kinasa C (PKC), inducción de apoptosis de células de cáncer de colon por incremento de la actividad caspasa-3, supresión de angiogénesis a través de la inhibición de la secreción del factor de crecimiento endotelial vascular (VEGF) y del factor de crecimiento y transformación (TGF- $\beta$ ) de células de cáncer de próstata (Ríos-Cañavate 2008)

## **2.2. Género *Ganoderma***

*Ganoderma* P. Karsten (1881) (Ganodermataceae) es uno de los géneros de hongos macromicetos más representativos en la naturaleza para el estudio de componentes con aplicaciones biotecnológicas, debido al alto contenido de metabolitos y su gran potencial en diversas aplicaciones en la industria, la biorremediación y la salud, entre otros (Baby, Johnson & Govindan 2015). El género *Ganoderma* pertenece a la Filo Basidiomycota, Clase Agaricomycetes, Orden Polyporales, Familia Ganodermataceae, Género *Ganoderma*.

El género comprende hongos de la pudrición blanca de la madera que descomponen la lignina, la celulosa, y algunos polisacáridos mediante enzimas modificadoras de las ligninas como las lacasas, lignina-peroxidasas y peroxidasas magnesio-dependientes (D'Souza *et al.* 1996). Este grupo de hongos ha sido reconocido, en principio por la medicina tradicional, por sus propiedades nutraceuticas, y como un buen complemento terapéutico contra enfermedades, presentando elementos contundentes para disminuir los efectos adversos en la salud humana, lo cual ha sido confirmado recientemente a través numerosos estudios formales que han permitido identificar los componentes químicos que explican su eficacia en el campo biomédico (Sheena *et al.*, 2005, Lin *et al.*, 2006, Russell & Patterson, 2006, Hsiao *et al.*, 2008, Chang *et al.* 2009, Chen *et al.*, 2012) e industrial y biotecnológico. Varias especies de *Ganoderma* han resultado ser efectivas en diversas aplicaciones biotecnológicas, por ejemplo, Ozel *et al.* (2012) reportaron la capacidad de *G. lucidum* para la

biodegradación de cianuro; Matos *et al.* (2007) investigaron la aplicación de *G. applanatum* para el tratamiento en la decoloración y defenolización de aguas residuales; Mendonca *et al.* (2008) y Rigas *et al.* (2007) estudiaron la implementación de *G. australe* en la biorremediación de suelos contaminados con compuestos orgánicos recalcitrantes y con lindano (plaguicida de uso en agricultura).

Del mismo modo, en numerosos estudios se han comprobado las propiedades estimuladoras de la respuesta inmune que poseen los hongos del género *Ganoderma* (Lull, Wichers & Savelkoul. 2005, Liao *et al.* 2006, Liao *et al.* 2008), demostrando su efectividad en la modulación de funciones inmunológicas e inhibición del crecimiento tumoral (Jeurink *et al.* 2008, Ding *et al.* 2009). Entre los componentes que le confieren estas propiedades se encuentran principalmente Proteínas Fúngicas Inmunomoduladoras (*Fungal Immunomodulatory Proteins*, FIPs), lectinas, lacasas y metabolitos derivados de las rutas metabólicas del ácido mevalónico y triterpenos, como los ácidos ganodéricos (Ramírez *et al.* 2003, Güzeldağ & Colak 2007, Arboleda Echavarría & Mejía Gallón 2010). Así mismo, algunas especies de *Ganoderma* también contienen distintos compuestos con comprobada actividad inhibitoria de mitosis en líneas celulares, como los polisacáridos, polifenoles, esteroides y estatinas, dentro de los cuales se destacan su actividad antimicrobiana y citotóxica, su efecto hipercolesterolémico y su acción antiinflamatoria (Shiao 2003, Russell & Paterson 2006, Schepetkin & Quinn 2006, Wachtel-Galor *et al.* 2011, Zonga *et al.* 2012).

### **2.3. Ácidos Ganodéricos**

En *Ganoderma* se han identificado más de 400 compuestos con actividad biológica, y que se relacionan con los efectos medicinales e industriales observados a partir de otros compuestos biológicos presentes en especies evolutivamente mayores como plantas y animales, no obstante, algunos son exclusivos de este género de hongos y han sido objeto de interés por algunos investigadores (Baby, Johnson & Govindan 2015).

A partir de esto, Kim & Kim (1999) describieron unos componentes exclusivos del género *Ganoderma*, los ácidos ganodéricos, que hacen referencia a moléculas de la familia de los triterpenos que poseen características bioquímicas importantes en el área biotecnológica, dado que tienen propiedades útiles en la medicina para el tratamiento de enfermedades cancerígenas por la inducción de muerte programada (apoptosis) en células malignas y así detener el avance de las mismas en la enfermedad (Tang *et al.* 2006, Tang *et al.* 2011, You *et al.* 2013). Además, estos metabolitos presentan eficacia en la estimulación del sistema inmune y propiedades antiinflamatorias (Lull, Wichers & Savelkoul 2005), y la protección del hígado contra agentes tóxicos (Ren *et al.* 2010).

En un estudio realizado por Sanodiya *et al.* (2009) en *Ganoderma lucidum*, se evaluó el efecto que poseían los ácidos ganodéricos relacionado al aporte de soluciones contra enfermedades como las alergias, cáncer o VIH/SIDA. Además, varias investigaciones han mostrado la relevancia de los ácidos ganodéricos para la prevención y/o acción contra células altamente susceptibles a procesos cancerígenos o autoinmunes (Min *et al.* 2000, Gao *et al.* 2002, Lin & Zhang 2004, Li *et al.* 2009, Xu *et al.* 2011b), obteniendo logros muy promisorios en el campo biomédico.

El metabolismo de los ácidos ganodéricos se deriva de la ruta metabólica del ácido mevalónico (KEGG map01062, map00900) cuyo producto final en plantas, animales y hongos esta descrita para la producción de metabolitos con núcleo esterólico a nivel celular como el colesterol, fitoesterol y ergosterol (Schmidt-Dannert 2014). Con la intención de dilucidar la producción de estos metabolitos, Liu *et al.* (2012) (ver Figura 1) describieron la ruta de la biosíntesis de los ácidos ganodéricos en *Ganoderma lucidum*, que inicia a partir del acetil-CoA hasta llegar al lanosterol, donde ocurre una transformación de la molécula de Lanosterol por parte del complejo del citocromo P450 (CYP450) y del complejo enzimático uridindifosfato glucosil transferasa [uridin diphosphate glycosyl transferases] (UGT) para generar las diferentes transformaciones (reacciones redox) hacia los triterpenos, y por último, los diferentes ácidos ganodéricos (Qian *et al.* 2013)

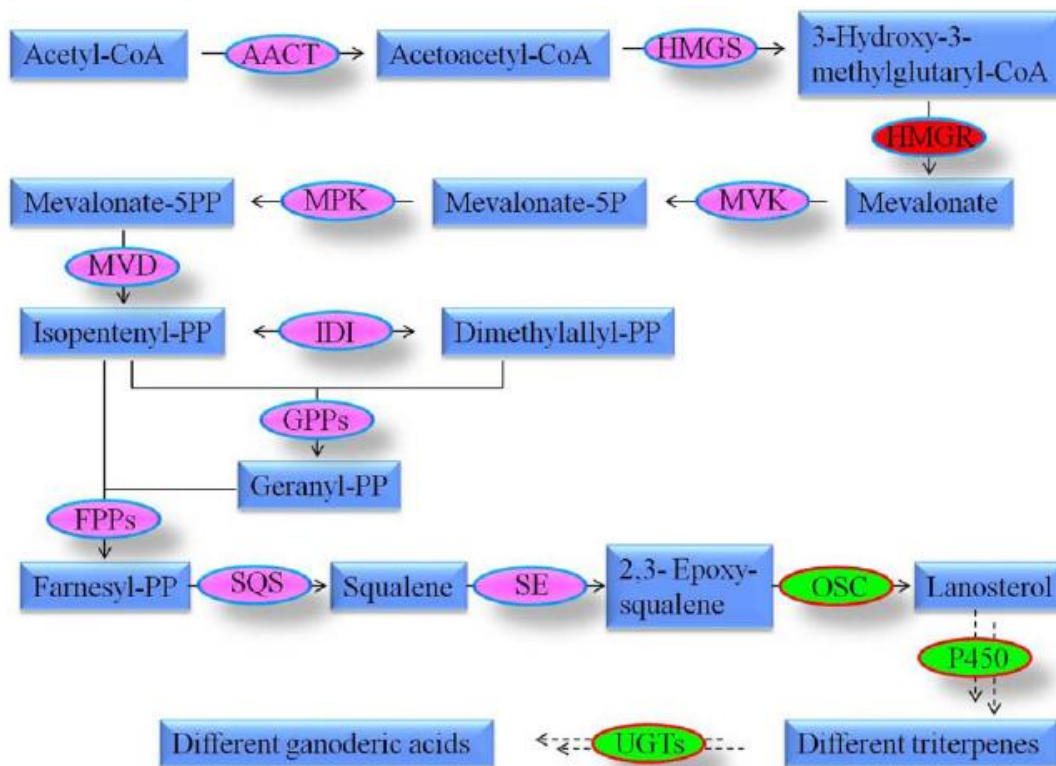


Figura 1. Descripción de la ruta metabólica de los ácidos ganodéricos en *G. lucidum*. Tomado de Liu *et al.* (2012).

## 2.4. Proteínas Fúngicas Inmunomoduladoras (FIPs)

Las investigaciones sobre las propiedades de inmunomodulación de las especies de *Ganoderma* han profundizado en el estudio de compuestos bioactivos como polisacáridos, ácido ganodérico, triterpenos y Proteínas Fúngicas Inmunomoduladoras (*Fungal Immunomodulatory Proteins*, FIPs), las cuales son proteínas con actividad reguladora del sistema inmunológico (Li *et al.* 2011). Estas proteínas modulan la activación y expresión de las células y mensajeros comprometidos en la respuesta inmunológica, conocido como inmunomodulación (Li *et al.* 2015). Dada la función inmunológica que pueden presentar este tipo de proteínas, se utilizan en campos de acción medicinal contra enfermedades como cáncer, enfermedades inflamatorias y alérgicas (Xu *et al.* 2011a, Bastiaan-Net 2013).

En el año 1989 Kino *et al.* reportaron la primera proteína fúngica inmunomoduladora (FIPs) denominada LZ-8, aislada del cuerpo fructífero del hongo medicinal *G. lucidum* (nombre común en chino Ling Zhi). A partir de esta fecha, se han reportado y caracterizado siete FIPs para especies del género *Ganoderma*, purificadas y/o clonadas (ver Tabla 1).

Tabla 1. Proteínas Fúngicas Inmunomoduladoras (FIPs) reportadas para el género *Ganoderma*.

<i>Ganoderma</i> sp.	Denominación de la FIP	Referencia	Accesión GenBank
<i>Ganoderma lucidum</i>	LZ-8	Kino <i>et al.</i> 1989.	P14945.2.
<i>Ganoderma tsugae</i>	FIP-gts	Lin <i>et al.</i> 1997.	No posee
<i>Ganoderma sinense</i>	FIP-gsi	Zhou <i>et al.</i> 2009.	AY449805.1
<i>Ganoderma microsporum</i>	FIP-gja	Wu <i>et al.</i> 2007.	3KCW_A.
<i>Ganoderma japonicum</i>	FIP-gmi	Sin publicar.	AAX9824.
<i>Ganoderma atrum</i>	FIP-gas	Sin publicar.	KM077027.1
<i>Ganoderma applanatum</i>	FIP-gap	Sin publicar.	AEP68179.

Trece proteínas FIP pertenecientes a otros géneros del Filo Basidiomycota también han sido reportadas y descritas a nivel molecular o de secuencia. En total, existen 20 reportes de secuencias de proteínas FIP en los Filos Basidiomycota y el Filo Ascomycota (ver Tabla 2), incluyendo al género *Ganoderma*.

Para el Filo Ascomycota, se han reportado dos FIPs en *Nectria haematococca* (FIP-nha) y *Stachybotrys chlorohalonata* (FIP-sc2), respectivamente. En Colombia, existe un reporte preliminar

de una secuencia putativa de FIP encontrada en el transcriptoma de *G. australe* por González-Muñoz *et al.* (2014).

Tabla 2. Reporte de FIPs para los Filos Basidiomycota y Ascomycota\*.

<b>Especie</b>	<b>Denominación de la FIP</b>	<b>Referencia</b>	<b>Accesión Genbank</b>
<i>Flammulina velutipes</i>	FIP-fve	Ko <i>et al.</i> 1995.	P80412.
<i>Volvariella volvacea</i>	FIP-vvo	Hsu <i>et al.</i> 1997	No posee.
<i>Auricularia polytricha</i>	FIP-Apl	Sheu <i>et al.</i> 2004	XM_007336979.1
<i>Antrodia camphoratae</i>	FIP-aca	Sheu <i>et al.</i> 2009.	AY569691.1.
<i>Pleurotus citrinopileatus</i>	FIP-PCiP	Sheu <i>et al.</i> 2007.	No posee.
<i>Poria cocos</i>	FIP-pcp	Chang, Yeh & Sheu 2009.	JN571084.1.
<i>Trametes versicolor</i>	FIP-tvc	Li <i>et al.</i> 2012, Li <i>et al.</i> 2011.	KC297708.1.
<i>Postia placenta</i>	rFIP-ppl	Li <i>et al.</i> 2015.	AJL35148.1
<i>Chroogomphus rutilus</i>	FIP-cru	Lin <i>et al.</i> 2016.	AKU37620.1.
<i>Tremella fuciformis</i>	FIP-tfu	Sin publicar.	EF152774.
<i>Lignosus rhinocerotis</i>	FIP-Lrh	Pushparajah <i>et al.</i> 2016.	No posee.
<i>Nectria haematococca*</i>	FIP-nha	Li <i>et al.</i> 2014.	XM_003043608.1.
<i>Stachybotrys chlorohalonata*</i>	FIP-sh2	Li <i>et al.</i> 2017.	AQQ80204.1.

#### 2.4.1. Estructura Molecular

Las proteínas FIPs poseen un peso molecular de 12,5-15 KDa, tienen una estructura primaria constituida de 111 a 121 aminoácidos, la estructura secundaria está compuesta de siete plegamientos beta ( $\beta$ ), dos hélices alfa ( $\alpha$ ) y un loop beta (Li *et al.* 2011)

Lin *et al.* (1997) estudiaron la estructura de las FIPs (ver Figura 2) en *Ganoderma tsugae*, realizando el alineamiento y comparación con las proteínas FIP reportadas con antelación para *G. lucidum* (Kino *et al.* 1989), *F. velutipes* (Ko *et al.* 1995) y *V. volvacea* (Hsu *et al.* 1997). Los autores describieron la homología que poseen las secuencias de FIPs reportadas por la conservación de un dominio en el extremo amino terminal del péptido. En esta parte de la secuencia se cree que es el sitio



activo donde se ejecuta los procesos celulares la proteína, dado que cuando los autores realizaron el estudio, hicieron un proceso mutagénico donde hicieron deleciones de dos leucinas (posición 5 y 9) y una fenilalanina (posición 7) en la FIP-gts natural. Con ello, determinaron que las FIP-gts mutantes, carentes de estos aminoácidos, no podían ejecutar el plegamiento que se observaba en la FIP-gts natural en el proceso de dimerización (unión a los receptores para ejecutar su actividad), concluyendo que esta región de la proteína era el sitio activo para ejecutar las acciones inmunomoduladoras (Lin *et al.* 1997). Esta parte molecular de la proteína es altamente conservada entre las secuencias reportadas de las FIPs (Li *et al.* 2015)

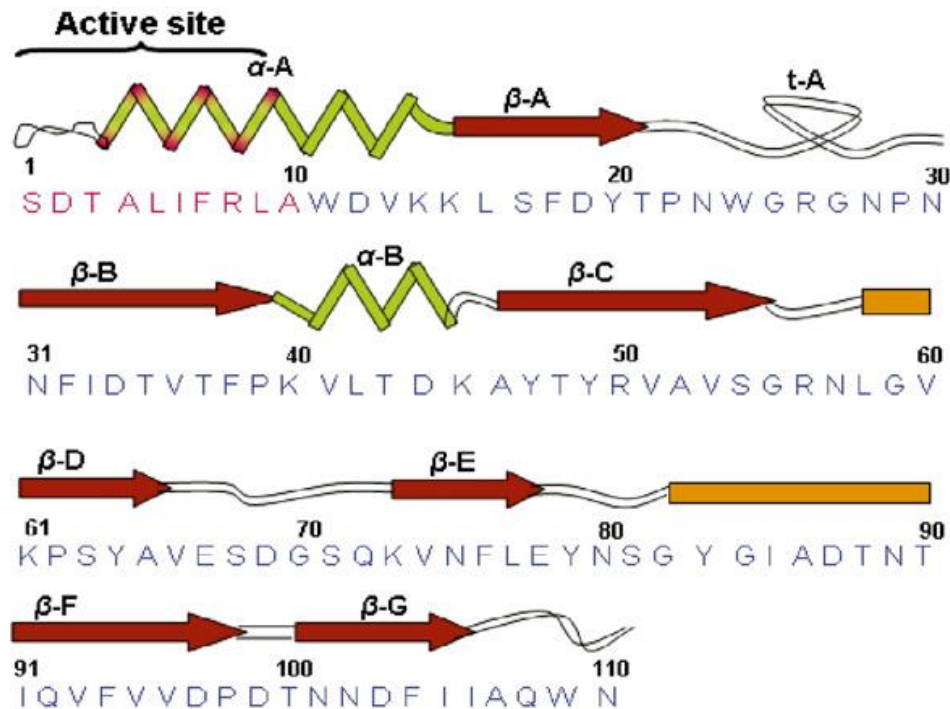


Figura 2. Estructura primaria y secundaria de la FIP-gts en *G. tsugae*. Tomado de Lin *et al.* (1997).

#### 2.4.2. Aplicaciones y prospectos

Se ha propuesto que la obtención de las FIPs puede ser útil para mejorar el sistema inmune y, de esta forma, potenciar la respuesta inmunológica y fomentar tratamientos novedosos para el cáncer y enfermedades autoinmunes. Así mismo, estas proteínas podrían ser empleadas como productos inmunológicos en insumos farmacéuticos o como vacunas que estimulen la resistencia de células NK (Natural Killer) del sistema inmune para la prevención de enfermedades, así como suplemento para aliviar el dolor o reducir los efectos secundarios generados por los medicamentos sintéticos, que poseen efectos adversos en pacientes diagnosticados con estas enfermedades (Li, Wang & Zhou 2011). Por ende, los tratamientos que se relacionan con la acción de las proteínas FIPs

deben estar dirigidos a mejorar los tratamientos como coadyuvantes en enfermedades como cáncer y diabetes (Patel & Goyal 2012, Kladar *et al.* 2016, Wang *et al.* 2017). De ahí se sugiere la implementación de estas proteínas en procesos tempranos de tratamiento para identificar patrones de transformaciones celulares malignas o progresos tumorales (De Silva, Rapior, Hyde, & Bahkali, 2012, Erjavec *et al.* 2012).

## 3. Materiales y métodos

### 3.1. Material Biológico

El basidioma de *G. australe* fue colectado en la unidad experimental de la Universidad del Valle y depositado en la colección del herbario CUVC de la misma universidad, con la colaboración de la doctora Ana Cristina Bolaños. El número de identificación asignado para depositar la muestra en el herbario fue 374 para depositar la muestra en el herbario. La Universidad del Vallen cuenta con permiso de colecta 1070, expedido por la agencia nacional de licencias ambientales (ANLA).

#### 3.1.1. Cultivo del Hongo

La cepa de *G. australe* se cultivó en el laboratorio de Biología Molecular de la Universidad de Manizales. El hongo se cultivó en agar Sabouraud dextrosa (SDA) durante 8 a 10 días a 25°C para estimular el crecimiento micelial. Posteriormente, se realizó el inóculo del micelio en frascos Schott (Duran) de 250 mL, que contenían 50 mL de Extracto de Malta. Se tomaron de cinco a diez discos de aproximadamente 90 mm de diámetro del medio sólido colonizado con el micelio, y se transfirieron al medio con Extracto de Malta. Los frascos Schott de 250 mL con el inóculo del hongo se ubicaron en una incubadora con agitación horizontal Labnet a 27°C y 100 rpm durante 15 días para obtener la mayor cantidad de biomasa necesaria para la extracción de ADN genómico.

#### 3.1.2. Extracción y cuantificación de ADN genómico

La extracción del ADN genómico (ADNg) de *G. australe* fue efectuada también en el laboratorio de biología molecular de la Universidad de Manizales. Para la extracción del ADNg, se tomó el micelio crecido del cultivo líquido. Se realizaron seis lavados del micelio con cloruro de magnesio (MgCl<sub>2</sub>), para limpieza y eliminación de polisacáridos y desechos celulares del crecimiento del hongo en el medio de cultivo que pudiesen interferir con la extracción de ADNg. Posteriormente, el micelio fue macerado con nitrógeno líquido (N<sub>2</sub>) y la extracción de ADNg se realizó utilizando el kit Wizard® Genomic DNA purification (Promega®), siguiendo la metodología descrita por el fabricante. Una vez obtenido el ADNg, se verificó la calidad mediante electroforesis en gel de agarosa al 0,8% p/v teñido con el colorante GelRed™, empleando el ADN del bacteriófago Lambda como referencia. Finalmente, se determinó la concentración y calidad del ADNg mediante un espectrofotómetro Nanodrop a longitudes de onda 260/280 y 260/234, y dos fluorómetros, Glomax® de Promega® y Qubit de ThermoFisher.

#### 3.1.3. Identificación molecular de *G. australe*

Se realizó una identificación molecular para el hongo por medio de la amplificación de la secuencia conservada ITS (Internal Transcribed Sequence) del ADN ribosomal, empleando los

indicadores ITS 4 y 5 de acuerdo a la metodología realizada por Park *et al.* (2012). Los productos de amplificación fueron secuenciados y se realizó una búsqueda por homología en la base de datos del NCBI contra los ITS reportados para *Ganoderma*, con el fin de confirmar la identidad del hongo con *G. australe*.

#### 3.1.4. Secuenciación genómica

La secuenciación del genoma de *G. australe* se realizó mediante la tecnología de secuenciación de nueva generación (*Next Generation Sequencing*, NGS) PacBio® en la plataforma Sequel. La preparación de las librerías y la secuenciación de las mismas se realizaron mediante el servicio contratado con el centro de secuenciación GENEWIZ Inc. Se prepararon las librerías SMRTbell con fragmentos de 20 kb, utilizando el marcador de selección Blue Pippin Size Selection, y éstas se secuenciaron en el equipo PacBio Sequel®, con una duración de la película de 6 horas en dos celdas SMRTCell de PacBio, a una profundidad de 100X con base en un tamaño de genoma estimado de 40 Mbp, tomando como referencia el genoma de *G. lucidum*. El resumen de la metodología de cultivo en el laboratorio y secuenciación del ADNg se puede observar en la Figura 3.

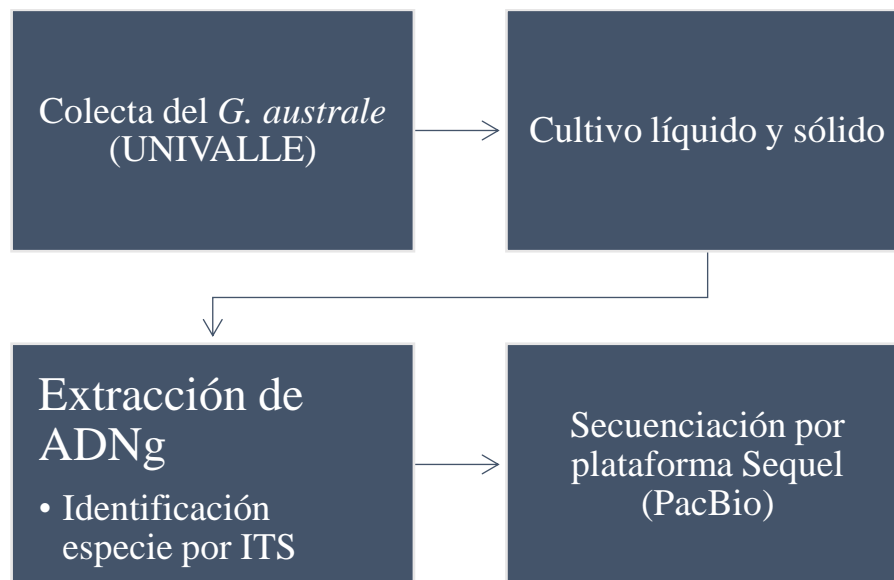


Figura 3. Metodología descrita para el cultivo de *G. australe* y secuenciación del ADNg del hongo.

#### 3.2. Ensamblaje y anotación del genoma

Las secuencias crudas del proceso de secuenciación se recibieron de parte del proveedor y se descargaron al clúster de computación de alto rendimiento (HPC) del Centro de Bioinformática y Biología Computacional de Colombia (BIOS). Los datos crudos del genoma se ensamblaron usando la herramienta Canu (Koren *et al.* 2017) utilizada para ensamblar lecturas largas de PacBio. Esta

herramienta utiliza un pipeline de corrección, corte y ensamblaje propio para lecturas largas de tecnologías de alto ruido como PacBio. El archivo de salida del ensamblaje por Canu se sometió a un mejoramiento por parte del módulo Arrow que hace parte de la herramienta SMRTLINK versión 4, con el fin de obtener un consenso del ensamblaje, alineando los datos crudos de la secuenciación a esta salida, y así, dar una mayor extensión al ensamblaje de Canu para generar scaffolds completos, y dar un complemento de soporte en la corrección de errores que ejecuta la herramienta de ensamblaje.

Las estadísticas de calidad del ensamblaje se obtuvieron utilizando Quast (Quality Assessment Tool for Genome Assemblies) (Gurevich *et al.* 2013), y la calidad del ensamblaje en cuanto a representatividad de genes conservados del linaje Basidiomycota se evaluó usando la herramienta Benchmarkin Universal Single-Copy Orthologs (BUSCO) (Simao 2015).

Cuando se obtuvo el ensamblaje *de novo* con Canu y refinado con las demás herramientas mencionadas, se procedió a anotar el genoma usando la metodología descrita en el pipeline de Maker2 (Holt & Yandell 2011), y así realizar la predicción del contenido génico del genoma a partir de predictores *ab initio* y evidencias proteicas de *Ganoderma* spp. reportadas en las bases de datos del Mycocosm (Grigoriev *et al.* 2013) y del National Center for Biotechnology Information (NCBI). El análisis de las regiones repetitivas del genoma (repeticiones interesparcidas y secuencias de ADN de baja complejidad) se evaluaron por fuera del pipeline de Maker2, usando la herramienta RepeatMasker v4 (<http://www.repeatmasker.org/>), asumiendo que los transposones estaban asociados al grupo taxonómico (Reino) Fungi.

Se utilizó el servicio de anotación automática del KEGG (KAAS) (Moriya *et al.* 2007) para la búsqueda y asociación de rutas metabólicas dentro del genoma, como una medida de cuán completo y correctamente ensamblado era el genoma obtenido. Para evaluar el ensamblaje del genoma con la anotación, se tomó como referencia la ruta metabólica del ácido ganodérico descrita para *G. lucidum* reportada por Liu *et al.* (2012), y se evaluó la presencia de genes relacionados con las enzimas claves de la ruta metabólica: farnesil difosfato sintasa (FDS) (EC 2.5.1.10), squaleno sintasa (SS) (EC 2.5.1.21), squaleno monooxigenasa (SQM) (EC 1.14.14.17) y lanosterol sintasa (LS) (EC 5.4.99.7). También se hizo una búsqueda de complejos enzimáticos que pertenecieran a citocromo P450 (CYP450) y al complejo uridindifosfato-glicosil-transferasas (UGT), con el fin de determinar si esta ruta se encontraba anotada de manera completa y correcta. Se utilizaron las secuencias de estas enzimas que estuvieran reportadas en las bases de datos del NCBI y UniProt para el filo Basidiomycota.

### 3.3. Anotación de proteínas FIPs

Se construyó una base de datos local con las secuencias proteicas asociadas a las enzimas FIPs reportadas y descargadas de las bases de datos UniProt (proteínas) y NCBI (cds-mRNA). Las secuencias se alinearon en el programa en línea Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) en formato Estocolmo. Con esto se construyeron los modelos para realizar búsquedas de las FIPs con el software HMMER v3.1b2 (Eddy 2011, Finn, Clements & Eddy 2011). Se realizaron también búsquedas por homología con BLASTP local.

La secuencia putativa de la FIP de *G. australe* se aisló del genoma mediante con las herramientas cdbfasta (<https://github.com/gpertea/cdbfasta>) y exonerate (Slater & Birney 2005) (<http://www.ebi.ac.uk/about/vertebrate-genomics/software/exonerate>). Se realizó una comparación con las secuencias reportadas en la base de datos de UniProt, donde se realizó un alineamiento en la herramienta Mega7 (Kumar, Stecher, & Tamura, 2016) con la herramienta MUSCLE.

La secuencia putativa se evaluó en el servidor de InterProScan en línea (<http://www.ebi.ac.uk/interpro/>) para búsqueda de dominios, y en Fgenesh y Promoter 2 (versiones en línea) para determinar la estructura del gen ([http://www.softberry.com/berry.phtml?topic=case\\_study\\_plants&no\\_menu=on](http://www.softberry.com/berry.phtml?topic=case_study_plants&no_menu=on)). La secuencia putativa de la FIP se llevó al servidor en línea de I-Tasser (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>), el cual predice la estructura secundaria y terciaria de secuencias de proteínas mediante un algoritmo que utiliza la base de datos de Protein Data Bank (PDB), junto a los reportes del consorcio GO y el Enzyme Commission Number (EC) para predecir un modelo consenso usando un estadístico llamado Cscore, el cual está en el rango de -5 (más bajo) a 2 (más alto) en niveles de similitud (Roy, Kucukural & Zhang 2010, Roy, Yang & Zhang 2012). La metodología completa para el ensamblaje y anotación del genoma de *G. australe* se puede observar en la Figura 4.

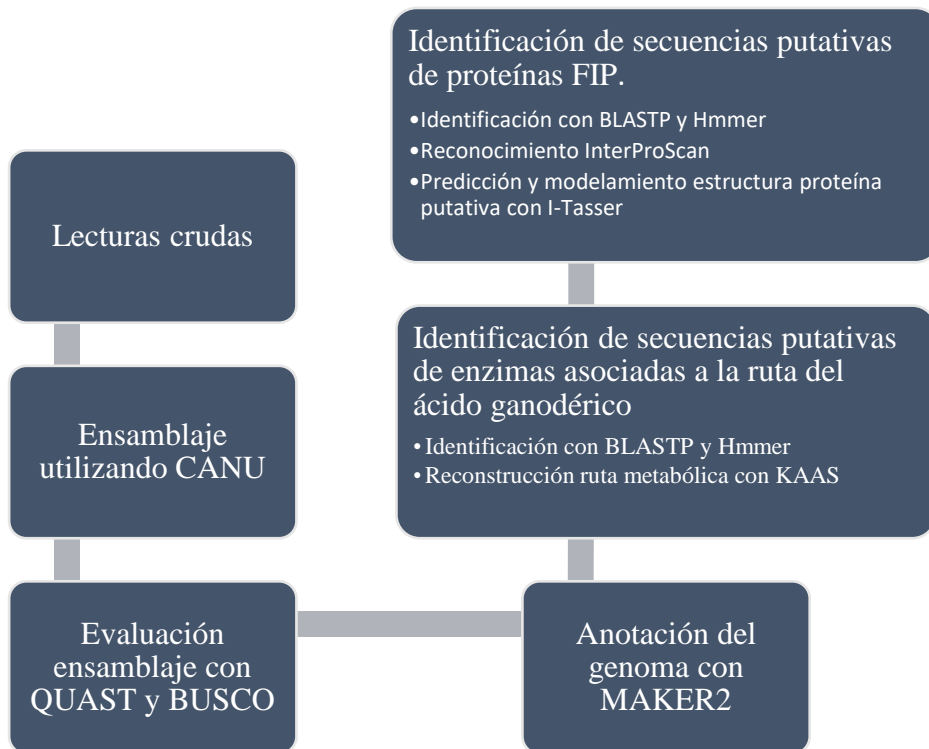


Figura 4. Metodología descrita para el ensamblaje y anotación del genoma de *G. australe*

## 4. Resultados y Discusiones

### 4.1. Obtención del ADN y secuenciación del genoma

El ADNg de *G. australe* se obtuvo de la extracción con el kit de Promega. El ADNg fue cuantificado mediante electroforesis de geles de Agarosa, donde se obtuvo un valor aproximado de la concentración de éste mediante la comparación del ADNg con la banda correspondiente al ADN del bacteriófago lambda de cantidad conocida. Los resultados de la corrida del gel de agarosa por electroforesis se observan en la Figura 5.

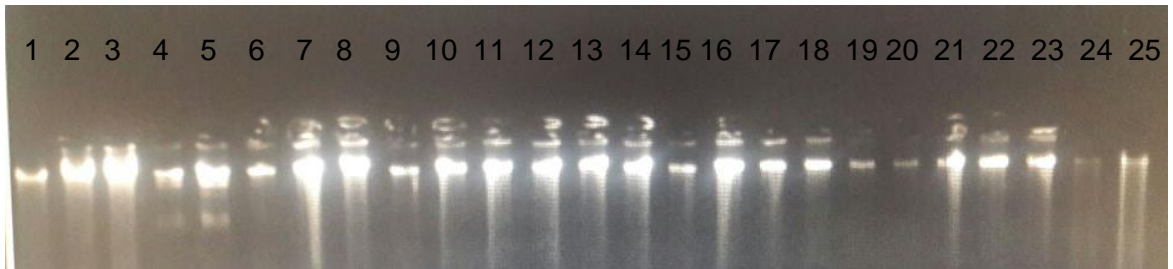


Figura 5. Electroforesis del ADNg extraído del micelio de *G. australe*. Pozos 1-3: ADN lambda a 15 ng, 30 ng y 45 ng; pozos 4-6: ADN lambda a 15 ng, 30 ng y 45 ng con HindIII; pozos 7-25: ADNg de *G. australe*.

Posteriormente, se evaluaron las calidades y concentración del ADNg mediante los equipos NanoDrop y Glomax/Qubit entre el laboratorio de la universidad de Manizales y el centro de secuenciación GeneWiz. Las muestras de ADNg se unificaron en 5 tubos y se cuantificaron de esa manera. Los resultados obtenidos se observan en la

Tabla 3.

Tabla 3. Concentración de ADNg cuantificados con el NanoDrop y Glomax/Qubit.

Muestra ADNg	Nanodrop(ng)	Glomax/Qubit(ng)
	[concentración ácidos nucleicos]	[concentración ADNg]
1	35.000	3.335
2	34.000	4.393
3	51.000	3.887
4	54.000	6.480
5	80.000	2.393



#### 4.1.1. Identificación molecular de *G. australe*

A partir de la extracción del ADNg del hongo, se procedió a realizar la identificación molecular del hongo utilizando las regiones de ITS 4 y 5 del ADN ribosomal. Se analizó la secuencia obtenida de la secuenciación por medio del programa CodonCode Aligner™, en la cual se obtuvo una banda de aproximadamente 650 pb (ver Figura 6) y de la cual se obtuvo una homología del 99% para *G. australe* en la secuencia reportada en el GenBank (*G. australe* voucher HMAS86596 GenBank: AY884180.1) para este tipo de ITS. Park *et al.* (2012) realizaron la identificación de 32 especies de *Ganoderma*, entre ellas, *G. lucidum*, *G. tsugae*, *G. tornatum*, utilizando las regiones de ITS 4 y 5 del ADN ribosomal. En los resultados presentados por los investigadores, obtuvieron que las bandas de los ITS 4 y 5 para los hongos de *Ganoderma* se encuentran en un rango entre 625 a 673 pb. A partir de esto, se confirmó que la región amplificada para *G. australe* estaba dentro del rango de los ITS reportados por Park *et al.* para *Ganoderma*, y que, además, fue bien alineada e identificada para la especie en un 99% para los ITS que están reportados en las bases de datos.

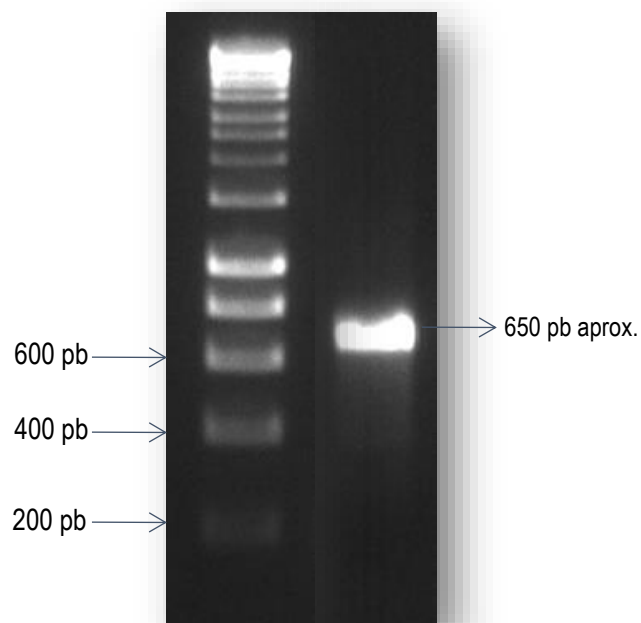


Figura 6. Identificación molecular de *G. australe* usando las regiones ITS 4 y 5 del ADN ribosomal.

El ADNg fue secuenciado por la tecnología Sequel de PacBio, en el cual se obtuvieron alrededor de seis gigabases de lecturas para las dos celdas de SMRTcell, las cuales se pueden observar en el Anexo I.

#### 4.1.2. *Ensamblaje del genoma*

Los datos de secuenciación se descargaron al clúster HPC del Centro de Bioinformática y Biología Computacional de Colombia – BIOS para ser analizados. Los datos crudos del genoma se ensamblaron siguiendo la metodología de Koren *et al.* (2017) para la herramienta Canu (Ver Anexo I, B). Se obtuvo un tamaño del genoma consenso total de 84 Mbp con la salida de la herramienta Arrow, con un porcentaje de GC de 55,57 %. Este valor representa un genoma de aproximadamente el doble del tamaño del reportado para la especie *G. lucidum* de 43 Mbp (Chen *et al* 2012; Lui *et al* 2012). El mayor scaffold generado del ensamblaje fue de 1,66 Mbp y el N50 estuvo alrededor de los 200 kbp. (Ver Tabla 4) Esto valores visualizan una excelente continuidad del genoma ensamblado ya que poseen pocos scaffolds (776) y que tienen una continuidad muy alta, generando scaffolds de tamaños mayores a 1 Mbp. El reporte final generado con la herramienta QUAST del ensamblaje se presenta en la Tabla 4.

Además de la herramienta Canu, se evaluaron otras estrategias de ensamblaje para el genoma de *G. australe*, que incluyeron los ensambladores *de novo* Falcon (Chin *et al* 2016) y HGAP (Chin *et al* 2013). Sin embargo, los resultados obtenidos con estos programas no fueron satisfactorios, pues se presentaron muchos fallos en la ejecución de estas herramientas dentro del clúster HPC de Bios, los cuales no fue posible subsanar. De todas maneras, el ensamblaje realizado con Canu fue ejecutado eficientemente en términos del rendimiento de la máquina (clúster HPC); además, se trata de un software ampliamente recomendado y utilizado frecuentemente para ensamblajes de tecnologías provenientes de secuenciación de “*long reads*” como PacBio y Oxford Nanopore. Además, permitió un ensamblaje óptimo *de novo* del genoma en términos de N50, estimación del contenido de GC y continuidad en los scaffolds.

Tabla 4. Reporte con QCAST de estadísticas del ensamblaje del genoma de *G. australe*.

<b>Parámetro</b>	<b>Valor</b>
<b># contigs (&gt;= 0 bp)</b>	776
<b># contigs (&gt;= 1000 bp)</b>	776
<b># contigs (&gt;= 5000 bp)</b>	776
<b># contigs (&gt;= 10000 bp)</b>	776
<b># contigs (&gt;= 25000 bp)</b>	755
<b># contigs (&gt;= 50000 bp)</b>	412
<b>Total length (&gt;= 0 bp)</b>	84.2 Mb
<b>Total length (&gt;= 1000 bp)</b>	84.2 Mb
<b>Total length (&gt;= 5000 bp)</b>	84.2 Mb
<b>Total length (&gt;= 10000 bp)</b>	84.2 Mb
<b>Total length (&gt;= 25000 bp)</b>	83.8 Mb
<b>Total length (&gt;= 50000 bp)</b>	71.1 Mb
<b># contigs</b>	776
<b>Largest contig</b>	1.66 Mb
<b>Total length</b>	84.2 Mb
<b>GC (%)</b>	55.57
<b>N50</b>	200.492
<b>N75</b>	70.371
<b>L50</b>	94
<b>L75</b>	276
<b># N's per 100 kbp</b>	0.00

El ensamblaje obtenido por la herramienta Canu se analizó con la herramienta BUSCO para evaluar la presencia de los genes ortólogos conservados del filo Basidiomycota (ver Tabla 5). La herramienta ejecutó una búsqueda de genes del linaje Basidiomycota, con base en modelos génicos de las especies *Corpinus rutilus*, *Laccaria bicolor* y *G. lucidum* como referencia para la predicción *ab initio* de genes con el programa Augustus, integrado en BUSCO. Los mejores valores de representatividad de genes conservados se obtuvieron con base en los modelos génicos de *Coprinus* y *G. lucidum* (ver Tabla 5) con un 96% y 95,8% de genes conservados encontrados en relación al

linaje de Basidiomycota. Es de resaltar que la búsqueda de genes con BUSCO presentó un alto valor de genes duplicados, alrededor de los 276 genes duplicados.

Tabla 5. Resultado de BUSCO para el ensamblaje de *G. australe*.

<b>Ensamblaje</b>	<b>Modelo génico Augustus (sp)</b>	<b>Genes encontrados</b>	<b>Completos</b>	<b>Duplicados</b>	<b>Faltantes</b>	<b>Genes de referencia en el linaje</b>
<b>Canu</b>	<b><i>Coprinus</i></b>	<b>1282</b>	<b>1006</b>	<b>276</b>	<b>23</b>	<b>1335</b>
<b>Canu</b>	<b><i>Laccaria</i></b>	<b>1252</b>	<b>987</b>	<b>265</b>	<b>17</b>	<b>1335</b>
<b>Canu</b>	<b><i>G. lucidum</i></b>	<b>1279</b>	<b>1007</b>	<b>272</b>	<b>28</b>	<b>1335</b>

Las estadísticas del ensamblaje obtenido para *G. australe* fueron contrastadas con los trabajos de Chen *et al* (2012) y Lui *et al* (2012) en los cuales reportaron el genoma completo de *G. lucidum* de 43 Mbp. Se observó que una gran diferencia en el tamaño genómico entre estas especies, ya que *G. australe* dobla el tamaño del genoma de *G. lucidum*. Se destaca que hay una carencia de trabajos genómicos publicados para la especie de estudio, por tanto, su genoma fue analizado con base al reporte de *G. lucidum*. Además, se analizó el genoma ensamblado en comparación con un proyecto que se hizo público por parte del Broad Institute (Kües *et al.*2015), el cual consistió en un estudio genómico de cuatro especies de *Ganoderma*, y del cual se sólo se presentó un reporte a la base de datos del NCBI para el genoma de *G. boninense* (Acceso: GCA\_001855635.1). Este genoma posee un tamaño de 60 Mbp aproximadamente (24 Mbp de diferencia con respecto a *G. australe*), y el cual fue ensamblado también con la herramienta Canu a partir de datos de PacBio RSII.

Los genomas de *G. lucidum* y *G. boninense* fueron comparados frente al genoma ensamblado de *G. australe* mediante al herramienta Mummer v3 (Kurtz *et al.* 2004), presentado en el Anexo II. Al observar la comparación de los genomas de *G. australe* vs *G. lucidum* (Anexo II, A), se pudo observar que los scaffolds de *G. australe* no pudieron ser organizados de forma correcta frente al genoma de *G. lucidum*. Además, hay un 80% de similitud entre las secuencias comparadas, dejando una franja que no pudo ser comparada en la parte superior de la gráfica. Sin embargo, podemos ver que los dos genomas están evaluados en casi su totalidad.

No sucedió lo mismo con la comparación de los genomas de *G. australe* vs *G. boninense* (Anexo II, B). Cuando se realizó la evaluación de estos genomas se pudo visualizar una mejor organización entre los scaffolds de los genomas, los cuales poseen un 80% de similitud, aunque se evidencia una mayor brecha entre los scaffolds en una región del genoma de *G. boninense*, como se aprecia en la figura (Anexo II, B).

Al analizar las secuencias de ITS halladas en el estudio de *G. australe*, y las secuencias reportadas en el NCBI para los genes de ADN ribosomal 18S, ITS1-5.8S-ITS2 y 28S, se descargaron las secuencias para observar las distancias filogenéticas que existen entre las especies de *G. australe*, *G. lucidum* y *G. boninense*. Se utilizaron, además, cuatro especies de *Ganoderma*, *G. sinense*, *G. tsugae*, *G. tornatum* y *G. applanatum*, y una especie como grupo externo, *S. cerevisiae* para enriquecer los métodos de comparación filogenética. Las secuencias fueron alineadas utilizando la herramienta MUSCLE dentro de Mega 7 (Kumar et al., 2016), y se ejecutaron los métodos de máxima parsimonia, Neighbour-Joining, UPGMA y máxima probabilidad para visualizar la filogenia de estas secuencias (ver Anexo II, C-D-E-F). Cada una de las secuencias de marcadores de ADN ribosomal e ITS evaluadas con Mega evidenciaron que las especies de *G. australe* y *G. applanatum* están con mayor cercanía en cuanto a características taxonómicas moleculares que *G. lucidum* y *G. boninense* como ya se habían mostrado en otros estudios (Jargalmaa et al., 2017; Park et al., 2012; Richter, Wittstein, Kirk, & Stadler, 2015), y por lo tanto, se presume que las diferencias evolutivas entre estas tres especies pudieron implicar cambios importantes a nivel genómico, que se reflejan en estas distancias taxonómicas (Jargalmaa et al., 2017). Existen antecedentes que demuestran que las diferencias evolutivas entre especies de un mismo género, eventualmente, podría también verse reflejadas en el tamaño de los genomas ((Kurtz et al. 2004 Anexo II, A-B)

Mohanta & Bae (2015) realizaron un reporte detallado de varias especies de hongos que poseían datos genómicos, principalmente de los filos Ascomycota y Basidiomycota. Para *Ganoderma*, los investigadores estimaron un genoma cercano a los 40 Mbp. Se observó, sin embargo, que el genoma ensamblado dista de esta estimación, y que se debe este tamaño genómico obtenido. Desde una perspectiva filogenética, *Ganoderma* es un clado taxonómico complejo que presenta principalmente datos morfológicos para su clasificación, careciendo aún de datos moleculares que aporten a determinar las especies dentro del clado (Richter, Wittstein, Kirk, & Stadler 2015).

*G. australe* es una especie que se encuentra filogenéticamente distante de *G. lucidum* ya que las dos especies pertenecen a distintos subgéneros dentro de *Ganoderma* (Park et al. 2012, Zhou et al. 2016, Jargalmaa et al. 2017), y presentan diferencias en años evolutivos, acumulación de información génica y número de generaciones presentes para cada especie (Hardison 2003, Miller et al. 2004). Es por esto, y observando esta información taxonómica, que puede resultar no confiable la

comparación entre los tamaños genómicos dentro del clado de *Ganoderma*, con base en los estudios moleculares de *G. lucidum*, que ha sido ampliamente estudiado (Park *et al.*, 2012; Zhou, Liu, Guo, Su, & Zhang, 2016), pero que deja un vacío importante para las especies que están en el subgénero *Eflvingia* como *G. australe*, *G. applanatum*, *G. atrum*, *G. boninense* para el ámbito genómico.

Además, teniendo en cuenta los resultados de BUSCO (ver Tabla 5), se observa que el genoma ensamblado de *G. australe* presenta un alto número de genes duplicados (>270) para los genes ortólogos del filo Basidiomycota, lo que puede indicar la presencia de repeticiones de haplotipos, reorganizaciones genómicas que generasen duplicaciones, fusiones genómicas o algún otro proceso que pudiese haber ocurrido dentro de este genoma, y que no se detecten con las herramientas bioinformáticas más utilizadas en la actualidad como BUSCO, limitando la determinación de este tipo de datos. Además, el análisis de BUSCO es limitado para este tipo de hongos por la poca información existente para ejecutar este tipo de análisis. Por consiguiente, se destaca la poca información genómica que hay disponible para el género *Ganoderma* en las bases de datos, siendo escasos los trabajos que han dado un reporte completo de los datos genómicos para especies que estén dentro de *Ganoderma*.

## **4.2. Anotación del Ensamblaje**

### **4.2.1. Anotación del ensamblaje con Maker2**

Se realizó la anotación del genoma con la herramienta Maker2, utilizando como modelo génico el reporte de *G. lucidum*. Los perfiles de configuración de corrida de Maker2 se pueden observar en el Anexo III, A-B

Se obtuvieron 10.713 modelos de genes predichos dentro del genoma de *G. australe*, en comparación con estudios de Chen *et al* (2012) y Lui *et al* (2012), que obtuvieron 12.000 secuencias modelos para el genoma de *G. lucidum*. El valor obtenido en la predicción de genes en el genoma de *G. australe* es menor al reportado para *G. lucidum*, pero se debe tener en cuenta la novedad de que *G. australe* no posee hasta la fecha un gran catálogo de secuencias reportadas en las bases de datos y, por ende, es probable que exista un sesgo atribuido al modelo génico para la predicción de genes basado en *G. lucidum*. Al respecto, habría que considerar una metodología que complemente este tipo de aproximaciones para genomas ensamblados *de novo* y que determinen los datos no encontrados para establecer los vacíos de información que se pudieran haber generado con esta aproximación.

Las secuencias predichas de proteínas de Maker2 se anotaron en el sistema de anotación automática de KEGG (KAAS), obteniendo una anotación de 7.709 secuencias que obtuvieron una relación funcional en cuanto a su actividad biológica en la base de datos (ver Anexo III, C). Las secuencias más representativas de la anotación estuvieron en el grupo de procesamiento de la información genética, la cual obtuvo una anotación de 2501 secuencias, lo que equivalió al 32% de

la anotación como se muestra en el diagrama circular (ver Anexo III, B ). Otros grupos importantes que se determinaron en la anotación de KAAS fueron procesos celulares (11%), procesamiento de la información ambiental (9%), familia de enzimas (5%) entre otros. Lui *et al* (2012) determinaron en la anotación del genoma de *G. lucidum* que sus proteínas predichas estaban en mayor cantidad relacionadas con procesos de metabolismo de carbohidratos, aminoácidos y xenobióticos. En cambio, en la anotación de *G. australe* estos grupos presentan una baja representatividad (<5%), dando a entender que se necesitan más estrategias de anotación de genes (ej. Blast, GO) para obtener una mejor precisión en la determinación de los genes que están en el ensamblaje del genoma. Aun así, se logró determinar por esta metodología de anotación la identificación de las secuencias putativas pertenecientes a la ruta metabólica de los ácidos ganodéricos.

Continuando con el análisis del genoma de *G. australe*, se determinaron las regiones asociadas a transposones en el ensamblaje con la herramienta RepeatMasker. En principio, la herramienta no fue posible ejecutarla dentro del pipeline de Maker2 debido a problemas de ejecución en el clúster HPC de BIOS, por lo que se decidió hacer la corrida por separado del pipeline. Se reconocieron 2,6 Mbp regiones enmascaradas, lo que corresponde al 3,18%. La mayor parte de estos elementos enmascarados corresponden al Retroelementos (2,18%), siendo los transposones LTR los más identificados (2,13%) (ver Anexo III, D).

Liu *et al* (2012) analizaron las regiones repetitivas en el genoma de *G. lucidum* de 39 Mbp, donde se encontraron con un 5% de regiones repetitivas en ese genoma, que equivalían a 2 Mbp. Sin embargo, Chen *et al* (2012) reportaron un mayor número de regiones repetitivas para el genoma de *G. lucidum*. Ellos reportaron para el genoma de 43 Mbp un 8% de regiones repetitivas que correspondían a 3,5%. Además, Kües *et al* (2015) ensamblaron un genoma de *G. lucidum* 39 Mbp, obteniendo 2,1% de regiones transponibles para ese genoma generado.

Comparado con ese resultado, el genoma de *G. australe* posee menos regiones repetitivas con respecto a *G. lucidum*, pero coinciden en que los transposones LTR son las regiones más encontradas en estos análisis. Otros análisis de regiones transponibles se deben efectuar cuando se efectúen otras estrategias de ensamblaje (ej. Falcon) u otras estrategias de secuenciación, con los cuales se pueda comparar la determinación de estos elementos para el genoma de *G. australe*.

A partir de las secuencias codificantes y proteínas predichas con Maker2, se evaluó la presencia de las enzimas claves de las rutas metabólicas dentro del genoma de *G. australe* y la identificación de secuencias putativas de proteínas FIP.

#### 4.2.2. Evaluación de la ruta metabólica del ácido ganodérico

Para la evaluación de la ruta metabólica del ácido ganodérico, se efectuó una búsqueda con las herramientas BLASTP versión 2.5 y Hmmer versión 3.1b2 para encontrar las enzimas claves de la ruta del ácido ganodérico en las proteínas predichas dentro del genoma de *G. australe*.

Las secuencias putativas encontradas por estas herramientas para la ruta metabólica del ácido ganodérico se llevaron a la herramienta de predicción de genes Fgenesh (Solovyev *et al.* 2006) para determinar la longitud del transcrito de cada enzima y su equivalente en aminoácidos. Estas secuencias también se llevaron al servidor en línea InterProScan (<http://www.ebi.ac.uk/interpro/>) para determinar los dominios proteicos dentro las secuencias de aminoácidos predichas, con el fin de corroborar la anotación funcional de las enzimas de la ruta metabólica. Todas las secuencias de las enzimas putativas obtuvieron algún resultado para las secuencias reportadas en las bases de datos (ver

Tabla 6), y para los dominios de proteínas con InterProScan. Las enzimas claves dentro de la ruta metabólica del ácido ganodérico como la FDS o la LS, presentaron una alta homología (evalue >90), y que fue confirmado con la identificación de los dominios proteicos para estas enzimas (ver

Tabla 6). Además, las asignaciones de las funciones celulares dentro del Gene Ontology (GO) fueron adecuadas para dar un atributo biológico que cumple cada enzima dentro de la ruta metabólica.

Tabla 6. Enzimas evaluadas para la ruta del ácido ganodérico dentro del genoma ensamblado de *G. australe*.

Enzima	NÚMERO aminoácidos	BLASTP / Hmmer	Scaffold	Interproscan	GO
<b>FDS</b>	376	5.67e-100	tig00000137 arrow	Polyprenyl Synthetase	0008299
<b>SS</b>	360	3.4e-90	tig00000176 arrow	Squalene synthase	0009058, 0016740, 0016765
<b>SQM</b>	466	8e-70	tig00001257 arrow	No predicho	0055114, 0004506, 0016491, 0050660,0016021
<b>LS</b>	784	1.34e-100	tig00001248 arrow	Squalene cyclase	0003824, 0016866



FDS: farnesil difosfato sintasa; SS: squaleno sintasa; SQM: squaleno monooxigenasa; LS: lanosterol sintasa; GO: gene ontology.

Las secuencias putativas obtenidas se llevaron al servicio de anotación de genomas del KEGG (KAAS). A partir de la anotación en el servidor de KAAS, se utilizó una opción para reconstruir rutas metabólicas, en especial, la ruta metabólica de los triterpenos (ver Figura 7), en donde están contenidas las enzimas de la ruta metabólica del ácido ganodérico. Se realizó la anotación completa de las secuencias putativas por parte de KEGG, donde se detallan las rutas metabólicas asociadas a las secuencias putativas (ver Figura 7a). Después, se realizó la reconstrucción de la ruta del ácido ganodérico que está dentro de la biosíntesis de los triterpenos (ver Figura 7b), resaltando en color verde las enzimas encontradas dentro de las secuencias putativas que se introdujeron, indicando la presencia de ellas en la ruta de la biosíntesis de los triterpenos.

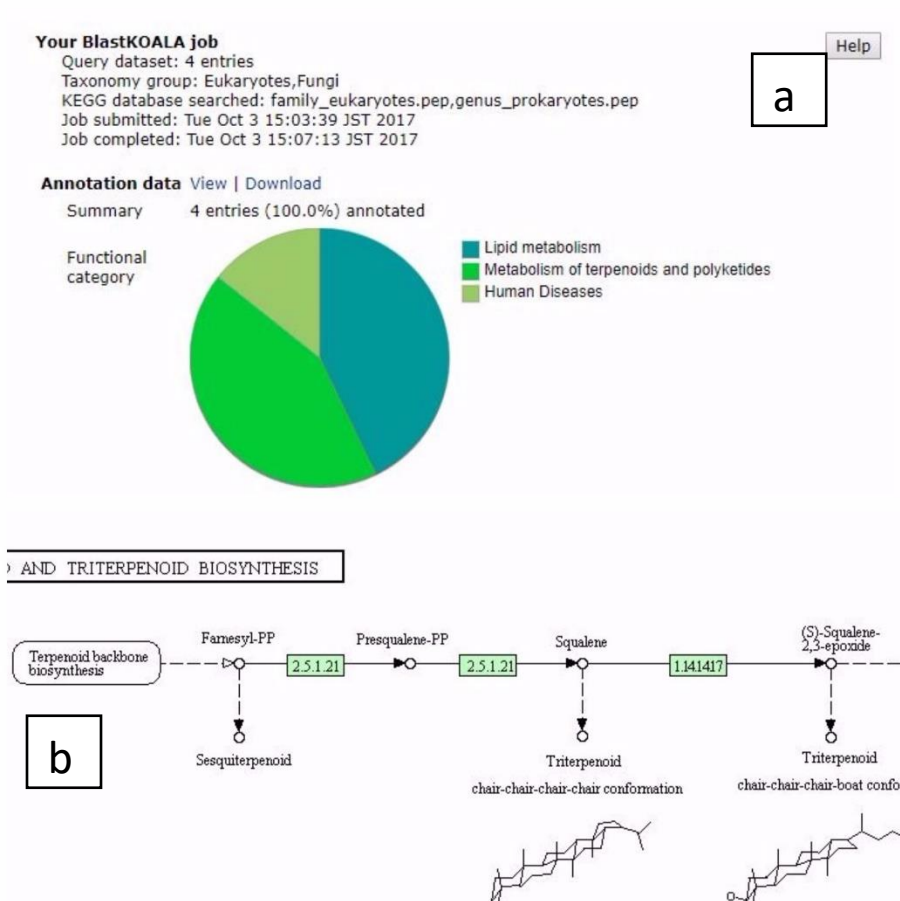


Figura 7. a. Resultado de la anotación de las secuencias putativas del genoma de *G. australe* con KEGG. b. Reconstrucción de las secuencias putativas dentro de la ruta de biosíntesis de triterpenos.

Investigaciones moleculares realizadas para las enzimas pertenecientes a la ruta de los ácidos ganodéricos, mostraron que se realizaron aislamientos y técnicas de clonación para las enzimas farnesil disfosfato sintasa (FDS) (Ding *et al.* 2008), squaleno sintasa (SS) (Zhao *et al.* 2007) y lanosterol sintasa (LS) (Shang *et al.* 2010) que estaban dentro de *G. lucidum*, y para la enzima squaleno monooxigenasa (SQM) (Godio, Fouces & Martín 2007) que estaba dentro de un hongo basidiomiceto *Hypholoma sublateritium*. En cada una de estas investigaciones se obtuvieron secuencias de proteínas de 360 aa para la enzima FDS, 468 aa para la enzima SS, 461 aa para la enzima SQM, y para la enzima LS una secuencia de 726 aa. Teniendo en cuenta estas investigaciones, se evaluaron esos resultados con los datos obtenidos dentro de la

Tabla 6, donde se observó que el genoma de *G. australe* posee variaciones en la longitud de las secuencias de estas enzimas, siendo algunas de mayor longitud en aminoácidos, en particular, la enzima FDS que posee 376 aa para *G. australe* en comparación a la enzima FDS de *G. lucidum* que posee 360 aa, y que es una enzima clave de la ruta metabólica por su actividad en la transformación de los sustratos del mevalonato para entrar en el paso de ciclación de la molécula metabólica al lanosterol. Dhar, Koul & Kaul (2013) describieron las características moleculares de la enzima FDS y su caracterización dentro de varias especies como ratones, bacterias, humanos y hongos, incluyendo a las levaduras *S. cerevisiae* y *S. pombe*. Estos autores enfatizaron que la estructura de la enzima en cada una de las especies descritas posee una gran variación en la estructura génica. Es por esto que podrían existir diferencias entre los genes de la enzima FDS entre *G. australe* y *G. lucidum*.

Además de la enzima FDS, las enzimas que derivan en la producción del ácido ganodérico SS, SQM y LS encontradas en *G. australe* presentaron longitudes de 360 aa, 466 aa y 768 aa con respecto a las enzimas identificadas y clonadas de *G. lucidum*, que fueron de 468 aa, 461 aa y 726 aa respectivamente. Estos hallazgos en *G. australe* revelaron una mayor o menor similitud en la longitud de la secuencia de las enzimas, dando la posibilidad de que exista una variación con las estructuras de los genes que codifican para cada proteína dentro de cada especie, y que se debe detallar más en futuros estudios. En consecuencia, no existen reportes de estas enzimas para el género *Ganoderma*, más allá del establecido con la especie *G. lucidum*, teniendo que ser remitidos a estudios realizados en especies modelos como *Arabidopsis thaliana*, *Escherichia coli* y *S. cerevisiae*. Concretamente para *S. cerevisiae*, se realizó un estudio por Lees, Skaggs, Kirsch & Bard (1995) donde efectuaron una clonación de cada enzima para la producción de ergosterol proveniente del genoma del hongo, determinando la longitud de las enzimas y obteniendo longitudes para la enzima SS de 444 aa, para la enzima SQM de 496 aa y para la enzima LS de 732 aa, infiriendo que la estructura génica de estas enzimas posee variaciones dentro de los genomas de cada especie que se estudió.

Al analizar las secuencias encontradas para el complejo de los CYP450, se hallaron 156 secuencias de genes putativas que estaban dentro del genoma ensamblado. Cheng *et al.* (2012) y Lui *et al.* (2012) determinaron la presencia de 220 secuencias de CYP450 dentro del genoma de *G. lucidum*, siendo las secuencias putativas encontradas dentro de *G. australe* un número menor al reportado para *G. lucidum*; no obstante, representa una primera aproximación de este complejo enzimático dentro del genoma de *G. australe*. Es importante mencionar que las secuencias contra las que se comparó este complejo enzimático para *G. lucidum* se relacionaron con estudios realizados en las especies *Postia placenta* y *Phanerochaete chrysosporium*, difiriendo de la búsqueda de CYP450 que se realizó dentro del genoma ensamblado de *G. australe*, la cual se efectuó con las secuencias de CYP450 reportadas en el NCBI para todo el filo Basidiomycota. Esto permitió a los investigadores encontrar más similitudes y diferencias en las familias de CYP450 que poseía *G. lucidum* dentro de su genoma en relación a las encontradas y caracterizadas en las otras especies mencionadas. Se necesita avanzar en la caracterización de los CYP450 encontrados en el genoma de *G. australe* y determinar la semejanza entre los CYP450 que se encontraron dentro de *G. lucidum*.

En relación al complejo enzimático de los UGT, se hallaron 26 secuencias de genes putativos pertenecientes a este grupo de enzimas. En estudios previos (Cheng *et al.* 2012 y Lui *et al.* 2012) se indicó que este complejo enzimático está involucrado en un paso clave para la producción de ácidos ganodéricos en las especies de *Ganoderma* por realizar la última fase de transformación de los triterpenos hacia los ácidos ganodéricos, debido a un proceso de glicosilación. Los autores indicaron que en el genoma de *G. lucidum* existe la presencia de seis secuencias que pertenecían a este complejo (Lui *et al.* 2012). Sin embargo, la búsqueda realizada en el genoma de *G. australe* se determinó con las proteínas reportadas en las bases de datos, y que necesitan una mayor caracterización para determinar similitudes con las proteínas descritas para *G. lucidum*.

### **4.3. Análisis de las proteínas FIPs**

Se realizó una búsqueda dentro del genoma ensamblado para las secuencias reportadas de las Proteínas Fúngicas Inmunomoduladoras (FIPs) en las bases de datos del NCBI. Se construyó una base de datos local con 15 secuencias reportadas para las proteínas FIPs (ver Anexo IV, A). Se obtuvieron 4 coincidencias (*match*) importantes con 3 secuencias provenientes del género *Ganoderma* y 1 del orden Polyporales, *C. rutilus* (ver Tabla 7.), presentando una alta homología (evalue > 60) y una identidad mayor al 80%. Estas coincidencias encontradas con Blast y Hmmer se hallaron dentro del mismo scaffold del ensamblaje (tig00001168). Igualmente, estas coincidencias se corroboraron con un análisis de homología utilizando Hmmer y BlastP con el archivo de proteínas predichas con Maker2, encontrando el mismo scaffold objetivo (tig00001168) (Anexo IV, B)

Tabla 7. Resultado de búsquedas por homología con BlastP y Hmmer para las proteínas FI P dentro del genoma de *G. australe*.

<b>Ensamblaje</b>	<b>blastp/ hmmmer</b>	<b>Scaffold</b>	<b>Organismo relacionado</b>	<b>Porcentaje identidad</b>
<b>Canu</b>	6.4e-66	tig00001168 arrow	<i>G. applanatum</i>	93.1%
<b>Canu</b>	2.5e-62	tig00001168 arrow	<i>C. rutilus</i>	92.6%
<b>Canu</b>	2.1e-53	tig00001168 arrow	<i>G. japonicum</i>	84.1%
<b>Canu</b>	2.9e-53	tig00001168 arrow	<i>G. atrum</i>	83.2%

El scaffold tig00001168 se aisló del genoma para evaluar la estructura del gen codificante para la FIP usando el predictor de genes Fgenesh en línea. Se obtuvo la predicción de una secuencia codificante de 342 bp (resaltada en naranja), la cual codificaría para una proteína 113 aa (ver Figura 8, CDS). Por medio del predictor de Fgenesh y de la herramienta en línea Promoter 2.0 (Knudsen 1999), se obtuvo la posición de la caja TATA, a 211 bp antes del ORF de la secuencia putativa (Figura 8, TSS y resaltada con amarillo - Anexo IV, C) y dos promotores que están ubicados antes del codón de inicio de la secuencia putativa. La herramienta Promoter identificó dos promotores CCAATC y CAATC (resaltados con color verde - Anexo IV, C) antes del codón de inicio de la secuencia codificante, resaltado en rojo (ver Anexo IV, C). Estos resultados son muy similares a los presentados por Murasugi *et al* (1991), donde se evaluó la estructura del gen de la FIP LZ-8 de *G. lucidum*. En este estudio localizaron dos regiones promotoras de CCAAT y una caja TATA para el gen de la FIP-*glu*. La diferencia que se presentó con la secuencia putativa del gen FIP de *G. australe* fue la presencia de uno de los promotores, ubicado después de la caja TATA (111bp antes del ORF del gen). En cambio, el segundo promotor está ubicado antes de la caja TATA, a una distancia de 88 bp. Además, Murasugi *et al* (1991) consideraron que la presencia y ubicación de los promotores en eucariotas estaban muy cercanas a la secuencia de la caja TATA, específicamente para hongos como levaduras, específicamente en *S. cerevisiae*, u hongos filamentosos (Struhl 1987), siendo inusual la presencia y ubicación de unos de los promotores después de la caja TATA. También, se observó, en la predicción de la estructura del gen de la FIP, el sitio de la poliadenilación del ARNm en la modificación post-transcripcional, y que estaría ubicado a 128 bp después del codón de parada de la secuencia putativa (resaltado con color azul) (ver Figura 8, PolyA y Anexo IV, C).

FGENESH 2.6 Prediction of potential genes in Laccaria genomic DNA  
 Seq name: test sequence  
 Length of sequence: 2000  
 Number of predicted genes 1: in +chain 1, in -chain 0.  
 Number of predicted exons 1: in +chain 1, in -chain 0.  
 Positions of predicted genes and exons: Variant 1 from 1, Score:46.966922

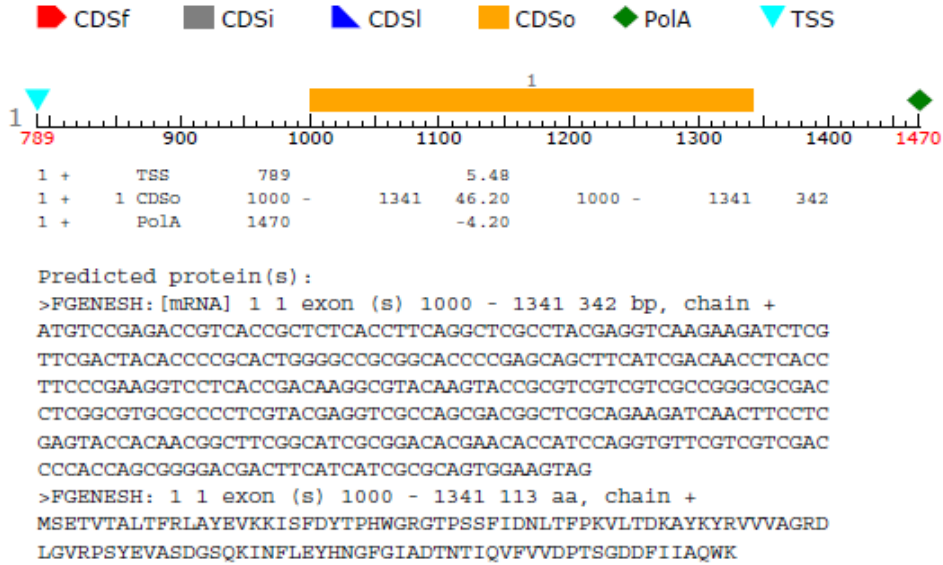


Figura 8. Estructura del gen putativo de la FIPs encontrado en el genoma de *G. australe* basado en la predicción con el programa Fgenesh.

Después de la predicción de la estructura génica de la FIP putativa, se realizó un alineamiento con las secuencias reportadas en el repositorio de UniProt, usando MUSCLE en Clustml Omega y Mega 7 (Kumar et al., 2016) (ver Anexo IV, D). Los alineamientos determinaron que la secuencia putativa posee una identidad muy alta (mayor al 90%) con las secuencias FIPs de *G. applanatum* y *C. rutilus*, el cual había sido determinado en el análisis de BlastP y Hmmer (ver Tabla 7).

La comparación de la secuencia putativa de la proteína FIP que se encontró dentro del genoma de *G. australe* mostró una alta identidad (85-93%) cuando se comparó con secuencias FIPs de especies del género *Ganoderma*, encontrándose cierta conservación en la secuencia para esta clase de proteínas dentro de este género. Cabe resaltar la alta similitud que mostró la secuencia putativa con la especie *Chroogomphus rutilus* (ver Tabla 7). Esta es una especie de Basidiomyceto que pertenece al orden Boletales, y que está distante filogenéticamente del orden Polyporales al que pertenece *G. australe* (en el análisis posee un 92% de identidad con la secuencia identificada, lo que la hace altamente similar a la secuencia putativa encontrada). Lin *et al.* (2016) aislaron y clonaron la FIP-cru de *C. rutilus* la cual presenta similitud con la FIP putativa de *G. australe* en cuanto su peso molecular y longitud de la secuencia obtenida. Al analizar la FIP-cru con respecto al clúster filogenético del

estudio, esta FIP fue más cercana a las secuencias del género *Ganoderma* y a la especie *Trametes versicolor*, infiriendo que provienen de un ancestro común por su agrupación en la misma clase filogenética y que conserva ciertas características en la secuencia dentro de este clado (Lin *et al.* 2016).

La secuencia putativa de la FIP se analizó con la herramienta EMBOSS en línea ([https://www.ebi.ac.uk/Tools/seqstats/emboss\\_pepstats/](https://www.ebi.ac.uk/Tools/seqstats/emboss_pepstats/)) para determinar las características del péptido predicho (ver Anexo IV, E). Se determinó que el péptido putativo posee un peso molecular de 12,7 KDa y un punto isoeléctrico de 6,53 para los 113 residuos de aminoácidos que conforman la secuencia. Al evaluar la secuencia putativa de *G. australe* respecto a la reportada para *G. applanatum*, se pudo observar una similitud en el peso molecular y la longitud de la secuencia (12,7 KDa y 113 aa). Yamashita & Hirose (2016) determinaron que el complejo de *G. australe* contiene a *G. applanatum* y que son especies cercanas filogenéticamente, agrupándose en el mismo clado taxonómico junto con *G. tornatum* y *G. adspersum* (Richter, Wittstein, Kirk & Stadler 2015), y que ayudaría a determinar la similitud de las secuencias FIP de las dos especies.

Además, las características determinadas para la FIP putativa de *G. australe* (Anexo IV, E) se asemejan a lo planteado por Wang *et al.* (2012), quienes introdujeron rasgos particulares de las proteínas FIP aisladas naturalmente del micelio y basidioma (cuerpo fructífero) de los hongos, como peso molecular, punto isoeléctrico y longitud de la secuencia de la proteína. Estos referentes para caracterizar las FIPs son claves en la identificación de este tipo de proteínas en estos estados de crecimiento, proporcionando una hoja de ruta fundamental para describir nuevas FIP en hongos aún no estudiados. Principalmente, las secuencias de FIPs se asemejan en un alto porcentaje en su estructura primaria (Uribe-Echeverry & López-Gartner 2017), como fue observado para las secuencias de *G. australe* y *G. applanatum*.

La secuencia putativa de la FIP se analizó con InterProScan y Pfam (ver Figura 9y Anexo IV,F) para la evaluación del dominio de la proteína. Se identificó el dominio canónico para la familia proteica de las FIPs llamado *Immunomodulatory protein FIP-Fve*, y se asoció la secuencia putativa con las ontologías de genes GO 0002682 y 0030246, que identifican la actividad biológica de la proteína (ver Figura 9). Cabe destacar que la ontología GO 0002682 es la única establecida para esta familia de proteína dentro del consorcio de Ontología de Genes.

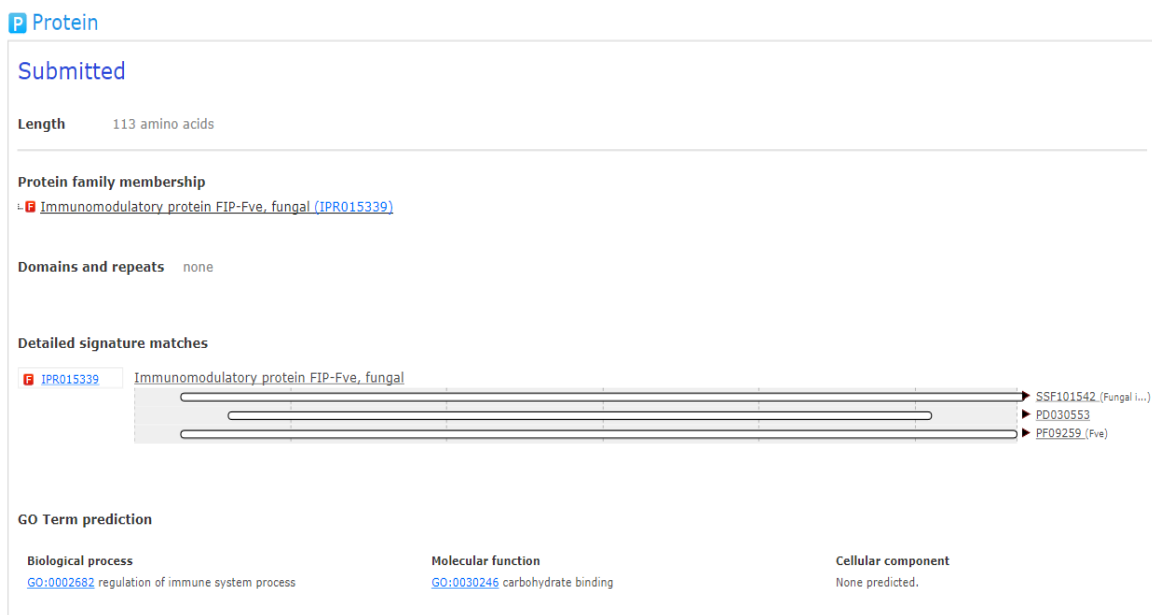


Figura 9. Predicción de dominios proteicos para la secuencia putativa de la FIP de *G. australe* con la herramienta InterProScan.

Por otra parte, la secuencia putativa de la proteína FIP se analizó con respecto a la predicción de estructura secundaria y terciarias en la herramienta en línea I-Tasser (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>). Este análisis se realizó para la secuencia putativa de *G. australe* (ver Figura 10 y Anexo IV, G), y para las secuencias reportadas de *G. applanatum* (ver Anexo IV, H) y *C. rutilus* (ver Anexo IV, H) por la similitud en la secuencias reportadas. Se determinó que la proteína putativa de *G. australe* está compuesta por dos hélices alfa y seis plegamientos beta en su estructura secundaria (ver Figura 10 y Anexo IV, G). El mismo procedimiento metodológico se realizó para la determinación de las estructuras secundarias para *G. applanatum* y *C. rutilus*, obteniendo estructuras muy similares a la de *G. australe*. con Cscore de 1.45 para *G. applanatum* y 1.4 para *C. rutilus*. También, al observar la estructura terciara predicha por I-Tasser, se comparó la tabla de la predicción de esta estructura con las proteínas FIP reportadas en el Protein Data Bank- PDB, en donde se logró un Cscore de 1.45 para el modelo de *G. australe*, lo que indica que el modelo predicho por la herramienta es muy similar a los comparados con la base de datos de PDB. Al realizar la determinación de las estructuras terciarias para *G. applanatum* y *C. rutilus*, se obtuvieron estructuras muy similares a la de *G. australe*. con Cscore de 1.45 para *G. applanatum* y 1.4 para *C. rutilus*.

Los mejores alineamientos de la secuencia putativa fueron con las dos proteínas FIP reportadas en esta base de datos para el género *Ganoderma*, correspondientes a *G. lucidum* y *G.*

*microsporium* (accesos PDB: 3F3HA y 3KCWA) con un porcentaje de identidad del 75% aproximadamente (ver Anexo IV, G-H). También se presentó una coincidencia con la proteína de *F. velutipes* reportada en el PDB, con una identidad del 62%. La estructura secundaria de las FIPs de *G. australe*, *G. applanatum* y *C. rutilus* fue muy similar en cuanto a su predicción, pero no alcanza a ser precisa en la totalidad de la estructura, difiriendo de las proteínas cristalizadas que están reportadas para el PDB (ej. *G. lucidum*). Podríamos decir que el predictor de I-Tasser hace un gran esfuerzo en su algoritmo para alcanzar este tipo de resultados, logrando grandes similitudes de secuencias establecidas con métodos químicos de alta precisión, pero que aún carece de una total exactitud a la hora de lograr las mismas estructuras que existen a nivel biológico. Por ende, se necesitan de modelos de aprendizaje de máquina (Machine learning) más robustos que ayuden a llenar estos vacíos entre la computación y la precisión biológica.

El modelo predictivo que arrojó I-Tasser para la estructura terciaria de *G. australe* se ve en la Figura 10, el cual se estructura en base a comparaciones con las proteínas reportadas en el PDB.



Figura 10. Modelo de estructura terciaria para la secuencia proteica putativa de la FIP de *G. australe*, con base en la predicción obtenida con la herramienta I-Tasser.



Los modelos predichos para la FIP putativa de *G. australe*, y para las secuencias de *G. applanatum* y *C. rutilus* establecieron que la proteínas poseen una estructura muy similar (>75%) a la estructura molecular establecida en *G.tsugae* para la FIP-gts por parte de Lin *et al.* (1997). El modelo de la FIP putativa de *G. australe* debe tener una actividad en el extremo N-terminal donde se considera que es el punto de interacción de la proteína para la promoción de la interacción celular y el desarrollo del proceso inmunomodulador (Cong *et al.* 2014, Uribe-Echeverry & López-Gartner 2017).

## 5. Conclusiones generales

### 5.1. Contribuciones de la Tesis

En este trabajo se logró la secuenciación y análisis preliminar del genoma de *Ganoderma australe*, a partir de una muestra de material biológico de origen colombiano. Esta especie de hongo macromiceto pertenece a un género de gran importancia industrial y medicinal, siendo éste el primer reporte realizado para esta especie. Además, se aporta una base de datos genómica del hongo que contribuirá a determinar componentes celulares importantes dentro del genoma del hongo, la caracterización de moléculas biológicas que ayuden en el área medicinal y su determinación para futuras aplicaciones. Esto permitirá a la comunidad científica nacional, a través del grupo de investigación en Ciencias biomédicas de la Universidad de Manizales, avanzar en el estudio de especies de hongos promisorios para el desarrollo de nuevas alternativas terapéuticas.

Se obtuvo 10.327 secuencias putativas de la caracterización del genoma de *G. australe*, en la cual se hizo un énfasis para la reconstrucción de la ruta metabólica del ácido ganodérico. Se aislaron varias secuencias putativas que pertenecen a esta ruta metabólica, presentando un reporte más completo sobre las enzimas que actúan en esta ruta metabólica dentro del genoma del hongo.

Se consiguió identificar la secuencia putativa de una proteína FIP dentro del genoma de *G. australe*. Este sería el primer reporte de este tipo de proteínas para la especie estudiada, implicando un avance en el estudio asociado a este tipo de proteínas y su aporte en el desarrollo de estrategias que ayuden en el tratamiento de enfermedades de alto impacto en el mundo.

### 5.2. Impactos Potenciales de la Tesis

Los productos biológicos de *Ganoderma* han tenido un crecimiento exponencial desde 1997 cuando poseían un valor de 1.7 millones de USD y que hoy se traducen en un mercado creciente y próspero. Es por ello que estudiando a fondo las propiedades inmunomoduladoras de las proteínas putativas encontradas en el genoma de *G. australe*, su expresión y características moleculares se puedan establecer estrategias de expresión y aislamiento para convertirse en un producto biotecnológico. Esto podría ser llevado a cabo en un término aproximado de 10 años. Al poder desarrollarse estas estrategias, conllevaría al avance en tratamiento de enfermedades de alto impacto tales como el cáncer, diabetes entre otras. Es importante remarcar la procedencia de este producto ya que sería un compuesto natural obtenido de un organismo vivo y que no poseería, hasta el momento, efectos contraproducentes para la salud humana.

### 5.3. Recomendaciones y trabajos futuros

Este primer reporte del ensamblaje de genoma para *Ganoderma australe* plantea la oportunidad de trabajos futuros para estudiar con más detalle el potencial genómico de la especie, ya que se dejan planteados baches de información que deben ser analizados en profundidad, tal como el contenido de las regiones repetitivas en el genoma, y su implicación en el tamaño genómico encontrado. Se deben efectuar otras estrategias de ensamblaje del genoma con el fin de validar su tamaño génico, y evaluar las regiones de transposones en estos ensamblajes para determinar un porcentaje común para la especie.

Se recomienda efectuar la anotación del genoma de *G. australe* junto con un transcriptoma que pertenezca a la misma especie, con el fin de evaluar la expresión de los genes evaluados de rutas metabólicas y proteínas de interés.

Con base en los resultados de este estudio con respecto a la ruta metabólica de los ácidos ganodéricos en *G. australe*, se podrán realizar estudios a futuro para caracterizar en más detalle las moléculas involucradas y utilizar las secuencias identificadas en modelos de expresión (ej. *P. pastoris*) para la obtención de estos metabolitos.

Para el genoma de *G. australe*, es vital continuar con la identificación de enzimas del complejo citocromo P450, ya que participan en la producción de metabolitos primarios y secundarios, especialmente la transformación de los productos de la ruta del ácido ganodérico con una reconocida importancia biológica, y así, establecer las estrategias de clonación de estas enzimas con el fin de lograr una producción industrial que conlleve a la comercialización de algún ácido ganodérico proveniente de *G. australe*.

Se debe avanzar en la expresión de la secuencia de la proteína FIP encontrada en el genoma del hongo, corroborar su estructura con trabajos de RNA-seq de la misma especie, y evaluar la expresión en modelos biológicos (ej. *S. cerevisiae* o *P. pastoris*) que ayuden a determinar su acción inmunomoduladora en líneas celulares.

## 6. Referencias bibliográficas

Alves, J. M. P., De Oliveira, A. L., Sandberg, T. O. M., Moreno-Gallego, J. L., De Toledo, M. A. F., De Moura, E. M. M., ... Gruber, A. (2016). GenSeed-HMM: A tool for progressive assembly using profile HMMS as seeds and its application in Alpavirinae viral discovery from metagenomic data. *Frontiers in Microbiology*, 7(MAR), 1–15. <https://doi.org/10.3389/fmicb.2016.00269>

Arboleda-Echavarría, C., & Gallón, A. I. M. (2010). Inducción de la actividad de lacasa en *Ganoderma* sp. y actividad antioxidante de su biomasa. *Revista Cubana de Farmacia*, 44(4), 519–532.

Baby, S., Johnson, A. J., & Govindan, B. (2015). Secondary metabolites from *Ganoderma*. *Phytochemistry*, 114, 66–101. <https://doi.org/10.1016/j.phytochem.2015.03.010>

Bastiaan-Net, S., Chanput, W., Hertz, A., Zwittink, R. D., Mes, J. J., & Wichers, H. J. (2013). Biochemical and functional characterization of recombinant fungal immunomodulatory proteins (rFIPs). *International Immunopharmacology*, 15(1), 167–175. <https://doi.org/10.1016/j.intimp.2012.11.003>

Blackwell, M. (2011). The fungi: 1, 2, 3 ... 5.1 million species? *American Journal of Botany*, 98(3), 426–438. <https://doi.org/10.3732/ajb.1000298>

Bishop, K. S., Kao, C. H. J., Xu, Y., Glucina, M. P., Paterson, R. R. M., & Ferguson, L. R. (2015). From 2000 years of *Ganoderma lucidum* to recent developments in nutraceuticals. *Phytochemistry*, 114, 56–65. <https://doi.org/10.1016/j.phytochem.2015.02.015>.

Cantarel, B. L., Korf, I., Robb, S. M. C., Parra, G., Ross, E., Moore, B., ... Yandell, M. (2008). MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, 18(1), 188–196. <https://doi.org/10.1101/gr.6743907>.

Cañas, A. I., & Camarero, S. (2010). Laccases and their natural mediators: Biotechnological tools for sustainable eco-friendly processes. *Biotechnology Advances*, 28(6), 694–705. <https://doi.org/10.1016/j.biotechadv.2010.05.002>.

Chang, H. H., Yeh, C. H., & Sheu, F. (2009). A novel immunomodulatory protein from *Poria cocos* induces Toll-like receptor 4-dependent activation within mouse peritoneal macrophages. *Journal of Agricultural and Food Chemistry*, 57(14), 6129–6139. <https://doi.org/10.1021/jf9011399>.

Chang, Y. H., Yang, J. S., Yang, J. L., Wu, C. L., Chang, S. J., Lu, K. W., ... Chung, J. G. (2009). *Ganoderma lucidum* extract promotes immune responses in normal BALB/c mice in vivo. *In Vivo*, 23(5), 755–759. <https://doi.org/23/5/755> [pii].

Chen, S., Xu, J., Liu, C., Zhu, Y., Nelson, D. R., Zhou, S., ... Sun, C. (2012). Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. *Nature Communications*, 3(May), 913–919. <https://doi.org/10.1038/ncomms1923>.

Chivian, D., Robertson, T., Bonneau, R., & Baker, D. (2005). Ab Initio Methods. In *Structural Bioinformatics* (Vol. 44, pp. 547–557). <https://doi.org/10.1002/0471721204.ch27>.

Cooper, E. L. (2004). Drug Discovery, CAM and Natural Products. *Evidence-Based Complementary and Alternative Medicine : eCAM*, 1(3), 215–217. <https://doi.org/10.1093/ecam/neh032>

Cuéllar, I. C., & Rodríguez, E. C. (2003). Reacciones adversas a medicamentos. *Univ. med*, 44(4), 212-216.

D'Souza, T. M., Boominathan, K., & Reddy, C. A. (1996). Isolation of laccase gene-specific sequences from white rot and brown rot fungi by PCR. *Applied and Environmental Microbiology*, 62(10), 3739–3744. Retrieved from <http://aem.asm.org/content/62/10/3739.short>.

Delcher, A. L., Harmon, D., Kasif, S., White, O., & Salzberg, S. L. (1999). Improved microbial gene identification with GLIMMER. *Nucleic acids research*, 27(23), 4636-4641. <https://doi.org/10.1093/nar/27.23.4636>

Delcher, A. L., Bratke, K. A., Powers, E. C., & Salzberg, S. L. (2007). Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics*, 23(6), 673–679. <https://doi.org/10.1093/bioinformatics/btm009>

De Silva, D. D., Rapior, S., Hyde, K. D., & Bahkali, A. H. (2012). Medicinal mushrooms in prevention and control of diabetes mellitus. *Fungal Diversity*, 56(1), 1–29. <https://doi.org/10.1007/s13225-012-0187-4>.

Dhar, M. K., Koul, A., & Kaul, S. (2013). Farnesyl pyrophosphate synthase: A key enzyme in isoprenoid biosynthetic pathway and potential molecular target for drug development. *New Biotechnology*, 30(2), 114–123. <https://doi.org/10.1016/j.nbt.2012.07.001>.

Dighton, J. (2016). *Fungi in ecosystem processes* (Vol. 31). CRC Press.

Dighton, J., & White, J. F. (Eds.). (2017). *The fungal community: its organization and role in the ecosystem*. CRC Press.

Ding, Y. X., Ou-Yang, X., Shang, C. H., Ren, A., Shi, L., Li, Y. X., & Zhao, M. W. (2008). Molecular cloning, characterization, and differential expression of a farnesyl-diphosphate synthase gene from the basidiomycetous fungus *Ganoderma lucidum*. *Bioscience, biotechnology, and biochemistry*, 72(6), 1571-1579. <https://doi.org/10.1271/bbb.80067>

Ding, Y., Seow, S. V., Huang, C. H., Liew, L. M., Lim, Y. C., Kuo, I. C., & Chua, K. Y. (2009). Coadministration of the fungal immunomodulatory protein FIP-Fve and a tumour-associated

antigen enhanced antitumour immunity. *Immunology*, 128(1 PART 2), 881–894. <https://doi.org/10.1111/j.1365-2567.2009.03099.x>.

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Computational Biology*, 7(10). <https://doi.org/10.1371/journal.pcbi.1002195>

Eleftherios, E., Vassilis, M. G., & Cleanthes, I. (2014). The potential use of mushrooms  $\beta$ -glucans in the food industry. *International Journal of Biotechnology for Wellness Industries*, 3(1), 15-18. <http://www.lifescienceglobal.com/pms/index.php/ijbwi/article/view/1823>

Elsayed, E. A., El Enshasy, H., Wadaan, M. A. M., & Aziz, R. (2014). Mushrooms: A potential natural source of anti-inflammatory compounds for medical applications. *Mediators of Inflammation*, 2014. <https://doi.org/10.1155/2014/805841>.

Erjavec, J., Kos, J., Ravnikar, M., Dreo, T., & Sabotič, J. (2012). Proteins of higher fungi - from forest to application. *Trends in Biotechnology*, 30(5), 259–273. <https://doi.org/10.1016/j.tibtech.2012.01.004>.

Escobar, M. C., Petrásovits, A., Peruga, A., Silva, N., Vives, M., & Robles, S. (2000). Mitos sobre la prevención y el control de las enfermedades no transmisibles en América Latina. *Salud pública de méxico*, 42(1), 56-64.

Finn, R. D., Clements, J., & Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic acids research*, 39 (supply 2), W29-W37. <https://doi.org/10.1093/nar/gkr367>

Gao, J. J., Min, B. S., Ahn, E. M., Nakamura, N., Lee, H. K., & Hattori, M. (2002). New triterpene aldehydes, lucialdehydes AC, from *Ganoderma lucidum* and their cytotoxicity against murine and human tumor cells. *Chemical and Pharmaceutical Bulletin*, 50(6), 837-840. <https://doi.org/10.1248/cpb.50.837>

Godio, R. P., Fouces, R., & Martín, J. F. (2007). A Squalene Epoxidase Is Involved in Biosynthesis of Both the Antitumor Compound Clavarinic Acid and Sterols in the Basidiomycete *H. sublateralium*. *Chemistry and Biology*, 14(12), 1334–1346. <https://doi.org/10.1016/j.chembiol.2007.10.018>.

González-M.A., Botero-Orozco, K. J. & López-Gartner, G. A. (2014). Finding of a novel fungal immunomodulatory protein coding sequence in *Ganoderma australe*. *Revista Colombiana de Biotecnología*, Universidad Nacional de Colombia (ISSN 0123-3475), 16(2), 90-95.

Grigoriev, I. V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otilar, R., ... & Smirnova, T. (2013). MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Research*, 42(D1), D699-D704. <https://doi.org/10.1093/nar/gkt1183>

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUASt: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), 1072-1075. <https://doi.org/10.1093/bioinformatics/btt086>

Güzeldağ, G., & Colak, O. (2007). Molecular identification of *Ganoderma lucidum* from Turkey. *International Journal of Agriculture and Biology*, 9(5), 767-770.

Hardison, R. C. (2003). Comparative genomics. *PLoS biology*, 1(2), e58. <https://doi.org/10.1371/journal.pbio.0000058>

Hibbett, D. S., Binder, M., Bischoff, J. F., Blackwell, M., Cannon, P. F., Eriksson, O. E., ... Zhang, N. (2007). A higher-level phylogenetic classification of the Fungi. *Mycological Research*, 111(5), 509–547. <https://doi.org/10.1016/j.mycres.2007.03.004>.

Holt, C., & Yandell, M. (2011). MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics*, 12(1), 491. <https://doi.org/10.1186/1471-2105-12-491>.

Hsiao, Y. M., Huang, Y. L., Tang, S. C., Shieh, G. J., Lai, J. Y., Wang, P. H., ... Ko, J. L. (2008). Effect of a fungal immunomodulatory protein from *Ganoderma tsugae* on cell cycle and interferon-gamma production through phosphatidylinositol 3-kinase signal pathway. *Process Biochemistry*, 43(4), 423–430. <https://doi.org/10.1016/j.procbio.2008.01.005>.

Hsin, I. L., Ou, C. C., Wu, T. C., Jan, M. S., Wu, M. F., Chiu, L. Y., ... Ko, J. L. (2011). GMI, an immunomodulatory protein from *Ganoderma microsporum*, induces autophagy in non-small cell lung cancer cells. *Autophagy*, 7(8), 873–882. <https://doi.org/10.4161/auto.7.8.15698>.

Hsu, H. C., Hsu, C. I., Lin, R., Kao, C., & Lin, J. (1997). Fip-vvo, a new fungal immunomodulatory protein isolated from *Volvariella volvacea*. *Biochem. J*, 323, 557-565.

Illana-Esteban, C. (2008). El hongo maitake (*Grifola frondosa*) y su potencial terapéutico. *Revista Iberoamericana de Micología*, 25(3), 141-144.

Jargalmaa, S., Eimes, J. A., Park, M. S., Park, J. Y., Oh, S.-Y., & Lim, Y. W. (2017). Taxonomic evaluation of selected *Ganoderma* species and database sequence validation. *PeerJ*, 5, e3596. <https://doi.org/10.7717/peerj.3596>

Jeurink, P. V., Noguera, C. L., Savelkoul, H. F. J., & Wichers, H. J. (2008). Immunomodulatory capacity of fungal proteins on the cytokine production of human peripheral blood mononuclear cells. *International Immunopharmacology*, 8(8), 1124–1133. <https://doi.org/10.1016/j.intimp.2008.04.004>.

Jeong, Y. T., Yang, B. K., Jeong, S. C., Kim, S. M., & Song, C. H. (2008). *Ganoderma applanatum*: a promising mushroom for antitumor and immunomodulating activity. *Phytotherapy research*, 22(5), 614-619. <https://doi.org/10.1002/ptr.2294>

JINN, T.-R., WU, C.-M., TU, W.-C., KO, J.-L., & TZEN, J. T. C. (2006). Functional Expression of FIP- *gts* , a Fungal Immunomodulatory Protein from *Ganoderma Tsugae* in Sf21 Insect Cells. *Bioscience, Biotechnology, and Biochemistry*, 70(11), 2627–2634. <https://doi.org/10.1271/bbb.60232>.

Kawagishi, H., Nomura, A., Mizuno, T., Kimura, A., & Chiba, S. (1990). Isolation and characterization of a lectin from *Grifola frondosa* fruiting bodies. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1034(3), 247-252. [https://doi.org/10.1016/0304-4165\(90\)90045-X](https://doi.org/10.1016/0304-4165(90)90045-X)

Kim, H. W. & Kim, B. K. (1999). Biomedicinal triterpenoids of *Ganoderma lucidum* (Curt.Fr.) P. Karst. (Aphyllphoromycetidae). *International Journal of Medicinal Mushrooms*, 1(2), 121–138.

Kino, K., Yamashita, A., Yamaoka, K., Watanabe, J., Tanaka, S., Ko, K., ... & Tsunoo, H. (1989). Isolation and characterization of a new immunomodulatory protein, ling zhi-8 (LZ-8), from *Ganoderma lucidum*. *Journal of Biological Chemistry*, 264(1), 472-478.

Kladar, N. V., Gavarić, N. S., & Božin, B. N. (2016). *Ganoderma*: insights into anticancer effects. *European Journal of Cancer Prevention*, 25(5), 462-471. <https://doi.org/10.1097/CEJ.0000000000000204>

Ko, J. L., Hsu, C. I., Lin, R. H., Kao, C. L., & Lin, J. Y. (1995). A New Fungal Immunomodulatory Protein, FIP-fve Isolated from the Edible Mushroom, *Flammulina velutipes* and its Complete Amino Acid Sequence. *European Journal of Biochemistry*, 228(2), 244-249.

König, G. M., Kehraus, S., Seibert, S. F., Abdel-Lateff, A., & Müller, D. (2006). Natural products from marine organisms and their associated microbes. *ChemBioChem*, 7(2), 229-238. <https://doi.org/10.1002/cbic.200500087>

Korf, I. (2013). SNAP: Semi-HMM-based Nucleic Acid Parser. Ian Korf homepage: <http://homepage.mac.com/iankorf>. <http://korflab.ucdavis.edu/software.html>

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33(7), 1870–1874. <https://doi.org/10.1093/molbev/msw054>

Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., & Salzberg, S. L. (2004). Versatile and open software for comparing large genomes. *Genome Biology*, 5(2), R12. <https://doi.org/10.1186/gb-2004-5-2-r12>

Kües, U., Nelson, D. R., Liu, C., Yu, G. J., Zhang, J., Li, J., ... & Sun, H. (2015). Genome analysis of medicinal *Ganoderma* spp. with plant-pathogenic and saprotrophic life-styles. *Phytochemistry*, 114, 18-37. <https://doi.org/10.1016/j.phytochem.2014.11.019>

Lees, N. D., Skaggs, B., Kirsch, D. R., & Bard, M. (1995). Cloning of the late genes in the ergosterol



biosynthetic pathway of *Saccharomyces cerevisiae*—A review. *Lipids*, 30(3), 221-226.  
<https://doi.org/10.1007/BF02537824>

Li, Y. Y., Mi, Z. Y., Tang, Y., Wang, G., Li, D. S., & Tang, Y. J. (2009). Lanostanoids isolated from *Ganoderma lucidum* mycelium cultured by submerged fermentation. *Helvetica Chimica Acta*, 92(8), 1586-1593. DOI: 10.1002/hlca.200900028

Li, M., Zhang, G., Wang, H., & Ng, T. (2010). Purification and characterization of a laccase from the edible wild mushroom *Tricholoma mongolicum*. *Journal of Microbiology and Biotechnology*, 20(7), 1069–1076. <https://doi.org/10.4014/jmb.0912.12033>.

Li, Q. Z., Wang, X. F., & Zhou, X. W. (2011). Recent status and prospects of the fungal immunomodulatory protein family. *Critical reviews in biotechnology*, 31(4), 365-375. <https://doi.org/10.3109/07388551.2010.543967>

Li, F., Wen, H. A., Liu, X., Zhou, F., & Chen, G. (2012). Gene cloning and recombinant expression of a novel fungal immunomodulatory protein from *Trametes versicolor*. *Protein expression and purification*, 82(2), 339-344. <https://doi.org/10.1016/j.pep.2012.01.015>

Li, Y. B., Liu, R. M., & Zhong, J. J. (2013). A new ganoderic acid from *Ganoderma lucidum* mycelia and its stability. *Fitoterapia*, 84, 115-122. <https://doi.org/10.1016/j.fitote.2012.11.008>

Li, S., Nie, Y., Ding, Y., Shi, L., & Tang, X. (2014). Recombinant Expression of a Novel Fungal Immunomodulatory Protein with Human Tumor Cell Antiproliferative Activity from *Nectria haematococca*. *International journal of molecular sciences*, 15(10), 17751-17764. doi:10.3390/ijms151017751

Li, S. Y., Shi, L. J., Ding, Y., Nie, Y., & Tang, X. M. (2015). Identification and functional characterization of a novel fungal immunomodulatory protein from *Postia placenta*. *Food and Chemical Toxicology*, 78, 64-70. <https://doi.org/10.1016/j.fct.2015.01.013>

Li, S., Jiang, Z., Xu, W., Xie, Y., Zhao, L., Tang, X., ... & Xin, F. (2017). FIP-sch2, a new fungal immunomodulatory protein from *Stachybotrys chlorohalonata*, suppresses proliferation and migration in lung cancer cells. *Applied microbiology and biotechnology*, 101(8), 3227-3235. <https://doi.org/10.1007/s00253-016-8030-6>

Liao, C. H., Hsiao, Y. M., Hsu, C. P., Lin, M. Y., Wang, J. C. H., Huang, Y. L., & Ko, J. L. (2006). Transcriptionally mediated inhibition of telomerase of fungal immunomodulatory protein from *Ganoderma tsugae* in A549 human lung adenocarcinoma cell line. *Molecular carcinogenesis*, 45(4), 220-229. DOI: 10.1002/mc.20161

Liao, C. H., Hsiao, Y. M., Lin, C. H., Yeh, C. S., Wang, J. C. H., Ni, C. H., ... & Ko, J. L. (2008). Induction of premature senescence in human lung cancer by fungal immunomodulatory

protein from *Ganoderma tsugae*. *Food and chemical toxicology*, 46(5), 1851-1859. <https://doi.org/10.1016/j.fct.2008.01.044>

LIN, J. Y., & CHOU, T. B. (1984). Isolation and characterization of a lectin from edible mushroom, *Volvariella volvacea*. *The Journal of Biochemistry*, 96(1), 35-40. <https://doi.org/10.1093/oxfordjournals.jbchem.a134826>

Lin, W. H., Hung, C. H., Hsu, C. I., & Lin, J. Y. (1997). Dimerization of the N-terminal amphipathic  $\alpha$ -helix domain of the fungal immunomodulatory protein from *Ganoderma tsugae* (Fip-gts) defined by a yeast two-hybrid system and site-directed mutagenesis. *Journal of Biological Chemistry*, 272(32), 20044–20048. <https://doi.org/10.1074/jbc.272.32.20044>.

Lin, Z. B., & Zhang, H. N. (2004). Anti-tumor and immunoregulatory activities of *Ganoderma lucidum* and its possible mechanisms. *Acta Pharmacologica Sinica*, 25, 1387-1395.

Lin, J. Y., Chen, M. L., Chiang, B. L. & Lin, B. F. (2006). *Ganoderma tsugae* supplementation alleviates bronchoalveolar inflammation in an airway sensitization and challenge mouse model. *International Immunopharmacology*, 6, 241–251. <https://doi.org/10.1016/j.intimp.2005.08.009>

Lin, J. W., Hao, L. X., Xu, G. X., Sun, F., Gao, F., Zhang, R., & Liu, L. X. (2009). Molecular cloning and recombinant expression of a gene encoding a fungal immunomodulatory protein from *Ganoderma lucidum* in *Pichia pastoris*. *World journal of microbiology and biotechnology*, 25(3), 383-390. <https://doi.org/10.1007/s11274-008-9902-4>

Lin, J. wei, Guan, S. yue, Duan, Z. wen, Shen, Y. hua, Fan, W. li, Chen, L. jing, ... Li, T. lai. (2016). Gene cloning of a novel fungal immunomodulatory protein from *Chroogomphus rutilus* and its expression in *Pichia pastoris*. *Journal of Chemical Technology and Biotechnology*, 91(11), 2761–2768. <https://doi.org/10.1002/jctb.4881>

Liu, J., Shimizu, K., Konishi, F., Noda, K., Kumamoto, S., Kurashiki, K., & Kondo, R. (2007). Anti-androgenic activities of the triterpenoids fraction of *Ganoderma lucidum*. *Food Chemistry*, 100(4), 1691-1696. <https://doi.org/10.1016/j.foodchem.2006.01.003>

Liu, D., Gong, J., Dai, W., Kang, X., Huang, Z., Zhang, H. M., ... & Xiong, X. (2012). The genome of *Ganoderma lucidum* provide insights into triterpene biosynthesis and wood degradation. *PloS one*, 7(5). <https://doi.org/10.1371/journal.pone.0036146>

López-Gartner, G., Agudelo-Valencia, D., Castaño, S., Isaza, G. A., Castillo, L. F., Sánchez, M., & Arango, J. (2015). Identification of a Putative Ganoderic Acid Pathway Enzyme in a *Ganoderma australe* Transcriptome by Means of a Hidden Markov Model. In *9th International Conference on Practical Applications of Computational Biology and Bioinformatics* (pp. 107-115). Springer International Publishing. [https://doi.org/10.1007/978-3-319-19776-0\\_12](https://doi.org/10.1007/978-3-319-19776-0_12)

Lull, C., Wichers, H. J., & Savelkoul, H. F. J. (2005). Antiinflammatory and immunomodulating properties of fungal metabolites. *Mediators of Inflammation*, 2005(2), 63–80. <https://doi.org/10.1155/MI.2005.63>.

Manoharachary, C., Sridhar, K., Singh, R., Adholeya, A., Suryanarayanan, T. S., Rawat, S., & Johri, B. N. (2005). Fungal biodiversity: Distribution, conservation and prospecting of fungi from India. *Current Science*, 89(1), 58–71.

Matos, A. J. F. S., Bezerra, R. M. F., & Dias, A. A. (2007). Screening of fungal isolates and properties of *Ganoderma applanatum* intended for olive mill wastewater decolourization and dephenolization. *Letters in applied microbiology*, 45(3), 270-275. DOI: 10.1111/j.1472-765X.2007.02181.x

McMeekin, D. (2004). The perception of *Ganoderma lucidum* in Chinese and Western culture. *Mycologist*, 18(4), 165-169. <https://doi.org/10.1017/S0269915X04004069>

Mendonça, R. T., Jara, J. F., González, V., Elissetche, J. P., & Freer, J. (2008). Evaluation of the white-rot fungi *Ganoderma australe* and *Ceriporiopsis subvermispora* in biotechnological applications. *Journal of industrial microbiology & biotechnology*, 35(11), 1323-1330. <https://doi.org/10.1007/s10295-008-0414-x>

Miller, W., Makova, K. D., Nekrutenko, A., & Hardison, R. C. (2004). Comparative Genomics. *Annual Review of Genomics and Human Genetics*, 5(1), 15–56. <https://doi.org/10.1146/annurev.genom.5.061903.180057>.

Min, B. S., Gao, J. J., Nakamura, N., & Hattori, M. (2000). Triterpenes from the spores of *Ganoderma lucidum* and their cytotoxicity against meth-A and LLC tumor cells. *Chemical & pharmaceutical bulletin*, 48(7), 1026-1033. <https://doi.org/10.1248/cpb.48.1026>

Mohanta, T. K., & Bae, H. (2015). The diversity of fungal genome. *Biological procedures online*, 17(1), 8. <https://doi.org/10.1186/s12575-015-0020-z>

Moradali, M. F., Mostafavi, H., Ghods, S., & Hedjaroude, G. A. (2007). Immunomodulating and anticancer agents in the realm of macromycetes fungi (macrofungi). *International immunopharmacology*, 7(6), 701-724. <https://doi.org/10.1016/j.intimp.2007.01.008>

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., & Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic acids research*, 35(suppl 2), W182-W185. <https://doi.org/10.1093/nar/gkm321>

Mota, T. R., Kato, C. G., Peralta, R. A., Bracht, A., de Morais, G. R., Baesso, M. L., ... & Peralta, R. M. (2015). Decolourization of Congo Red by *Ganoderma lucidum* Laccase: Evaluation of Degradation Products and Toxicity. *Water, Air, & Soil Pollution*, 226(10), 351. <https://doi.org/10.1007/s11270-015-2612-2>

- Mueller, G. M. & Schmit, J. P. (2007). Fungal biodiversity: what do we know? What can we predict? *Biodiversity and Conservation*, 16(1), 1-5. <https://doi.org/10.1007/s10531-006-9117-7>
- Mukherjee, S., Stamatis, D., Bertsch, J., Ovchinnikova, G., Verezemskaya, O., Isbandi, M., ... & Reddy, T. B. K. (2016). Genomes OnLine Database (GOLD) v. 6: data updates and feature enhancements. *Nucleic acids research*, gkw992. <https://gold.jgi.doe.gov/>
- Murasugi, A., Tanaka, S., Komiyama, N., Iwata, N., Kino, K., Tsunoo, H., & Sakuma, S. (1991). Molecular cloning of a cDNA and a gene encoding an immunomodulatory protein, Ling Zhi-8, from a fungus, *Ganoderma lucidum*. *Journal of Biological Chemistry*, 266(4), 2486–2493.
- Ng, T. B. (2004). Peptides and proteins from fungi. *Peptides*, 25(6), 1055-1073. <https://doi.org/10.1016/j.peptides.2004.03.013>
- Ng, T. B., Cheung, R. C. F., Wong, J. H., Chan, Y. S., Dan, X., Pan, W., ... & Liu, F. (2016). Fungal proteinaceous compounds with multiple biological activities. *Applied microbiology and biotechnology*, 100(15), 6601-6617. <https://doi.org/10.1007/s00253-016-7671-9>
- Ngai, P. H., & Ng, T. B. (2004). A mushroom (*Ganoderma capense*) lectin with spectacular thermostability, potent mitogenic activity on splenocytes, and antiproliferative activity toward tumor cells. *Biochemical and biophysical research communications*, 314(4), 988-993.
- Nova, E., Montero, S., Gómez, S., & Marcos, A. (2004). La estrecha relación entre la nutrición y el sistema inmunitario. Soporte Nutricional en el Paciente Oncológico. Gómez Candela C, Sastre Gallego A (eds). Barcelona: Glosa, 9-21.
- OMS. 2014. Informe sobre la situación mundial de las enfermedades no transmisibles 2014. Consultado abril 21 de 2016 disponible en URL: <http://www.who.int/nmh/publications/ncd-status-report-2014/es/>
- Özel, Y. K., Gedikli, S., Aytar, P., Ünal, A., Yamaç, M., Çabuk, A., & Kolankaya, N. (2010). New fungal biomasses for cyanide biodegradation. *Journal of bioscience and bioengineering*, 110(4), 431-435. <https://doi.org/10.1016/j.jbiosc.2010.04.011>
- Park, Y. J., Kwon, O. C., Son, E. S., Yoon, D. E., Han, W., Yoo, Y. B., & Lee, C. S. (2012). Taxonomy of *Ganoderma lucidum* from Korea based on rDNA and partial  $\alpha$ -Tubulin gene sequence analysis. *Mycobiology*, 40(1), 71–75. <https://doi.org/10.5941/MYCO.2012.40.1.071>
- Patel, S., & Goyal, A. (2012). Recent developments in mushrooms as anti-cancer therapeutics: a review. *3 Biotech*, 2(1), 1-15. <https://doi.org/10.1007/s13205-011-0036-2>
- Patlolla, J., & Rao, C. (2012). Triterpenoids for cancer prevention and treatment: current status and future prospects. *Current pharmaceutical biotechnology*, 13(1), 147-155. <https://doi.org/10.2174/138920112798868719>

Pillai, T. G., John, M., & Thomas, G. S. (2011). Prevention of cisplatin induced nephrotoxicity by terpenes isolated from *Ganoderma lucidum* occurring in Southern Parts of India. *Experimental and toxicologic pathology*, 63(1), 157-160. <https://doi.org/10.1016/j.etp.2009.11.003>

Pohleven, J., Obermajer, N., Sabotič, J., Anžlovar, S., Sepčić, K., Kos, J., ... & Brzin, J. (2009). Purification, characterization and cloning of a ricin B-like lectin from mushroom *Clitocybe nebularis* with antiproliferative activity against human leukemic T cells. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1790(3), 173-181. <https://doi.org/10.1016/j.bbagen.2008.11.006>

Pushparajah, V., Fatima, A., Chong, C. H., Gambule, T. Z., Chan, C. J., Ng, S. T., ... & Lim, R. L. H. (2016). Characterisation of a new fungal immunomodulatory protein from Tiger Milk mushroom, *Lignosus rhinocerotis*. *Scientific reports*, 6, 30010. doi:10.1038/srep30010

Qian, J., Xu, H., Song, J., Xu, J., Zhu, Y., & Chen, S. (2013). Genome-wide analysis of simple sequence repeats in the model medicinal mushroom *Ganoderma lucidum*. *Gene*, 512(2), 331-336. <https://doi.org/10.1016/j.gene.2012.09.127>

Ramírez, N. E., Vargas, M. C., Ariza, J. C., & Martínez, C. (2003). Caracterización de la lacasa obtenida por dos métodos de producción con *Pleurotus ostreatus*. *Revista Colombiana de Biotecnología*, 5(2), 64-72.

Ren, A., Qin, L., Shi, L., Dong, X., Mu, D. S., Li, Y. X., & Zhao, M. W. (2010). Methyl jasmonate induces ganoderic acid biosynthesis in the basidiomycetous fungus *Ganoderma lucidum*. *Bioresource technology*, 101(17), 6785-6790. <https://doi.org/10.1016/j.biortech.2010.03.118>

Richter, C., Wittstein, K., Kirk, P. M., & Stadler, M. (2015). An assessment of the taxonomy and chemotaxonomy of *Ganoderma*. *Fungal Diversity*, 71(1). <https://doi.org/10.1007/s13225-014-0313-6>

Rigas, F., Papadopoulou, K., Dritsa, V., & Doulia, D. (2007). Bioremediation of a soil contaminated by lindane utilizing the fungus *Ganoderma australe* via response surface methodology. *Journal of hazardous Materials*, 140(1), 325-332. <https://doi.org/10.1016/j.jhazmat.2006.09.035>

Ríos-Cañavate, J. L. (2008). *Ganoderma lucidum*, un hongo con propiedades inmunoestimulantes. *Revista de Fitoterapia*, 8(2), 135-146.

Roberts, R. J., Carneiro, M. O., & Schatz, M. C. (2013). The advantages of SMRT sequencing. *Genome Biology*, 14(6), 405. <https://doi.org/10.1186/gb-2013-14-6-405>

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature protocols*, 5(4), 725-738. doi:10.1038/nprot.2010.5

Roy, A., Yang, J., & Zhang, Y. (2012). COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. *Nucleic acids research*, 40(W1), W471-W477. <https://doi.org/10.1093/nar/gks372>

- Russell, R. & Patterson, M. (2006). *Ganoderma*: A therapeutic Fungal Factory. *Phytochemistry*, 67, 1985–2001. <https://doi.org/10.1016/j.phytochem.2006.07.004>
- Salamov, A. A., & Solovyev, V. V. (2000). *Ab initio* gene finding in Drosophila genomic DNA. *Genome research*, 10(4), 516-522. <https://doi.org/10.1101/gr.10.4.516>
- Saltarelli, R., Ceccaroli, P., Buffalini, M., Vallorani, L., Casadei, L., Zambonelli, A., ... & Stocchi, V. (2015). Biochemical characterization and antioxidant and antiproliferative activities of different *Ganoderma* collections. *Journal of molecular microbiology and biotechnology*, 25(1), 16-25. <https://doi.org/10.1159/000369212>
- Sanodiya, B. S., Thakur, G. S., Baghel, R. K., Prasad, G. B. K. S., & Bisen, P. S. (2009). *Ganoderma lucidum*: a potent pharmacological macrofungus. *Current pharmaceutical biotechnology*, 10(8), 717-742. <https://doi.org/10.2174/138920109789978757>
- Schepetkin, I. A. & Quinn, M. T. (2006). Botanical polysaccharides: Macrophage immunomodulation and therapeutic potential. *International Immunopharmacology*, 6, 317-333. <https://doi.org/10.1016/j.intimp.2005.10.005>
- Schmidt-Dannert, C. (2014). Biosynthesis of terpenoid natural products in fungi. In *Biotechnology of Isoprenoids* (pp. 19-61). Springer International Publishing. [https://doi.org/10.1007/10\\_2014\\_283](https://doi.org/10.1007/10_2014_283)
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19), 3210-3212. <https://doi.org/10.1093/bioinformatics/btv351>
- Shang, C. H., Shi, L., Ren, A., Qin, L., & Zhao, M. W. (2010). Molecular cloning, characterization, and differential expression of a lanosterol synthase gene from *Ganoderma lucidum*. *Bioscience, biotechnology, and biochemistry*, 74(5), 974-978. <https://doi.org/10.1271/bbb.90833>
- Sheena, N., Lakshmi, B. & Janardhanan, K. (2005). Therapeutic potencial of *Ganoderma lucidum* (Fr.) P. Karst. *Natural Product Radiance*, 4(5), 382-386. <http://nopr.niscair.res.in/handle/123456789/8124>
- Sheu, F., Chien, P. J., Chien, A. L., Chen, Y. F., & Chin, K. L. (2004). Isolation and characterization of an immunomodulatory protein (APP) from the Jew's Ear mushroom *Auricularia polytricha*. *Food chemistry*, 87(4), 593-600. <https://doi.org/10.1016/j.foodchem.2004.01.015>
- Sheu, F., Chien, P. J., Hsieh, K. Y., Chin, K. L., Huang, W. T., Tsao, C. Y., ... & Chang, H. H. (2009). Purification, cloning, and functional characterization of a novel immunomodulatory protein from *Antrodia camphorata* (bitter mushroom) that exhibits TLR2-dependent NF-κB activation and M1 polarization within murine macrophages. *Journal of agricultural and food chemistry*, 57(10), 4130-4141. DOI: 10.1021/jf900469a

Sheu, F., Chien, P. J., Wang, H. K., Chang, H. H., & Shyu, Y. T. (2007). New protein PCiP from edible golden oyster mushroom (*Pleurotus citrinopileatus*) activating murine macrophages and splenocytes. *Journal of the Science of Food and Agriculture*, 87(8), 1550-1558. DOI: 10.1002/jsfa.2887

Shiao, M. S. (2003). Natural products of the medicinal fungus *Ganoderma lucidum*: Occurrence, biological activities, and pharmacological functions. *Chemical Record*, 3(3), 172–180. <https://doi.org/10.1002/tcr.10058>

Slater, G. S. & Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6(31). doi: 10.1186/1471-2105-6-31

Smiderle, F. R., Olsen, L. M., Ruthes, A. C., Czelusniak, P. A., Santana-Filho, A. P., Sasaki, G. L., ... & Iacomini, M. (2012). Exopolysaccharides, proteins and lipids in *Pleurotus pulmonarius* submerged culture using different carbon sources. *Carbohydrate Polymers*, 87(1), 368-376. <https://doi.org/10.1016/j.carbpol.2011.07.063>

Solovyev, V., Kosarev, P., Seledsov, I., & Vorobyev, D. (2006). Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome biology*, 7(1), S10. <https://doi.org/10.1186/gb-2006-7-s1-s10>

Sono, M., Roach, M. P., Coulter, E. D., & Dawson, J. H. (1996). Heme-containing oxygenases. *Chemical Reviews*, 96(7), 2841-2888.

Stanke, M. & Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics*, 19(suppl\_2), ii215-ii225. <https://doi.org/10.1093/bioinformatics/btg1080>

Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., & Morgenstern, B. (2006). AUGUSTUS: *ab initio* prediction of alternative transcripts. *Nucleic acids research*, 34(suppl\_2), W435-W439. <https://doi.org/10.1093/nar/gkl200>

Strobel, G.A. (2003). Endophytes as sources of bioactive products. *Microbes and Infection* 5, 535–544. [https://doi.org/10.1016/S1286-4579\(03\)00073-X](https://doi.org/10.1016/S1286-4579(03)00073-X)

Struhl, K. (1987). Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. *Cell*, 49(3), 295–297. [https://doi.org/10.1016/0092-8674\(87\)90277-7](https://doi.org/10.1016/0092-8674(87)90277-7)

Suárez Arango, C. S., & Nieto, I. J. (2013). Cultivo biotecnológico de macrohongos comestibles: una alternativa en la obtención de nutraceuticos. *Revista Iberoamericana de Micología*, 30(1), 1-8.

Syed, K., Nelson, D. R., Riley, R., & Yadav, J. S. (2013). Genomewide annotation and comparative genomics of cytochrome P450 monooxygenases (P450s) in the polypore species

*Bjerkandera adusta*, *Ganoderma* sp. and *Phlebia brevispora*. *Mycologia*, 105(6), 1445-1455.  
<https://doi.org/10.3852/13-002>

Sze, S. C. W., Ho, J. C. K., & Liu, W. K. (2004). *Volvariella volvacea* lectin activates mouse T lymphocytes by a calcium dependent pathway. *Journal of cellular biochemistry*, 92(6), 1193-1202. DOI: 10.1002/jcb.20153

Tang, W., Liu, J. W., Zhao, W. M., Wei, D. Z., & Zhong, J. J. (2006). Ganoderic acid T from *Ganoderma lucidum* mycelia induces mitochondria mediated apoptosis in lung cancer cells. *Life sciences*, 80(3), 205-211. <https://doi.org/10.1016/j.lfs.2006.09.001>

Tang, Y. J., Zhang, W., Liu, R. S., Zhu, L. W., & Zhong, J. J. (2011). Scale-up study on the fed-batch fermentation of *Ganoderma lucidum* for the hyperproduction of ganoderic acid and *Ganoderma* polysaccharides. *Process Biochemistry*, 46(1), 404-408. <https://doi.org/10.1016/j.procbio.2010.08.013>

Uribe-Echeverry, P. T., & Lopez-Gartner, G. A. (2017). Fungal immunomodulatory proteins in the context of biomedicine. *Frontiers in Bioscience*, 9, 286–306. <https://doi.org/10.2741/e803>

Vincent, M., Philippe, E., Everard, A., Kassis, N., Rouch, C., Denom, J., ... & Magnan, C. (2013). Dietary supplementation with *Agaricus blazei* murill extract prevents diet-induced obesity and insulin resistance in rats. *Obesity*, 21(3), 553-561. DOI: 10.1002/oby.20276

Wachtel-Galor S., Yuen J., Buswell J.A., Benzie I.F.F. (2011). *Ganoderma lucidum* (Lingzhi or Reishi) - A Medicinal Mushroom. in: *Herbal Medicine: Biomolecular and Clinical Aspects*. Benzie I.F.F., Wachtel-Galor S. editors. 2nd edition, Boca Raton (FL), CRC Press, Chapter 9.

Wang, H. X., & Ng, T. B. (2006). Purification of a laccase from fruiting bodies of the mushroom *Pleurotus eryngii*. *Applied Microbiology and Biotechnology*, 69(5), 521-525. <https://doi.org/10.1007/s00253-005-0086-7>

Wang, H. X., Liu, W. K., Ng, T. B., Ooi, V. E. C., & Chang, S. T. (1996). The immunomodulatory and antitumor activities of lectins from the mushroom *Tricholoma mongolicum*. *Immunopharmacology*, 31(2), 205-211. [https://doi.org/10.1016/0162-3109\(95\)00049-6](https://doi.org/10.1016/0162-3109(95)00049-6)

Wang, X. F., Su, K. Q., Bao, T. W., Cong, W. R., Chen, Y. F., Li, Q. Z., & Zhou, X. W. (2012). Immunomodulatory effects of fungal proteins. *Current Topics in Nutraceuticals Research*, 10(1), 1.

Wang, C. H., Hsieh, S. C., Wang, H. J., Chen, M. L., Lin, B. F., Chiang, B. H., & Lu, T. J. (2014). Concentration variation and molecular characteristics of soluble (1, 3; 1, 6)- $\beta$ -d-glucans in submerged cultivation products of *Ganoderma lucidum* mycelium. *Journal of agricultural and food chemistry*, 62(3), 634-641. DOI: 10.1021/jf404533b



Wang, T. Y., Yu, C. C., Hsieh, P. L., Liao, Y. W., Yu, C. H., & Chou, M. Y. (2017). GMI ablates cancer stemness and cisplatin resistance in oral carcinomas stem cells through IL-6/Stat3 signaling inhibition. *Oncotarget*, 8(41), 70422. doi: 10.18632/oncotarget.19711

Wasser, S. P. (2014). Medicinal mushroom science: Current perspectives, advances, evidences, and challenges. *Biomedical journal*, 37(6), 345.

Wu, M.-Y., Hsu, M.-F., Huang, C.-S., Fu, H.-Y., Huang, C.-T., & Yang, C.-S. (2007). A 2.0 Å Structure of GMI, a Member of the Fungal Immunomodulatory Protein Family from *Ganoderma Microsporium*. *Protein Crystallography*, II, 132

Xu, J. W., Zhao, W., & Zhong, J. J. (2010). Biotechnological production and application of ganoderic acids. *Applied microbiology and biotechnology*, 87(2), 457-466. <https://doi.org/10.1007/s00253-010-2576-5>

Xu, X., Yan, H., Chen, J., & Zhang, X. (2011a). Bioactive proteins from mushrooms. *Biotechnology advances*, 29(6), 667-674.

Xu, Z., Chen, X., Zhong, Z., Chen, L., & Wang, Y. (2011b). *Ganoderma lucidum* polysaccharides: immunomodulation and potential anti-tumor activities. *The American journal of Chinese medicine*, 39(01), 15-27.

Xu, S., Dou, Y., Ye, B., Wu, Q., Wang, Y., Hu, M., ... & Guo, J. (2017). *Ganoderma lucidum* polysaccharides improve insulin sensitivity by regulating inflammatory cytokines and gut microbiota composition in mice. *Journal of Functional Foods*, 38, 545-552. <https://doi.org/10.1016/j.jff.2017.09.032>

Xu, L., Wang, H., & Ng, T. (2012). A laccase with HIV-1 reverse transcriptase inhibitory activity from the broth of mycelial culture of the mushroom *Lentinus tigrinus*. *BioMed Research International*.

Yau, T., Dan, X., Ng, C. C. W., & Ng, T. B. (2015). Lectins with potential for anti-cancer therapy. *Molecules*, 20(3), 3791-3810. <http://dx.doi.org/10.3390/molecules20033791>

Yamashita, S., & Hirose, D. (2016). Phylogenetic analysis of *Ganoderma australe* complex in a Bornean tropical rainforest and implications for mechanism of coexistence of various phylogenetic types. *Fungal Ecology*, 24, 1–6. <https://doi.org/10.1016/j.funeco.2016.04.006>

You, B. J., Lee, M. H., Tien, N., Lee, M. S., Hsieh, H. C., Tseng, L. H., ... & Lee, H. Z. (2013). A novel approach to enhancing ganoderic acid production by *Ganoderma lucidum* using apoptosis induction. *PloS one*, 8(1), e53616. <https://doi.org/10.1371/journal.pone.0053616>

Zhang, G., Sun, J., Wang, H., & Ng, T. B. (2010). First isolation and characterization of a novel lectin with potent antitumor activity from a *Russula* mushroom. *Phytomedicine*, 17(10), 775-781. <https://doi.org/10.1016/j.phymed.2010.02.001>

Zhao, M. W., Liang, W. Q., Zhang, D. B., Wang, N., Wang, C. G., & Pan, Y. J. (2007). Cloning and characterization of squalene synthase (SQS) gene from *Ganoderma lucidum*. *Journal of microbiology and biotechnology*, 17(7), 1106-1112.

Zhao, J., Zhou, L., Wang, J., Shan, T., Zhong, L., Liu, X., & Gao, X. (2010). Endophytic fungi for producing bioactive compounds originally from their host plants. *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*, 1, 567-576.

Zhong, J. J., & Xiao, J. H. (2009). Secondary metabolites from higher fungi: discovery, bioactivity, and bioproduction. *Advances in Biochemical Engineering / Biotechnology*, 113, 79–150. [https://doi.org/10.1007/10\\_2008\\_26](https://doi.org/10.1007/10_2008_26)

Zhou, X., Xie, M., Hong, F., & Li, Q. Z. (2009). Genomic cloning and characterization of a FIP-gsi gene encoding a fungal immunomodulatory protein from *Ganoderma sinense* Zhao et al.(Aphyllphoromycetideae). *International Journal of Medicinal Mushrooms*, 11(1).

Zhou, X.-W., Liu, Y., Guo, M.-Y., Su, K.-Q., & Zhang, Y.-M. (2016). Species Clarification of the Widely Cultivated *Ganoderma* in China Based on rDNA and FIP Gene Sequence Analysis. *International Journal of Agriculture and Biology*, (January 2016), 932–938. <https://doi.org/10.17957/IJAB/15.0189>

Zonga A., Caob H., Wanga F. (2012). Anticancer polysaccharides from natural resources: A review of recent research. *Carbohydrate Polymers*, 90, 1395-1410.

# ANEXOS

## Anexo I.

### A. Resultado de la secuenciación del genoma de *G. australe* por la tecnología de secuenciación Sequel de PacBio.

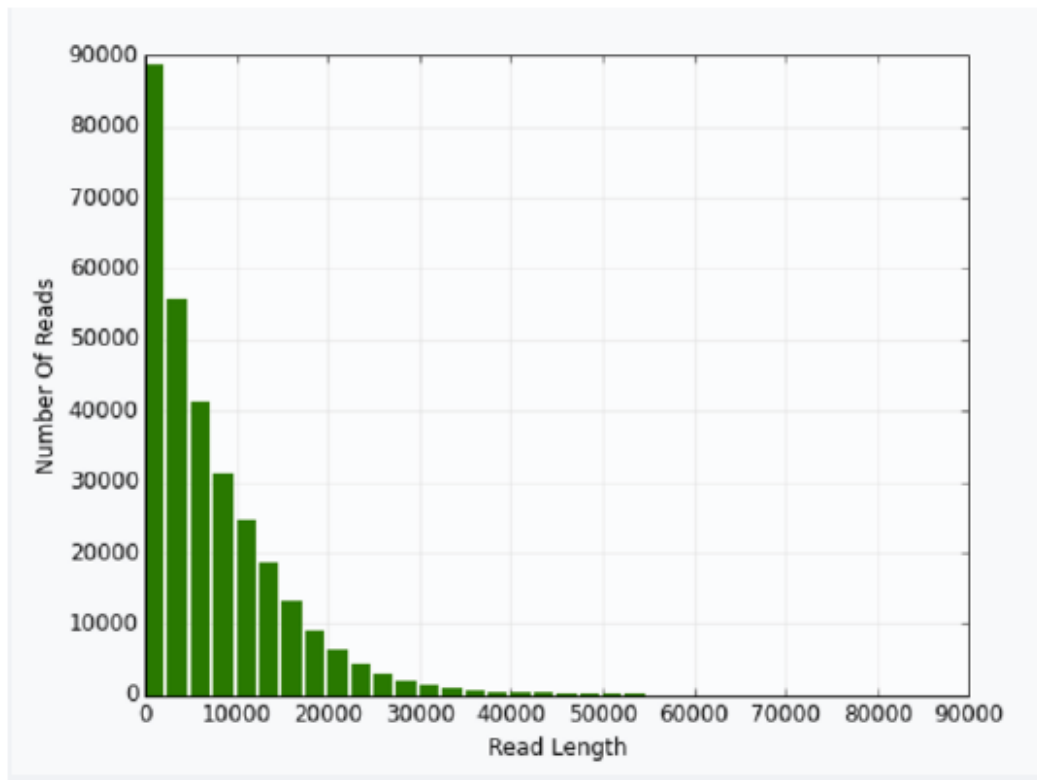
## 2. Data Overview

### 2.1 Polymerase Read Statistics

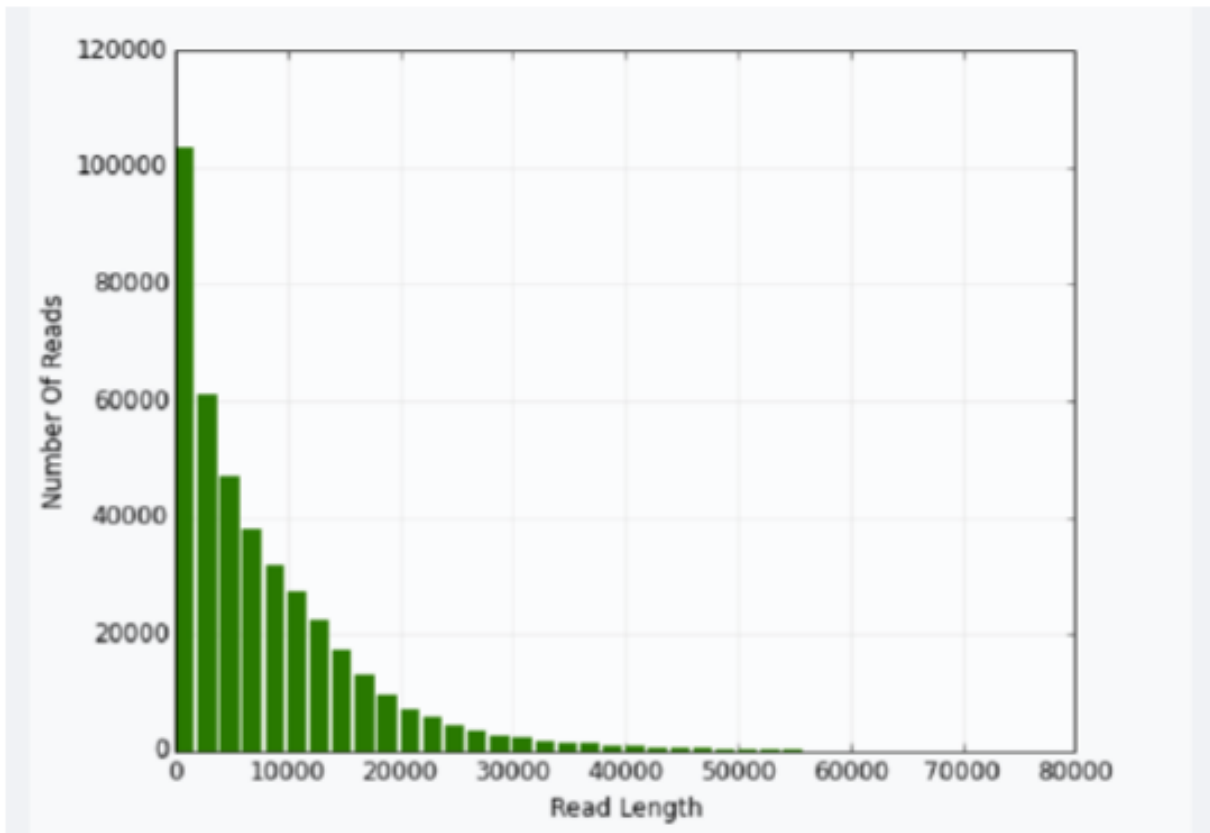
	Polymerase Read Bases	Polymerase Reads	Polymerase Read Length (mean)	Polymerase Read N50	Insert Length (mean)
Cell1	2,349,470,488	304,754	7,709	13,250	6,933
Cell2	3,324,105,593	406,351	8,180	14,250	7,195

### 2.2 Read Length Distribution

Cell 1:



Cell 2:



**B. Comandos utilizados para el ensamblaje del genoma de *G. australe* utilizando la herramienta Canu.**

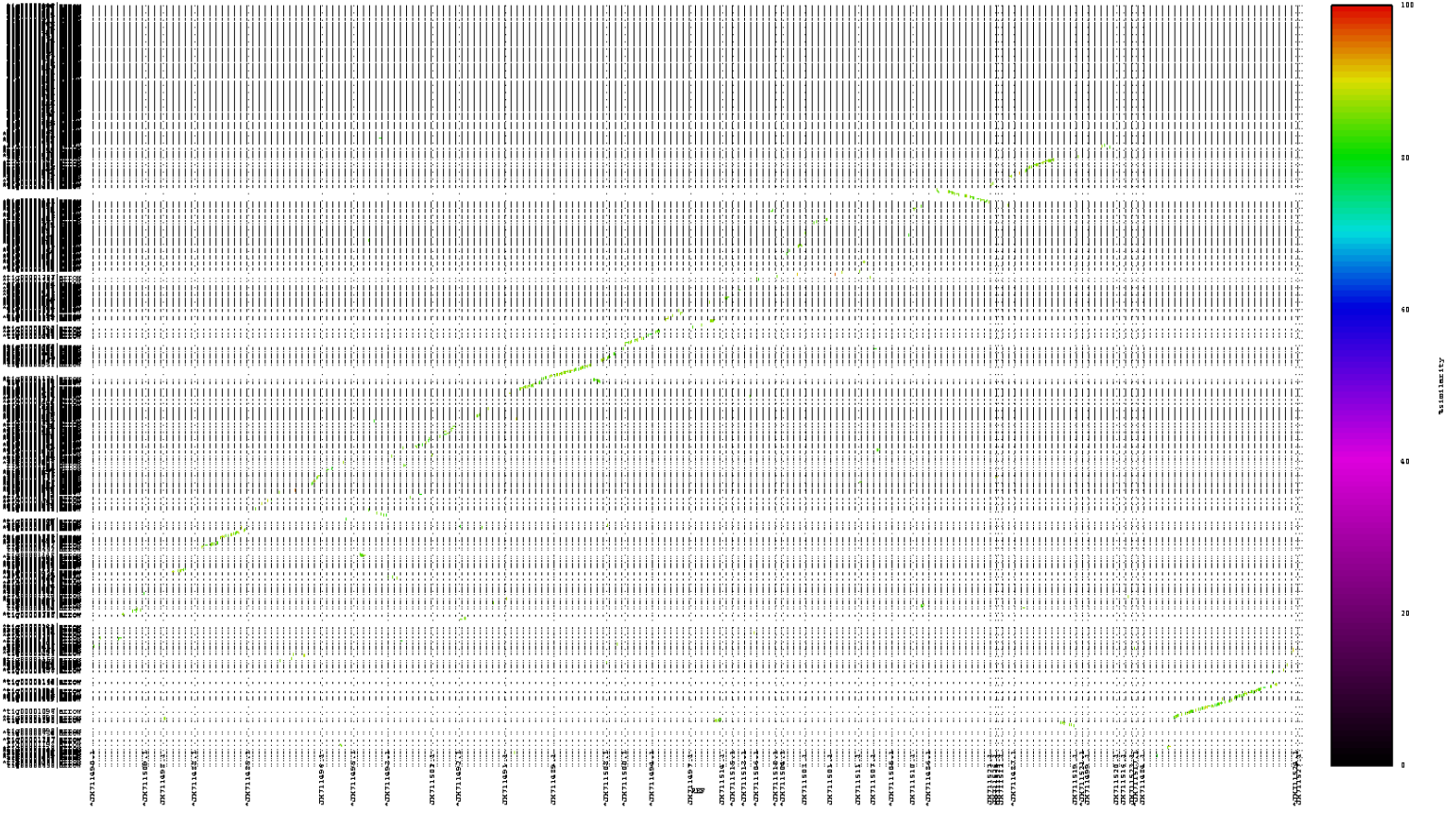
Los datos ingresados para realizar el ensamblaje con la herramienta Canu fue el siguiente:

Configuración se dejó por defecto.

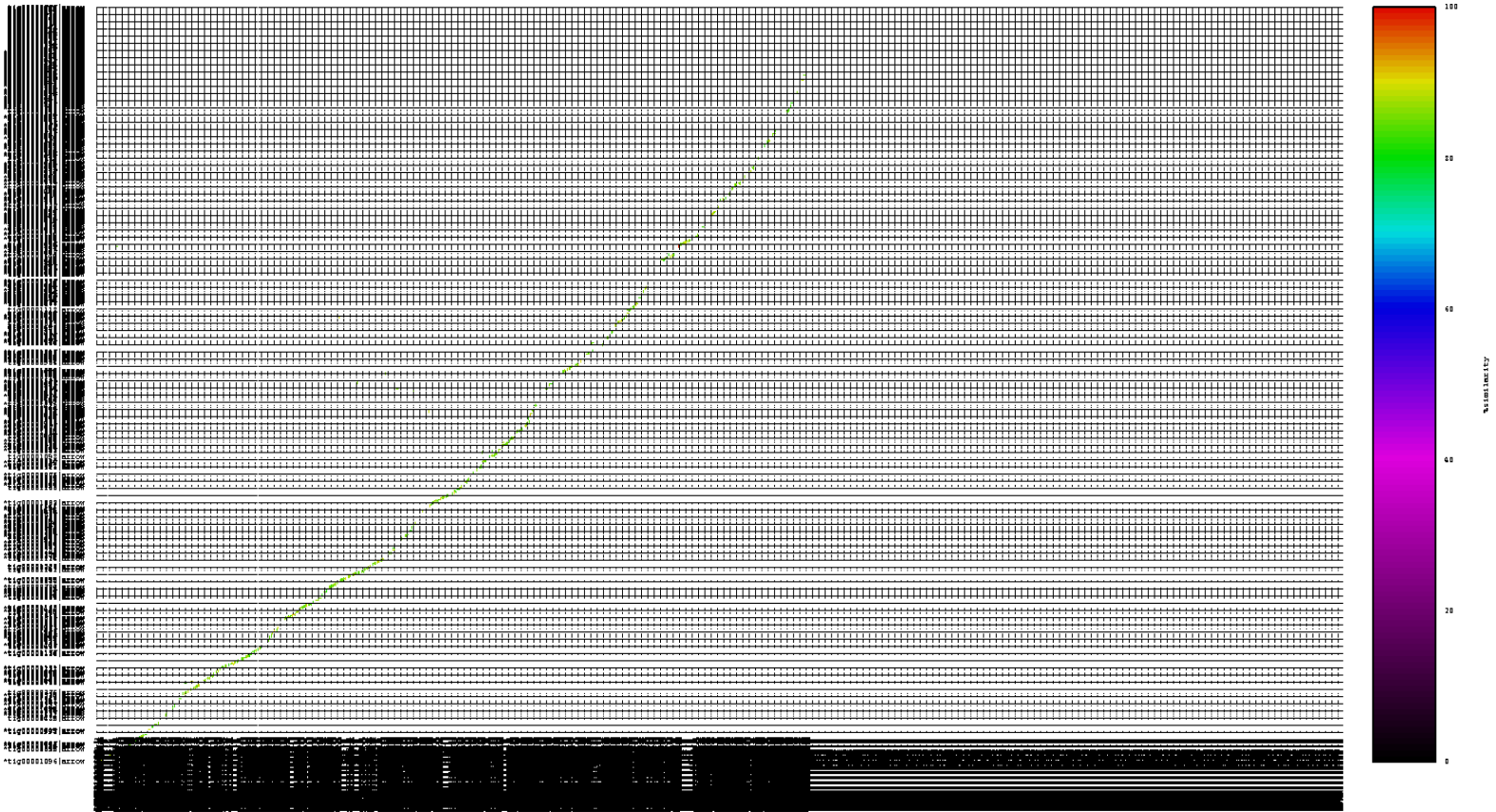
```
canu -p gaust -d canu_gaust genomeSize=40m -pacbio-raw /BIOS-Share/home/dagudelov/ensamblaje_pacbio/Gaust_pacbio.fasta
```

## Anexo II. Análisis del ensamblaje del genoma de *G. australe*

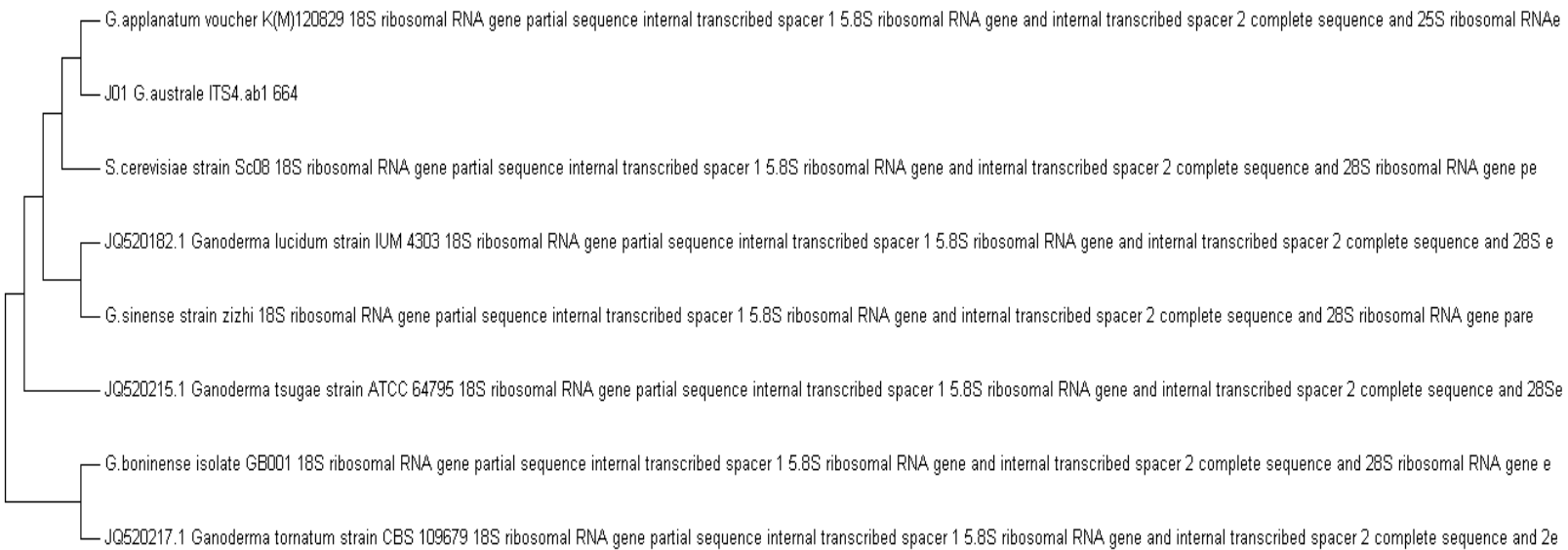
### A. Comparación genómica utilizando la herramienta Mummer de los genomas de *G. australe* vs *G. lucidum*



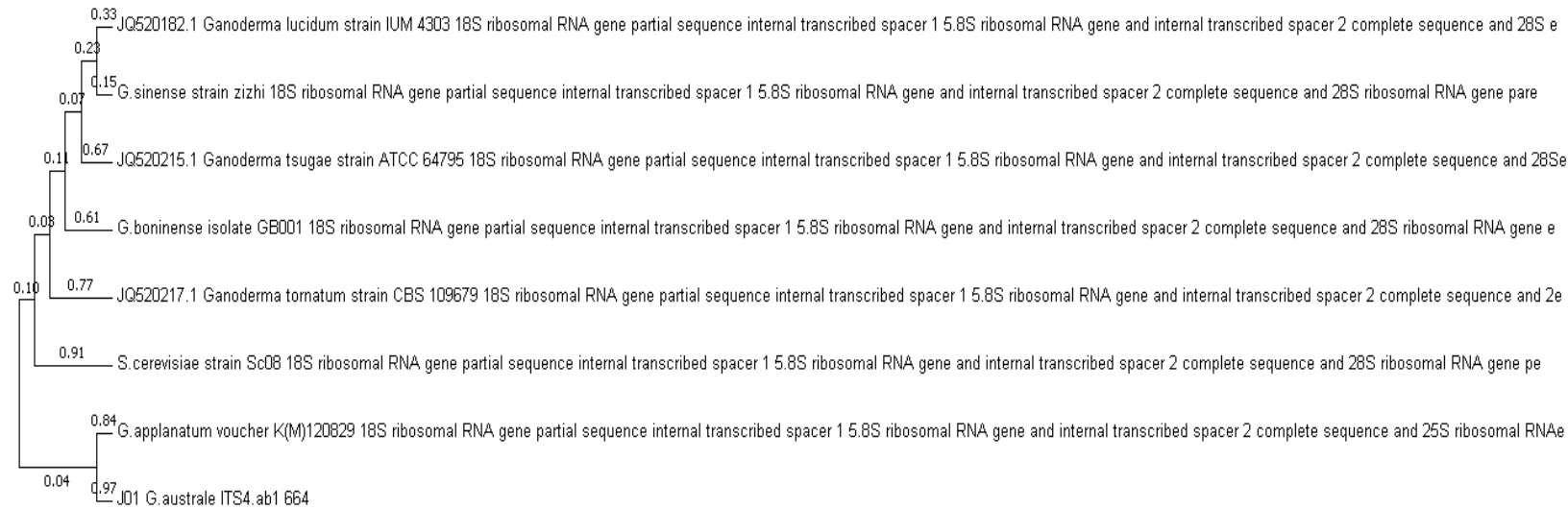
**B. Comparación genómica utilizando la herramienta Mummer de los genomas de *G. australe* vs *G. boninense***



### C. Filogenia generada con la herramienta Mega7 con el método de máxima parsimonia

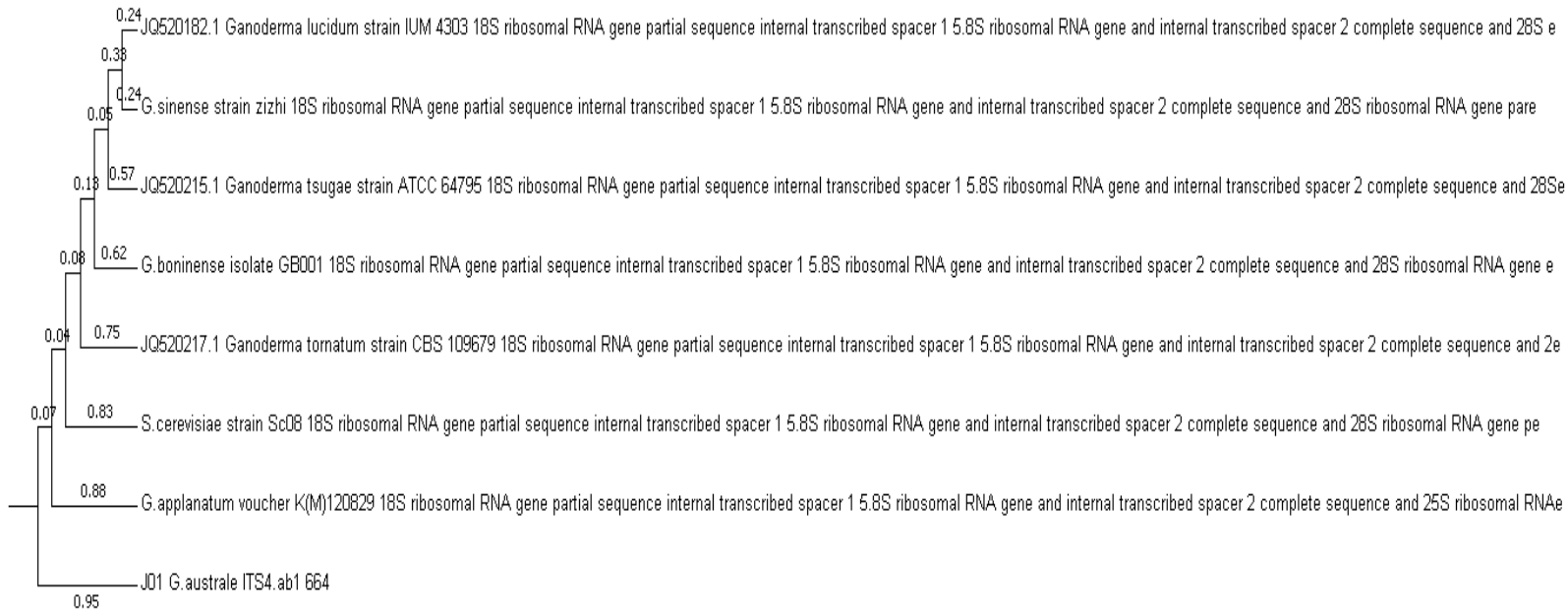


## D. Filogenia generada con la herramienta Mega7 con el método de Neighbour-Joining

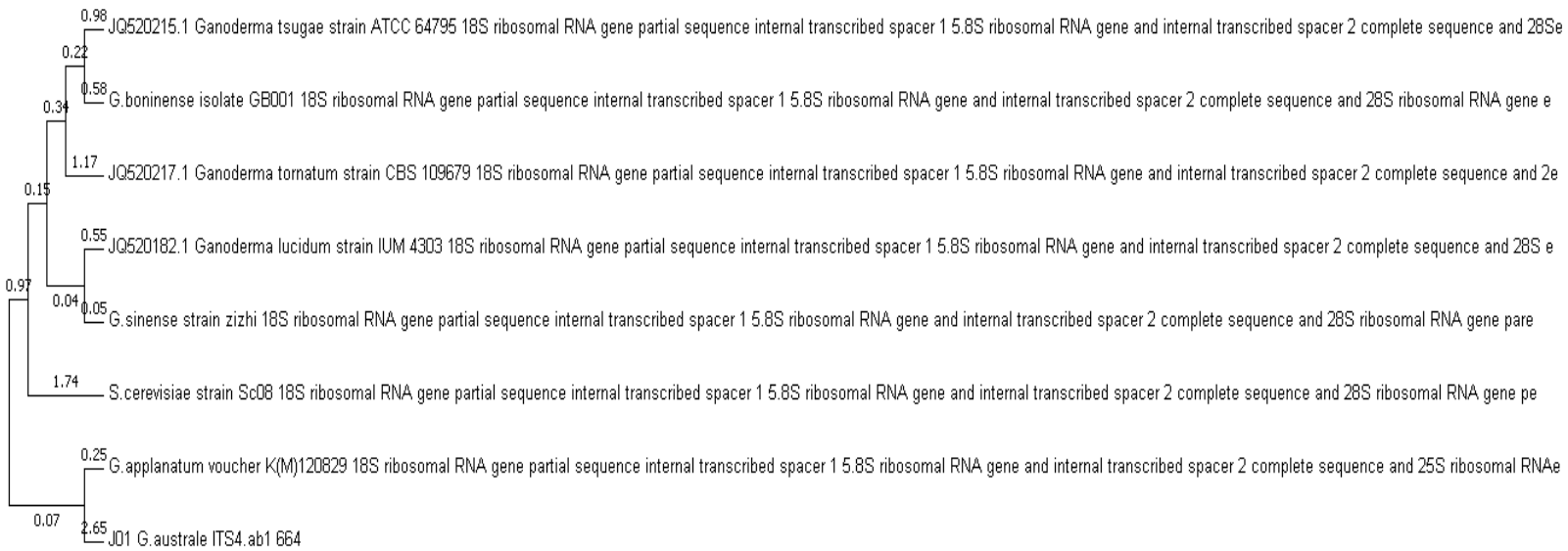




## E. Filogenia generada con la herramienta Mega7 con el método de UPGMA



**F. Filogenia generada con la herramienta Mega7 con el método del máximo probable (Maximum-likelihood)**



### Anexo III. Configuración para la anotación del ensamblaje del genoma de *G. australe* con la herramienta Maker.

#### A. Archivo de configuración para el entrenamiento del modelo con el genoma de *G. lucidum*.

```
#-----Genome (these are always required)
genome=/BIOS-
Share/home/dagudelov/ensamblaje_pacbio/canu/maker/G_lucidum_genome_assemb
led.fasta #genome sequence (fasta file or fasta embeded in GFF3 file)
organism_type=eukaryotic #eukaryotic or prokaryotic. Default is
eukaryotic

#-----Re-annotation Using MAKER Derived GFF3
maker_gff= #MAKER derived GFF3 file
est_pass=0 #use ESTs in maker_gff: 1 = yes, 0 = no
altest_pass=0 #use alternate organism ESTs in maker_gff: 1 = yes, 0
= no
protein_pass=0 #use protein alignments in maker_gff: 1 = yes, 0 = no
rm_pass=0 #use repeats in maker_gff: 1 = yes, 0 = no
model_pass=0 #use gene models in maker_gff: 1 = yes, 0 = no
pred_pass=0 #use ab-initio predictions in maker_gff: 1 = yes, 0 = no
other_pass=0 #passthrough anything else in maker_gff: 1 = yes, 0 =
no

#-----EST Evidence (for best results provide a file for at least
one)
est=/BIOS-
Share/home/dagudelov/ensamblaje_pacbio/canu/maker/Gansp1_ESTs_20110326_CH
GH.good_ESTs.fasta #set of ESTs or assembled mRNA-seq in fasta format
altest= #EST/cDNA sequence file in fasta format from an alternate
organism
est_gff= #aligned ESTs or mRNA-seq from an external GFF3 file
altest_gff= #aligned ESTs from a closely relate species in GFF3 format

#-----Protein Homology Evidence (for best results provide a file for
at least one)
protein=/BIOS-
Share/home/dagudelov/DBs/Ganoderma_Mycocosm_Filtered_Models/Proteins/Gans
p1_GeneCatalog_proteins_20110326.aa.fasta #protein sequence file in fasta
format (i.e. from mutiple oransisms)
protein_gff= #aligned protein homology evidence from an external
GFF3 file

#-----Repeat Masking (leave values blank to skip repeat masking)
model_org=all #select a model organism for RepBase masking in
RepeatMasker
```

```

    rmlib= #provide an organism specific repeat library in fasta format
for RepeatMasker
    repeat_protein=/BIOS-
Share/Software/bioinformatics/maker/2.31.8/data/te_proteins.fasta
#provide a fasta file of transposable element proteins for RepeatRunner
    rm_gff= #pre-identified repeat elements from an external GFF3 file
    prok_rm=0 #forces MAKER to repeatmask prokaryotes (no reason to
change this), 1 = yes, 0 = no
    softmask=1 #use soft-masking rather than hard-masking in BLAST (i.e.
seg and dust filtering)

#-----Gene Prediction
snaphmm= #SNAP HMM file
gmhmm= #GeneMark HMM file
augustus_species= #Augustus gene prediction species model
fgenesh_par_file= #FGENESH parameter file
pred_gff= #ab-initio predictions from an external GFF3 file
model_gff=/BIOS-
Share/home/dagudelov/ensamblaje_pacbio/canu/maker/modelo1/Gansp1.filtered
_proteins.FilteredModels1.gff3 #annotated gene models from an external
GFF3 file (annotation pass-through)
    est2genome=1 #infer gene predictions directly from ESTs, 1 = yes, 0
= no
    protein2genome=1 #infer predictions from protein homology, 1 = yes,
0 = no
    trna=1 #find tRNAs with tRNAscan, 1 = yes, 0 = no
    snoscan_rrna= #rRNA file to have Snoscan find snoRNAs
    unmask=1 #also run ab-initio prediction programs on unmasked
sequence, 1 = yes, 0 = no

#-----Other Annotation Feature Types (features MAKER doesn't
recognize)
    other_gff= #extra features to pass-through to final MAKER generated
GFF3 file

#-----External Application Behavior Options
    alt_peptide=C #amino acid used to replace non-standard amino acids
in BLAST databases
    cpus=32 #max number of cpus to use in BLAST and RepeatMasker (not
for MPI, leave 1 when using MPI)

#-----MAKER Behavior Options
    max_dna_len=100000 #length for dividing up contigs into chunks
(increases/decreases memory usage)
    min_contig=1 #skip genome contigs below this length (under 10kb are
often useless)

    pred_flank=200 #flank for extending evidence clusters sent to gene
predictors

```

```

    pred_stats=0 #reporqt AED and QI statistics for all predictions as
well as models
    AED_threshold=1 #Maximum Annotation Edit Distance allowed (bound by
0 and 1)
    min_protein=0 #require at least this many amino acids in predicted
proteins
    alt_splice=0 #Take extra steps to try and find alternative splicing,
1 = yes, 0 = no
    always_complete=0 #extra steps to force start and stop codons, 1 =
yes, 0 = no
    map_forward=0 #map names and attributes forward from old GFF3 genes,
1 = yes, 0 = no
    keep_preds=0 #Concordance threshold to add unsupported gene
prediction (bound by 0 and 1)

    split_hit=10000 #length for the splitting of hits (expected max
intron size for evidence alignments)
    single_exon=0 #consider single exon EST evidence when generating
annotations, 1 = yes, 0 = no
    single_length=250 #min length required for single exon ESTs if
'single_exon is enabled'
    correct_est_fusion=0 #limits use of ESTs in annotation to avoid
fusion genes

    tries=2 #number of times to try a contig if there is a failure for
some reason
    clean_try=0 #remove all data from previous run before retrying, 1 =
yes, 0 = no
    clean_up=0 #removes theVoid directory with individual analysis
files, 1 = yes, 0 = no
    TMP= #specify a directory other than the system default temporary
directory for temporary files

```

## B. Archivo de configuración para la anotación del genoma de *G. australe*.

```
#-----Genome (these are always required)
genome=/BIOS-
Share/home/dagudelov/ensamblaje_pacbio/canu/canu_gaust/canu_arrow.consensus.fasta #genome sequence (fasta file or fasta embedded in GFF3 file)
organism_type=eukaryotic #eukaryotic or prokaryotic. Default is eukaryotic

#-----Re-annotation Using MAKER Derived GFF3
maker_gff= #MAKER derived GFF3 file
est_pass=0 #use ESTs in maker_gff: 1 = yes, 0 = no
altest_pass=0 #use alternate organism ESTs in maker_gff: 1 = yes, 0
= no
protein_pass=0 #use protein alignments in maker_gff: 1 = yes, 0 = no
rm_pass=0 #use repeats in maker_gff: 1 = yes, 0 = no
model_pass=0 #use gene models in maker_gff: 1 = yes, 0 = no
pred_pass=0 #use ab-initio predictions in maker_gff: 1 = yes, 0 = no
other_pass=0 #passthrough anything else in maker_gff: 1 = yes, 0 =
no

#-----EST Evidence (for best results provide a file for at least
one)
est= #set of ESTs or assembled mRNA-seq in fasta format
altest=/BIOS-
Share/home/dagudelov/ensamblaje_pacbio/canu/maker/Gansp1_ESTs_20110326_CH
GH.good_ESTs.fasta #EST/cDNA sequence file in fasta format from an
alternate organism
est_gff= #aligned ESTs or mRNA-seq from an external GFF3 file
altest_gff= #aligned ESTs from a closely relate species in GFF3 format

#-----Protein Homology Evidence (for best results provide a file for
at least one)
protein=/BIOS-
Share/home/dagudelov/DBs/Ganoderma_prot_sequences.fasta #protein sequence
file in fasta format (i.e. from mutiple oransisms)
protein_gff= #aligned protein homology evidence from an external
GFF3 file

#-----Repeat Masking (leave values blank to skip repeat masking)
model_org=all #select a model organism for RepBase masking in
RepeatMasker
rmlib= #provide an organism specific repeat library in fasta format
for RepeatMasker
repeat_protein=/BIOS-
Share/Software/bioinformatics/maker/2.31.8/data/te_proteins.fasta
#provide a fasta file of transposable element proteins for RepeatRunner
rm_gff= #pre-identified repeat elements from an external GFF3 file
prok_rm=0 #forces MAKER to repeatmask prokaryotes (no reason to
change this), 1 = yes, 0 = no
```

```
softmask=1 #use soft-masking rather than hard-masking in BLAST (i.e.
seg and dust filtering)
```

```
#-----Gene Prediction
snaphmm=/BIOS-
Share/home/dagudelov/ensamblaje_pacbio/canu/maker/G_lucidum_genome_assemb
led.maker.output/Foxy1.hmm #SNAP HMM file
gmhmm=/BIOS-
Share/home/dagudelov/ensamblaje_pacbio/canu/maker/output/gmhmm.mod
#GeneMark HMM file
augustus_species=laccaria_bicolor #Augustus gene prediction species
model
fgenesh_par_file= #FGENESH parameter file
pred_gff= #ab-initio predictions from an external GFF3 file
model_gff= #annotated gene models from an external GFF3 file
(annotation pass-through)
est2genome=0 #infer gene predictions directly from ESTs, 1 = yes, 0
= no
protein2genome=0 #infer predictions from protein homology, 1 = yes,
0 = no
trna=1 #find tRNAs with tRNAscan, 1 = yes, 0 = no
snoscan_rrna= #rRNA file to have Snoscan find snoRNAs
unmask=1 #also run ab-initio prediction programs on unmasked
sequence, 1 = yes, 0 = no
```

```
#-----Other Annotation Feature Types (features MAKER doesn't
recognize)
other_gff= #extra features to pass-through to final MAKER generated
GFF3 file
```

```
#-----External Application Behavior Options
alt_peptide=C #amino acid used to replace non-standard amino acids
in BLAST databases
cpus=32 #max number of cpus to use in BLAST and RepeatMasker (not
for MPI, leave 1 when using MPI)
```

```
#-----MAKER Behavior Options
max_dna_len=100000 #length for dividing up contigs into chunks
(increases/decreases memory usage)
min_contig=1 #skip genome contigs below this length (under 10kb are
often useless)
```

```
pred_flank=200 #flank for extending evidence clusters sent to gene
predictors
pred_stats=0 #reporqt AED and QI statistics for all predictions as
well as models
AED_threshold=1 #Maximum Annotation Edit Distance allowed (bound by
0 and 1)
min_protein=0 #require at least this many amino acids in predicted
proteins
```

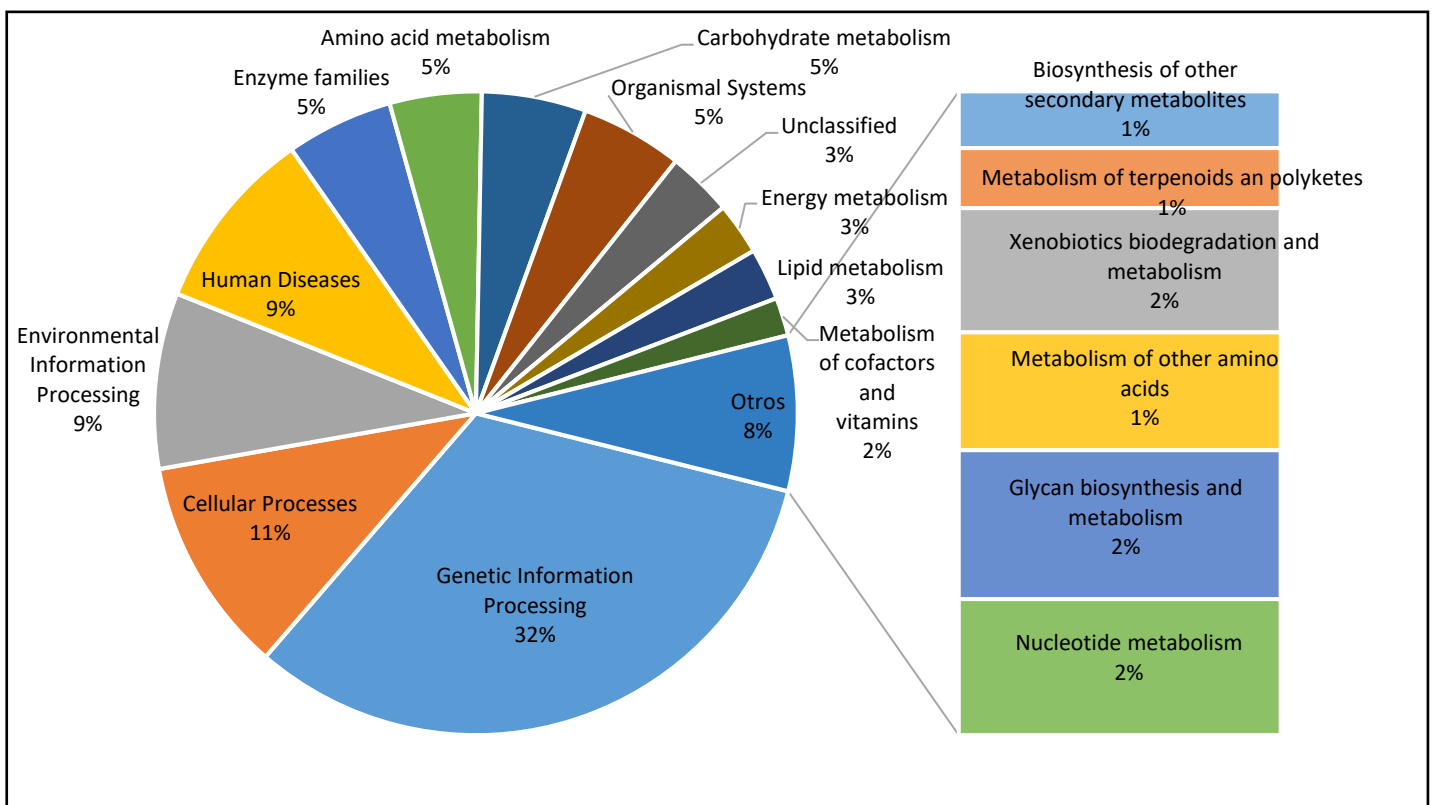
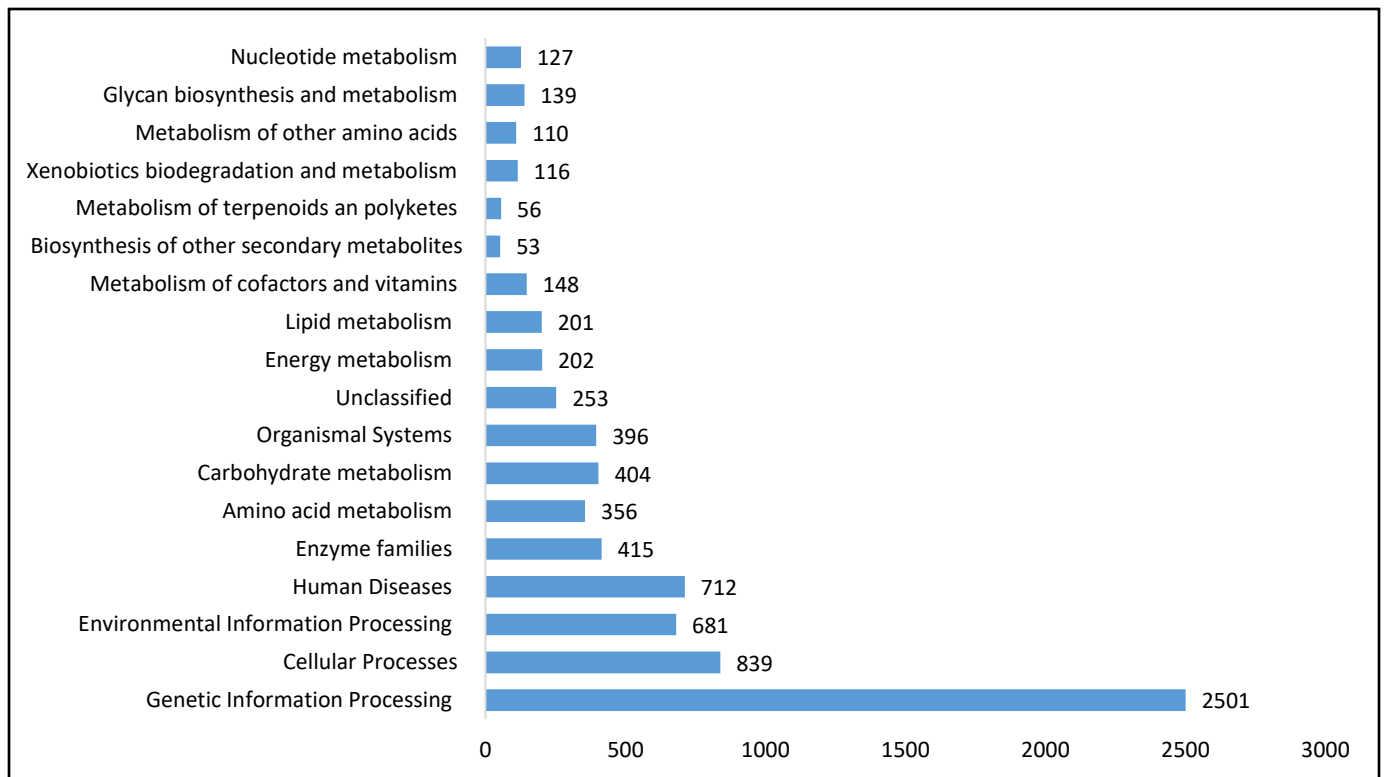
```
    alt_splice=0 #Take extra steps to try and find alternative splicing,
1 = yes, 0 = no
    always_complete=0 #extra steps to force start and stop codons, 1 =
yes, 0 = no
    map_forward=0 #map names and attributes forward from old GFF3 genes,
1 = yes, 0 = no
    keep_preds=0 #Concordance threshold to add unsupported gene
prediction (bound by 0 and 1)

    split_hit=10000 #length for the splitting of hits (expected max
intron size for evidence alignments)
    single_exon=0 #consider single exon EST evidence when generating
annotations, 1 = yes, 0 = no
    single_length=250 #min length required for single exon ESTs if
'single_exon is enabled'
    correct_est_fusion=0 #limits use of ESTs in annotation to avoid
fusion genes

    tries=2 #number of times to try a contig if there is a failure for
some reason
    clean_try=0 #remove all data from previous run before retrying, 1 =
yes, 0 = no
    clean_up=0 #removes theVoid directory with individual analysis
files, 1 = yes, 0 = no
    TMP= #specify a directory other than the system default temporary
directory for temporary files
```



### C. Anotación de las secuencias predichas por Maker2 en el servidor de KAAS.



**D. Tabla consenso de las regiones repetitivas identificadas con la herramienta RepeatMasker en el genoma de *G. australe***

```
=====
file name: canu_arrow.consensus.fasta
sequences:          776
total length:      84254787 bp (84254787 bp excl N/X-runs)
GC level:          55.57 %
bases masked:      2683035 bp ( 3.18 %)
```

```
=====
```

	number of elements*	length occupied	percentage of sequence
-----			
Retroelements	1606	1835486 bp	2.18 %
SINEs:	0	0 bp	0.00 %
Penelope	9	8750 bp	0.01 %
LINEs:	65	41223 bp	0.05 %
CRE/SLACS	0	0 bp	0.00 %
L2/CR1/Rex	0	0 bp	0.00 %
R1/LOA/Jockey	0	0 bp	0.00 %
R2/R4/NeSL	0	0 bp	0.00 %
RTE/Bov-B	0	0 bp	0.00 %
L1/CIN4	0	0 bp	0.00 %
LTR elements:	1541	1794263 bp	2.13 %
BEL/Pao	0	0 bp	0.00 %
Ty1/Copia	226	190587 bp	0.23 %
Gypsy/DIRS1	1314	1603039 bp	1.90 %
Retroviral	0	0 bp	0.00 %
DNA transposons	177	113270 bp	0.13 %
hobo-Activator	3	261 bp	0.00 %
Tc1-IS630-Pogo	89	75125 bp	0.09 %
En-Spm	0	0 bp	0.00 %

MuDR-IS905	0	0 bp	0.00 %
PiggyBac	0	0 bp	0.00 %
Tourist/Harbinger	19	6828 bp	0.01 %
Other (Mirage, P-element, Transib)	0	0 bp	0.00 %
Rolling-circles	0	0 bp	0.00 %
Unclassified:	23	6249 bp	0.01 %
Total interspersed repeats:		1955005 bp	2.32 %
Small RNA:	24	31431 bp	0.04 %
Satellites:	0	0 bp	0.00 %
Simple repeats:	13117	595790 bp	0.71 %
Low complexity:	2003	100876 bp	0.12 %

=====

\* most repeats fragmented by insertions or deletions  
have been counted as one element  
Runs of >=20 X/Ns in query were excluded in % calcs

The query species was assumed to be fungi  
RepeatMasker version open-4.0.6 , default mode

run with rmblastn version 2.2.27+

**Anexo IV. Estadísticas obtenidas del aislamiento de la secuencia FIP putativa del genoma ensamblado de *G. australe*.**

**A. Secuencias utilizadas en Blast y Hmmer para la búsqueda de FIPs en el genoma de *G. australe***

gi|47156029|gb|AY569691.1| DE Taiwanofungus camphoratus immunomodulatory protein (Aca1) mRNA, complete cds, Synonyms include Antrodia camphorata, Antrodia ci

gi|598011088|ref|XM\_007336979.1| DE Auricularia subglabra TFB-10046 SS5 immunomodulatory protein partial mRNA

gi|460002059|gb|KC297708.1| DE Trametes versicolor immunomodulatory protein (YZip) mRNA, complete cds

gi|38455234|gb|AY449805.1| DE Ganoderma sinense strain Zi-Zhi immunomodulatory protein gene, partial sequence

gi|344055436|gb|JN571084.1| DE Wolfiporiaporia cocos isolate PCP 080110 immunomodulatory protein (PCP) mRNA, complete cds

gi|119710155|gb|EF152774.1| DE Tremella fuciformis immunomodulatory protein (TFP1) mRNA, complete cds

gi|756420805|gb|KJ818121.1| DE Postia placenta isolate Mad-698-R immunomodulatory protein FIP-pp1 mRNA, partial cds

gi|302889536|ref|XM\_003043608.1| DE Nectria haematococca mpVI 77-13-4 hypothetical protein, mRNA

gi|283488735|gb|GU388420.1| DE Flammulina velutipes immunomodulatory protein mRNA, partial cds

gi|597978930|ref|XM\_007362804.1| DE Dichomitus squalens LYAD-421 SS1 immunomodulatory protein FIP-Fve (DICSQDRAFT\_125142), partial mRNA

gi|910269404|gb|KT210088.1| DE Chroogomphus rutilus strain ACCC 51531 fungal immunomodulatory protein (FIP) gene, complete cds

gi|348167217|gb|JN167598.1| DE Ganoderma applanatum fungal immunomodulatory protein (FIP) gene, complete cds

gi|187961979|gb|EU680479.1| DE Ganoderma lucidum immunomodulatory protein 8 gene, complete cds

gi|746719557|gb|KM077027.1| DE Ganoderma atrum immunomodulatory protein gene, complete cds

gi|62739081|gb|AY987805.1| DE Ganoderma japonicum immunomodulatory protein (Fip-gja) gene, complete cds

- Secuencias de proteínas del repositorio UniProt usadas en la búsqueda de FIPs en el genoma de *G. australe*

>sp|P80412|F.velutipes

SATSLTFQLAYLVKKIDFDYTPNWGRGTPSSYIDNLTFPKVLTDKKYSYRWWWNGSDLGV  
ESNFAVTPSGGQTINFLQYNKGYVADTKTIQVFVWIPDTGNSEEEYIIAEWKKT

>sp|P14945|LZ8\_G.lucidum

MSDTALIFRLAWDVKKLSFDYTPNWGRGNPNFIDTVTFPKVLTDKAYTYRVAVSGRNLG  
VKPSYAVESDGSQKVNFLYNSGYGIADTNTIQVFVWDPDTNNDFIIAQWN

>tr|A0A0K1L6W2|C.rutilus

MSETVTALAFRLAWEVKKISFDYTPNWGRGSPSSYIDNLTFPKVLTDKAYTCRWWWSGRD  
LGVRPSYEVASDGSQKINFLEYHNGYGIADTNTIQVYVWDPSTGDDFIIAQWN

>tr|G5CJT8|G.applanatum

MSETVTALTFRRLAWEVKKISFDYTPNWGRGSPSSYIDNLTFPKVLTDKAYTYRWWWSGRD  
LGVRPSYEVASDGSQKINFLEYHNGYGIADTNTIQVYVWDPSTGDDFIIAQWN

>tr|E7FH75|G.microsporum

MSDTALIFTLAWNVKQLAFDYTPNWGRGRPSSFIDTVTFPTVLTDKAYTYRWWWSGKDLG  
VRPSYAVESDGSQKINFLEYNSGYGIADTNTIQVYVIDPDTGNNFIVAQWNYLEQKLISE  
EDLNSAVDHHHHHH

>tr|A0A0C5BFA5|P.placenta

CITYVLQVESWRSLDRLVPRAGKPFVDFDYTPKWKRVEWGAPGFYLDVTFPKVLKDKR  
YKYRWWWSGSTDLGIREPYDVQADGSQRINFLEYNQGYGIRDNTNTVTVYLVDPETATEYKL  
TQSQ

>tr|A0A0B4ZUG9|G.atrum

MSDTALIFRLAWDVKKLSFDYTPNWGRGNPSRFVDNVTFPQVLADKAYTYRWWWSGRDLDG  
VRPSYAVGSDGSQKVNFLYNSQSYGIADTNTIQVFVIDPDTGADFIIAQWN

**B. Búsqueda de Hmmer contra las FIPs de Swissprot (*G. lucidum* y *F. velutipes*)**

phmmer results for job BB9716B6-B87A-11E7-BAD2-92B253F04F9B.1:

```
Target Num-hits Bias Bit-Score E-value Tax-Id Species Description
=====
=====
LZ8_GANLU 1 0.4 190.0 2.6e-55 5315 Ganoderma lucidum
Immunomodulatory protein Ling Zhi-8
-----
-----
Target-env-start Target-env-end Target-ali-start Target-ali-end
Query-start Query-end E-value(cond.)
-----
-----
1 111 4 110 6 112 1.0e-60
QUERY
taltfrlayevkkisfdytphwgrgtpssfidnltpkvltdkaykyrvvvagrldgvrpsyevsdgsqkin
fleyhngfgiadntntiqvfvdptsgddfiaqw
MATCH tal frla++vkk+sfdytp+wrg p++fid +tfpkvltdkay yrv
v+gr+lgv+psy v sdgsqk+nfley++g+giadtntiqvfvdv +++dfiaqw
PP
89*****
*****
TARGET
TALIFRLAWDVKKLSFDYTPNWGRGNPNFIDTVTFPKVLTDKAYTYRVAVSGRNLGVKPSYAVESDGSQKVN
FLEYN SGYGIADTNTIQVFVDPDTNNDFIIAQW
-----
-----
=====
=====
FVE_FLAVE 1 0.5 163.3 5.1e-47 38945 Flammulina velutipes
Immunomodulatory protein FIP-Fve
```

```

-----
-----
Target-env-start Target-env-end Target-ali-start Target-ali-end
Query-start Query-end E-value(cond.)
-----
-----
1 112 2 112 5 113 2.1e-52
QUERY
vtaltfrlayevkkisfdytphwgrgtpssfidnltpkvltdkaykyrvvvagrldlgvrpsyevasdgsqki
nfleyhngfgiadntiqvfvdptsg..ddfiaqwk
MATCH t+ltf+lay vkki fdytp+wgrgtpss+idnltpkvltdk y yrvvv g dlgv+
++ v g q infl+y+ g+g+adt tiqvfvv p +g +++iia+wk
PP
699*****
*****998873368*****8
TARGET
ATSLTFQLAYLVKKIDFDYTPNWGRGTPSSYIDNLTFPKVLTDKKYSYRVVWNGSDLGVESNFAVTPSGGQTI
NFLQYNKGYGVADTKTIQVFVWIPDTGnsEEYIIAEWK

```

```

-----
-----
=====
=====

```

Search Details

```

=====
Date Started: 2017-09-20 06:18:13
Cmd: phmmer -E 1 --domE 1 --incE 0.01 --incdomE 0.03 --mx BLOSUM62
--pextend 0.4 --popen 0.02 --seqdb swissprot
Database: swissprot, version 2017_09, downloaded on 2017-09-18
Query: >Query
MSETVTALTFRLAYEVKKISFDYTPHWGRGTPSSYIDNLTFPKVLTDK
AYKYRVVWAGRDLGVRPSYEVASDGSQKINFLEYHNGFGIADTNTIQV
FVVDPTSGDDFIIAQWK

```

Stats

=====

nhits:2

elapsed:0.07

Z:555594

Z\_setby:0

n\_past\_msv:11140

unpacked:2

user:0

domZ\_setby:0

nseqs:555594

n\_past\_bias:10214

sys:0

n\_past\_fwd:5

nmodels:1

nincluded:2

n\_past\_vit:524

nreported:2

domZ:2



### C. Predicción de la estructura del gen putativo de la FIP de *G. australe*

- Predicción de promotor y caja TATA con la herramienta Fgenesh de Softberry

```
>tig00001168|arrow:subseq(267872,2000)
```

```
Length of sequence-      2000
```

```
Thresholds for TATA+ promoters - 0.02, for TATA-/enhancers - 0.04
```

```
1 promoter/enhancer(s) are predicted
```

```
Promoter Pos:      820 LDF- 0.02 TATA box at      786      19.80
```

```
Transcription factor binding sites/RegSite DB:
```

```
for promoter at position -      820
```

```
734 (+) RSP00004      tagaCACGTaga
743 (-) RSP00010      cACGTG
746 (-) RSP00011      ctccACGTGgt
709 (-) RSP00016      caTGCAC
585 (-) RSP00026      gcttttgaTGACtTcaaacac
742 (-) RSP00065      ACGTGgcgc
742 (-) RSP00066      ACGTGccgc
743 (-) RSP00069      tACGTG
755 (+) RSP00096      GGTTT
580 (-) RSP00102      aaaTGACGaaaatgc
793 (+) RSP00161      WAAAG
577 (-) RSP00214      TGACGTgg
592 (-) RSP00269      atcttatgtcattgaTGACGacctcc
549 (+) RSP00284      CCGTCC
798 (-) RSP00305      CCTTTT
570 (-) RSP00308      CAACA
560 (-) RSP00400      CGGTTG
565 (-) RSP00431      rdygRCRGTTTRs
572 (+) RSP00437      ACGTCA
665 (+) RSP00469      GNGGTG
679 (+) RSP00469      GNGGTG
812 (-) RSP00483      GCCGC
627 (-) RSP00483      GCCGC
789 (-) RSP00501      CGAGCAg
```



CTCTGTAAACGCCGTGTCAGCGACCTTTGTACCGGAAACCCGTGGCATGTATTGATCCTA	1740
CCATGTCTCCACGGACGGCGAAGCCGTGGACGCACCCGTCCCAGTACGTGCGCTCGTAGC	1800
CAGGCGGAACTTCCCGCTAAACATCTCATCTGCAAGTTTTTGGCCCTCCTGCGGGAACA	1860
TCGGGTCCGTCTCGCACGAGTTGATCAGCAGCGGCCCTTCGACTTCTCAAAATATTCCT	1920
ACAGAATCCAAGTTTTCCGCTGGTCAGTCACCGTCACTCCTGGACGAACTAATGCATGGTA	1980
TCGCTGCTCACGTGGAGATC	2040

PREDICTED TRANSCRIPTION START SITES:

tig00001168\_arrow\_subseq\_267872,2000\_, 2000 nucleotides

Position	Score	Likelihood
600	0.605	Marginal prediction
900	1.399	Highly likely prediction
1600	0.631	Marginal prediction

Position Score Likelihood

where 'Position' is a position in the sequence, 'Score' is the prediction score for a transcription start site occurring within 100 base pairs upstream from that position and 'Likelihood' is a descriptive label associated with that score. The scores are always positive numbers; they are labelled as follows:

below 0.5	ignored
0.5 - 0.8	Marginal prediction
0.8 - 1.0	Medium likely prediction
above 1.0	Highly likely prediction



- Alineamiento de FIPs con MUSCLE en formato ClustalO

CLUSTAL multiple sequence alignment by MUSCLE (3.8)

```

tr|A0A0C5BFA5|Postia          -----CITYVLQVESWRSLDRLVPRAGKPFVDFDYTPKWKRVEMGAPGQFYLDEVTFPK
sp|P80412|Flammulina         SA---TSLTFQLAYLVKK-----IDFDYTPNWGR---GTPSS-YIDNLTFPK
genemark-tig00001168|Ganoderma MSETVTALTFRLAYEVKK-----ISFDYTPNWGR---GTPSS-FIDNLTFPK
tr|A0A0K1L6W2|Chroogomphus   MSETVTALAFRLANVKK-----ISFDYTPNWGR---GSPSS-YIDNLTFPK
tr|G5CJT8|Ganoderma          MSETVTALTFRLANVKK-----ISFDYTPNWGR---GSPSS-YIDNLTFPK
tr|E7FH75|Ganoderma          MSD--TALIFTLANVVKQ-----LAFDYTPNWGR---GRPSS-FIDVTFTP
tr|A0A0B4ZUG9|Ganoderma      MSD--TALIFRLANDVKK-----LSFDYTPNWGR---GNPSR-FVDNVTFPQ
sp|P14945|LZ8_Ganoderma      MSD--TALIFRLANDVKK-----LSFDYTPNWGR---GNPNN-FIDVTTFPK
                                .: : * ..                : ***** * * * * . : * :***

```

```

tr|A0A0C5BFA5|Postia          VLKDKRYKYRVVSGSTDLGIREPYDVQADGSQRINFLEYNQGYGIRDNTNTVTVYLVDPET
sp|P80412|Flammulina         VLTDKKYSYRVVWNGSDLGVESNFAVTPSGGQTINFLQYNKGYGVADTKTIQVFWIPDT
genemark-tig00001168|Ganoderma VLTDKAYKYRVVWAGRDLGVRPSYEVASDGSQKINFLEYNHNGYGIADTNTIQVFWDPST
tr|A0A0K1L6W2|Chroogomphus   VLTDKAYTCRVVWVSGRDLGVRPSYEVASDGSQKINFLEYNHNGYGIADTNTIQVYVDPST
tr|G5CJT8|Ganoderma          VLTDKAYTYRVVWVSGRDLGVRPSYEVASDGSQKINFLEYNHNGYGIADTNTIQVYVDPST
tr|E7FH75|Ganoderma          VLTDKAYTYRVVWVSGKDLGVRPSYAVESDGSQKINFLEYNHNGYGIADTNTIQVYVIDPD
tr|A0A0B4ZUG9|Ganoderma      VLADKAYTYRVVWVSGRDLGVRPSYAVGSDGSQKVNFLQYNSYGIADTNTIQVVIDPDT
sp|P14945|LZ8_Ganoderma      VLTDKAYTYRVAVSGRNLGVKPSYAVESDGSQKVNFLQYNSYGIADTNTIQVFWDPDT
** * * * . ** . . : **:   : * ..*.* :***:~::~*: **:* :~::~ * :

```

```

tr|A0A0C5BFA5|Postia          AT--EYKLTQSQ-----
sp|P80412|Flammulina         GNSEEIYIAENKKT-----
genemark-tig00001168|Ganoderma GD--DFIIAQWK-----
tr|A0A0K1L6W2|Chroogomphus   GD--DFIIAQWN-----
tr|G5CJT8|Ganoderma          GD--DFIIAQWN-----
tr|E7FH75|Ganoderma          GN--NFIVAQWNYLEQKLISEEDLNSAVDHHHHHH
tr|A0A0B4ZUG9|Ganoderma      GA--DFIIAQWN-----
sp|P14945|LZ8_Ganoderma      NN--DFIIAQWN-----
                                :: :: :

```

**E. Propiedades de la secuencia putativa FIP encontrada en el genoma ensamblado de *G. australe*.**

PEPSTATS of FGENESH from 1 to 113

Molecular weight = 12732.37                      Residues = 113

Average Residue Weight = 112.676    Charge = 0.0

Isoelectric Point = 6.5237

A280 Molar Extinction Coefficients = 19940 (reduced)                      19940

(cystine bridges)

A280 Extinction Coefficients 1mg/ml = 1.566 (reduced)                      1.566

(cystine bridges)

Improbability of expression in inclusion bodies = 0.635

Residue	Number	Mole%	DayhoffStat
A = Ala	7	6.195	0.720
B = Asx	0	0.000	0.000
C = Cys	0	0.000	0.000
D = Asp	9	7.965	1.448
E = Glu	4	3.540	0.590
F = Phe	8	7.080	1.967
G = Gly	8	7.080	0.843
H = His	2	1.770	0.885
I = Ile	7	6.195	1.377
J = ---	0	0.000	0.000
K = Lys	7	6.195	0.939
L = Leu	6	5.310	0.718
M = Met	1	0.885	0.521
N = Asn	4	3.540	0.823
O = ---	0	0.000	0.000
P = Pro	5	4.425	0.851
Q = Gln	3	2.655	0.681
R = Arg	5	4.425	0.903
S = Ser	8	7.080	1.011

T = Thr	10	8.850	1.451
U = ---	0	0.000	0.000
V = Val	11	9.735	1.475
W = Trp	2	1.770	1.361
X = Xaa	0	0.000	0.000
Y = Tyr	6	5.310	1.562
Z = Glx	0	0.000	0.000

Property	Residues	Number	Mole%
Tiny	(A+C+G+S+T)	33	29.204
Small	(A+B+C+D+G+N+P+S+T+V)	62	54.867
Aliphatic	(A+I+L+V)	31	27.434
Aromatic	(F+H+W+Y)	18	15.929
Non-polar	(A+C+F+G+I+L+M+P+V+W+Y)	61	53.982
Polar	(D+E+H+K+N+Q+R+S+T+Z)	52	46.018
Charged	(B+D+E+H+K+R+Z)	27	23.894
Basic	(H+K+R)	14	12.389
Acidic	(B+D+E+Z)	13	11.504



## F. Búsqueda de homología de la secuencia putativa FIP dentro del repositorio de Pfam



### Sequence search results

[Show](#) the detailed description of this results page.

We found **1** Pfam-A match to your search sequence (**all** significant)

Fve

[Show](#) the search options and sequence that you submitted.

[Return](#) to the search form to look for Pfam domains on a new sequence.

### Significant Pfam-A Matches

[Show](#) or [hide](#) all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Fve	Fungal immunomodulatory protein Fve	Domain	n/a	4	112	6	112	3	111	111	195.6	1.6e-58	n/a	<a href="#">Hide</a>
#HMM	talifrlaylvkkldfdytpnwgrrgtppssyidltlfpkvltdkky yrvvvsgrdlgykssyavesdgsqkvnfleyngygvadtktiqvfvvipdtnnseyiiaew													
#MATCH	tal+frlay+vkk++fdytp wgrgtppss+id+ltfpkvltdk+y yrvvv+grdlgv++sy v sdgsqk+nfley +g+g+adt+tiqvfvv+p +++ ++iia+w													
#PP	g*****9998..*****													
#SEQ	TALTFRLAYEVKKISFDYTPHWGRGTTPSSFIDNLTFPKVLTDKAYKRVVVVAGRDLGVRPSYEVASDGSQKINFLEYHNGFGIADTNTIQVFVVDPTSGL--DFIIAQL													

Comments or questions on the site? Send a mail to [pfam-help@ebi.ac.uk](mailto:pfam-help@ebi.ac.uk).  
European Molecular Biology Laboratory



## G. Resultados predicción secuencia putativa de *G. australe* en I-Tasser

### Submitted Sequence in [FASTA format](#)

```
>seq
MSETVTALTFRLAYEVKKISFDYTPHWGRGTPSSFIDNLTFPKVLTDKAYKYRVVWAGRD
LGVRPSYEVASDGSQKINFLEYHNGFGIADTNTIQVFVVDPTSGDDFIIAQWK
```

### Predicted Secondary Structure

	20	40	60	80	100
Sequence	MSETVTALTFRLAYEVKKISFDYTPHWGRGTPSSFIDNLTFPKVLTDKAYKYRVVWAGRD	LGVRPSYEVASDGSQKINFLEYHNGFGIADTNTIQVFVVDPTSGDDFIIAQWK			
Prediction	CCCHHH	HH	SS	SS	SS
Conf. Score	96204645567865431110224777678883111013323501127760799998155355465534627887257679951774402476689999868999703887429				
	H:Helix; S:Strand; C:Coil				

### Top 10 homologous GO templates in PDB

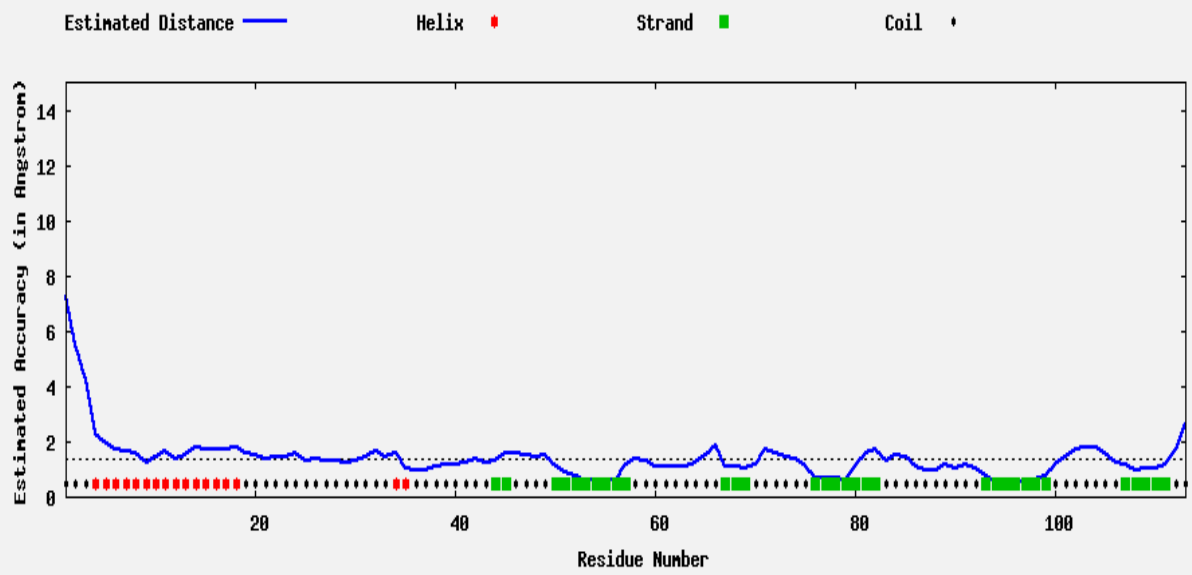
Rank	Cscore <sup>GO</sup>	TM-score	RMSD <sup>a</sup>	IDEN <sup>a</sup>	Cov	PDB Hit	Associated GO Terms
1	0.46	0.9587	0.50	0.74	0.97	<a href="#">3kcwA</a>	<a href="#">GO:0002682</a> <a href="#">GO:0030246</a>
2	0.45	0.9470	0.69	0.76	0.97	<a href="#">3f3hA</a>	<a href="#">GO:0002682</a> <a href="#">GO:0030246</a>
3	0.44	0.9106	1.17	0.68	0.97	<a href="#">1osyA</a>	<a href="#">GO:0005529</a> <a href="#">GO:0002682</a> <a href="#">GO:0030246</a>
4	0.33	0.6296	3.13	0.12	0.85	<a href="#">1t6bX</a>	<a href="#">GO:0009405</a> <a href="#">GO:0005576</a> <a href="#">GO:0005515</a> <a href="#">GO:0005615</a> <a href="#">GO:0046872</a>
5	0.32	0.6240	2.97	0.12	0.84	<a href="#">1tzoA</a>	<a href="#">GO:0009405</a> <a href="#">GO:0005615</a> <a href="#">GO:0005576</a> <a href="#">GO:0005515</a> <a href="#">GO:0046872</a>
6	0.31	0.6195	3.04	0.12	0.82	<a href="#">3etbJ</a>	<a href="#">GO:0009405</a> <a href="#">GO:0005576</a> <a href="#">GO:0005515</a> <a href="#">GO:0046872</a> <a href="#">GO:0005615</a>
7	0.30	0.5948	3.98	0.07	0.94	<a href="#">1moeA</a>	<a href="#">GO:0005515</a>
8	0.30	0.5762	3.48	0.17	0.83	<a href="#">1bg1A</a>	<a href="#">GO:0046872</a> <a href="#">GO:0016787</a> <a href="#">GO:0005990</a> <a href="#">GO:0016798</a> <a href="#">GO:0005515</a> <a href="#">GO:0004565</a> <a href="#">GO:0008152</a> <a href="#">GO:0003824</a> <a href="#">GO:0004553</a> <a href="#">GO:0005975</a> <a href="#">GO:0009341</a> <a href="#">GO:0030246</a> <a href="#">GO:0043169</a>
9	0.29	0.5860	3.01	0.04	0.81	<a href="#">2wngA</a>	<a href="#">GO:0005515</a>
10	0.29	0.5069	4.20	0.06	0.82	<a href="#">2d9cA</a>	<a href="#">GO:0005515</a>

### Consensus prediction of GO terms

Molecular Function	<a href="#">GO:0005515</a>	<a href="#">GO:0046872</a>	<a href="#">GO:0005529</a>
GO-Score	0.54	0.54	0.44
Biological Process	<a href="#">GO:0002682</a>	<a href="#">GO:0009405</a>	
GO-Score	0.83	0.54	
Cellular Component	<a href="#">GO:0005615</a>		
GO-Score	0.54		

## Generated 3D models

[Download the estimated local accuracy of models](#)



- [Download Model 1](#)
- C-score=1.44 ([Read more about C-score](#))
- Estimated TM-score = 0.91±0.06
- Estimated RMSD = 1.5±1.4Å

## H. Resultados predicción de las secuencias FIPs de *G. applanatum* y *C. rutilus* en I-Tasser

- Resultados para la FIP de *G. applanatum*



### Submitted Sequence in FASTA format

```
>seq
MSETVTALTFRLAWEVKKISFDYTPNWGRGSPSSYIDNLTFPKVLTDKAYTYRVVWVSGRD
LGVRPSEYEVASDGSQKINFLYHNGYGIADTNTIQVYVVDVDPSTGDDFIIAQWN
```

### Predicted Secondary Structure

	20	40	60	80	100
Sequence	MSETVTALTFRLAWEVKKISFDYTPNWGRGSPSSYIDNLTFPKVLTDKAYTYRVVWVSGRDLGVRPSEYEVASDGSQKINFLYHNGYGIADTNTIQVYVVDVDPSTGDDFIIAQWN				
Prediction	CCCCHHHHHHHHHHHHHHCCCCCCCCCCCCCHHCCCCCCCCSSCCCCSSSSSSSSCCCCCCCCSSSSCCCCSSSSSSSSCCCCCCCCSSSSSSSSCCCCCCCCSSSSSSCC				
Conf. Score	96205655567865432110224777678883200013323501127760799997065355465534627887258679961774402476689999868999714887439				
	H:Helix; S:Strand; C:Coil				

### Predicted Solvent Accessibility

	20	40	60	80	100
Sequence	MSETVTALTFRLAWEVKKISFDYTPNWGRGSPSSYIDNLTFPKVLTDKAYTYRVVWVSGRDLGVRPSEYEVASDGSQKINFLYHNGYGIADTNTIQVYVVDVDPSTGDDFIIAQWN				
Prediction	85442320103013324414252334234343442133131342135441323112425523234524246734441320423533223434303010030644441123428				
	Values range from 0 (buried residue) to 9 (highly exposed residue)				

### Top 10 Identified structural analogs in PDB

Click to view	Rank	PDB Hit	TM-score	RMSD <sup>a</sup>	IDEN <sup>a</sup>	Cov	Alignment
	1	<a href="#">3f3hA</a>	0.964	0.40	0.791	0.974	<a href="#">Download</a>
	2	<a href="#">3kcwA</a>	0.943	0.73	0.791	0.974	<a href="#">Download</a>
	3	<a href="#">1osyB</a>	0.938	0.84	0.673	0.974	<a href="#">Download</a>
	4	<a href="#">1pqbH</a>	0.638	2.91	0.121	0.841	<a href="#">Download</a>
	5	<a href="#">3etbM</a>	0.631	2.71	0.126	0.814	<a href="#">Download</a>
	6	<a href="#">1tznL</a>	0.631	2.67	0.123	0.814	<a href="#">Download</a>
	7	<a href="#">1acc_3</a>	0.621	2.68	0.128	0.796	<a href="#">Download</a>
	8	<a href="#">6etzA</a>	0.602	3.35	0.135	0.858	<a href="#">Download</a>
	9	<a href="#">3jvfC</a>	0.597	3.90	0.057	0.911	<a href="#">Download</a>

- Resultados para la FIP de *C. rutilus*



