UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE FARMACIA

TESIS DOCTORAL

Secuenciación metagenómica y nuevos procedimientos bioinformáticos para entender la evolución de hongos liquenizados

Metagenome sequencing with new bioinformatic approaches to understand the evolution of lichen forming fungi

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

David Pizarro Martínez

Directores

Ana María Crespo de las Casas
Pradeep Kumar Divakar

Madrid

© David Pizarro Martínez, 2019

Universidad Complutense de Madrid
Facultad de Farmacia

Departamento de Farmacología, Farmacognosia y Botánica

SECUENCIACIÓN METAGENÓMICA Y NUEVOS PROCEDIMIENTOS BIOINFORMÁTICOS PARA ENTENDER LA EVOLUCIÓN DE HONGOS LIQUENIZADOS

METAGENOME SEQUENCING WITH NEW BIOINFORMATIC APPROACHES TO UNDERSTAND THE EVOLUTION OF LICHEN FORMING FUNGI

MEMORIA PARA OPTAR AL GRADO DE DOCTOR PRESENTADA POR:

DAVID PIZARRO MARTÍNEZ

BAJO LA DIRECCIÓN DE LOS DOCTORES

ANA MARÍA CRESPO DE LAS CASAS y

PRADEEP KUMAR DIVAKAR

Madrid, 2019

Universidad Complutense de Madrid
Facultad de Farmacia

Departamento de Farmacología, Farmacognosia y Botánica

SECUENCIACIÓN METAGENÓMICA Y NUEVOS PROCEDIMIENTOS BIOINFORMÁTICOS PARA ENTENDER LA EVOLUCIÓN DE HONGOS LIQUENIZADOS

METAGENOME SEQUENCING WITH NEW BIOINFORMATIC APPROACHES TO UNDERSTAND THE EVOLUTION OF LICHEN FORMING FUNGI

MEMORIA PARA OPTAR AL GRADO DE DOCTOR PRESENTADA POR:

DAVID PIZARRO MARTÍNEZ

BAJO LA DIRECCIÓN DE LOS DOCTORES

ANA MARÍA CRESPO DE LAS CASAS y

PRADEEP KUMAR DIVAKAR

Madrid, 2019

UNIVERSIDAD COMPLUTENSE MADRID

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA PARA OBTENER EL TÍTULO DE DOCTOR

D./Dña. David Pizarro Martínez,
estudiante en el Programa de Doctorado Farmacia,
de la Facultad de Farmacia de la Universidad Complutense de Madrid, como autor/a de la tesis presentada para la obtención del título de Doctor y titulada:

Secuenciación metagenómica y nuevos procedimientos bioinformáticos para entender la evolución de hongos liquenizados

y dirigida por: Ana Mª Crespo de las Casas y Pradeep K. Divakar

DECLARO QUE:

La tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la Ley de Propiedad Intelectual (R.D. legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, modificado por la Ley 2/2019, de 1 de marzo, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita.

Del mismo modo, asumo frente a la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad del contenido de la tesis presentada de conformidad con el ordenamiento jurídico vigente.

En Madrid, a 10 de septiembre de 2019

Fdo.: 

Esta DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD debe ser insertada en la primera página de la tesis presentada para la obtención del título de Doctor.

Dra. Ana María Crespo de las Casas, Profesora Emérita del Departamento de Farmacología, Farmacognosia y Botánica de la Facultad de Farmacia de la Universidad Complutense de Madrid, y Dr. Pradeep Kumar Divakar, Profesor Contratado Doctor del mismo departamento,

CERTIFICAN:

Que el presente trabajo de investigación titulado "Genómica y nuevos procedimientos bioinformáticos para entender la evolución de hongos liquenizados" constituye la memoria de Tesis Doctoral que presenta D. David Pizarro Martínez para optar al grado de Doctor con Mención Europea, y que ha sido realizado en el Departamento de Farmacología, Farmacognosia y Botánica de la Facultad de Farmacia de la Universidad Complutense de Madrid, bajo nuestra dirección.

Asimismo, concluido el trabajo experimental, bibliográfico y de redacción, autorizamos su presentación y defensa en la Universidad Complutense de Madrid, para que sea juzgado por el Tribunal correspondiente. Para que conste a los efectos oportunos, firmamos el presente certificado en Madrid a 10 de Septiembre de 2019.

Dra. Ana María Crespo de las Casas Dr. Pradeep Kumar Divakar

Este trabajo de Tesis Doctoral ha sido realizado en el Departamento de Farmacología, Farmacognosia y Botánica de la Facultad de Farmacia de la Universidad Complutense de Madrid gracias a una beca pre-doctoral de Formación de Personal Investigador (FPI) del Ministerio de Economía (FPI CODE) concedida a David Pizarro Martínez.

Durante el desarrollo de la Tesis Doctoral se ha realizado dos estancias breves, ambas financiadas por el Ministerio de Economía e incluidas en el programa formativo de Formación de Personal Investigador (FPI). Una estancia de cuatro meses fue realizada en el año 2016 en la Universidad Göethe (Frankfurt, Alemania), en el grupo de investigación de Bioinformática aplicada dirigido por el Dr. Ingo Ebersberger. La segunda estancia tuvo una duración de dos meses y se realizó durante el año 2017 en el Field Museum de Chicago (EEUU) bajo la tutela del Dr. H Thorsten Lumbsch director del Centro de Investigación Integrativa.

TABLA DE CONTENIDOS

Dedico esta tesis a mis padres y a mis hermanas por su continuo e incondicional apoyo.
Sin sus palabras alentadoras y sus consejos, no estaría donde estoy hoy.

AGRADECIMIENTOS

Este viaje llegó a su fin gracias a muchas personas que colaboraron con paciencia, generosidad, entendimiento, comprensión y tolerancia.

En primer lugar, agradecer especialmente a mis supervisores Ana Mª Crespo, por la generosidad de compartir su infinito conocimiento, y Pradeep K. Divakar por su incesante aporte de ideas y fomentar colaboraciones científicas. Gracias, es un verdadero orgullo haber podido trabajar y aprender de vosotros. Igualmente agradecer sinceramente al profesor Thorsten Lumbsch por gran cantidad de razones, pero principalmente, por sus consejos y por servir de guia en esta investigación. El ha sido como un tercer director de tesis.

Doy también las gracias a Francesco dal Grande de quien tanto he aprendido, su ayuda, generosidad y buen hacer han sido de gran importancia. A Imke Schmitt, Steve Leavitt, Felix Grewe, JP, Mats Wedin, Paul Dyer, Daniele Armaleo, Elisa Lagostina, Todd Widhelm, por haber cedido generosamente gran cantidad de datos genómicos y permitirme realizar muchos de los análisis en sus clústeres de computación, sin los cuales esta tesis no hubiese sido posible.

Mi más sincero agradecimiento al profesor Ingo Ebersberger por recibirme en el Departamento de Bioinformática Aplicada de la Universidad Goethe y por enseñarme una gran parte de las técnicas empleadas en esta tesis. Doy las gracias a todas las personas de este grupo de investigación, en especial a Bastian y Andreas por sus charlas de genómica de líquenes.

Agradezco también a todos los profesores del extinto Departamento de "Biología Vegetal II", especialmente a Ruth del Prado, Victor Rico, Leopoldo Sancho, Paloma Cubas, Carmen Navarro, Ana Pintado por su apoyo en temas

científicos siempre que los he necesitado, y su generosidad en las colaboraciones docentes. También agradecer al personal PAS del departamento, Charo, Pilar, Pepe Pizarro y Manolo por brindarme su ayuda durante estos años.

También agradezco a todos los compañeros del Departamento con quienes he compartido gran parte de mi tiempo y quienes siempre han estado dispuestos a echar una mano. Gracias a Alors y Tino por vuestro apoyo moral en estos largos años, a Carlos y JC por compartir vuestro saber. Gracias a Johana, Zu, Clara, Guille, Bea, Rober, Alberto, Elena, Pepe, Nuria, Rosina, Manuela, Soumi, Giti.

Doy las gracias a Fede, quien siempre ha estado a mi lado, por su amor, por sus palabras de ánimo y por su comprensión, sin tu infinito apoyo esto no hubiera sido posible.

A mis amigos, quienes han estado ahí con palabras amables, siempre dando ánimo cuando más necesario era y por ayudar a evadirme con sus incesantes planes, sus músicas y sus fiestas de guardar.

Finalmente, dar las gracias al Ministerio competente de la Administración del Estado Español por haberme concedido una FPI con la cual he tenido la oportunidad de realizar esta tesis.

RESUMEN

SUMMARY

RESUMEN

Los líquenes son un grupo ampliamente diverso resultado de la simbiosis entre un hongo, algas y otros organismos asociados. Este conjunto de organismos ha colonizado la mayor parte de los ambientes terrestres y se han presentado, además, como fuente natural de nuevos compuestos con interés farmacológico. Hasta la fecha pocos estudios se han llevado a cabo usando herramientas genómicas, probablemente, debido a la dificultad de obtener los cultivos del micobionte por su lento crecimiento. En esta tesis se presenta una alternativa usando secuenciación de metagenomas del holobionte completo con la finalidad de recuperar el genoma del micobionte. Las secuencias genómicas de cada una de las especies de Parmeliaceos incluidas han sido evaluadas mediante la comparación del tamaño, la integridad de genomas y el número de genes con genomas de líquenes obtenidos a partir de cultivos aposimbióticos. Posteriormente, los genomas han sido utilizados para contestar diferentes preguntas relacionadas con la biología evolutiva de estos hongos. En primer lugar, se ha llevado a cabo un estudio filogenómico usando un gran conjunto de datos de genes de una sola copia, donde se han resuelto muchas de las relaciones evolutivas internas de los seis clados principales de la familia Parmeliaceae. Dado el poco consenso que había entorno a la reproducción sexual en estos organismos simbióticos, mediante el análisis comparativo de los genomas disponibles de la clase Lecanoromycetes se ha detectado el heterotalismo como estrategia dominante en la reproducción sexual. La estructura génica del locus MAT se ha confirmado similar a la de otros hongos pertenecientes a la subdivisión Pezizomycotina y se han identificado varios genes auxiliares que podrían ser específicos de la clase Lecanoromycetes. Por último, se aborda la diversidad de clústeres de genes biosintéticos implicados en la producción de metabolitos secundarios en Lecanoromycetes. En concreto, se ha estudiado la familia de proteínas de policétidos sintasas no reductoras (NR-PKS), la cuales están implicadas en la producción de gran cantidad de sustancias liquénicas como el ácido úsnico o las melaninas fúngicas. El análisis filogenético de la NR-PKS del ácido úsnico desveló que solamente las especies productoras de este compuesto albergan dicho clúster en el genoma, incluso en especies relacionadas evolutivamente, algo que solo puede ser explicado mediante repetidas pérdidas de genes a lo largo de la evolución de estos organismos.

SUMMARY

Lichens are a widely diverse group produced by the symbiosis between a fungus, algae and other associated organisms. This set of organisms has colonized most parts of terrestrial environments and has also been presented as a natural source of new compounds with pharmacological interest. To date, few studies have been carried out using genomic approach, probably due to the difficulty of obtaining mycobiont cultures because of their slow growth. In this thesis is presented an alternative option using metagenome sequencing of the complete holobiont in order to recover the genome of the mycobiont. The binned genomes of each Parmeliaceae specie has been evaluated by comparing size, genome completeness and number of genes with lichen genomes obtained from aposymbiotic cultures. Subsequently, genomes have been used to answer different questions related to the evolutionary biology of these fungi. On one hand, a phylogenomic study has been carried out using a large set of single-copy gene, where monophyly of the six major clade of Parmeliaceae was confirmed and most backbone relationships in the topology have been resolved. On the other hand, given the lack of scientific consensus regarding sexual reproduction in these symbiotic organisms, through comparative analysis of the available genomes, heterothallism has been detected as a dominant strategy in sexual reproduction among the Lecanoromycetes class. The gene synteny of the MAT locus has been confirmed to be similar to other fungi and we have detected that several auxiliary mat genes could be specific to the Lecanoromycetes class. Finally, it has been carried out an in-silico analysis in order to uncover the diversity of biosynthetic genes cluster involved in the production of secondary metabolites in Lecanoromycetes. Specifically, the non-reducing polyketide synthase (NR-PKS) proteins have been studied because they are involved in the production of a large amount of lichen substances such as usnic acid or fungal melanin. The phylogenetic analysis of the NR-PKS revealed that only the species usnic acid-producers harbor the biosynthetic cluster of this compound in the genome, even in evolutionarily related species, something that can only be explained by repeated losses throughout the evolution of these organisms.

2.INTRODUCCIÓN GENERAL

2.GENERAL INTRODUCTION

2

2.1 LOS LÍQUENES (Lecanoromycetes)

El término simbiosis viene de la palabra griega "symbio" que significa "vivir junto" y fue acuñado por el micólogo alemán Heinrich Anton de Bary en 1879 refiriéndose a «la vida en conjunción de dos organismos disímiles, normalmente en íntima asociación, y por lo general con efectos benéficos para al menos uno de ellos». Este término, por lo tanto, engloba diferentes relaciones biológicas donde ambos componentes (apobiontes) de la simbiosis obtienen un beneficio de dicha relación. Las relaciones simbióticas son una de las más importantes adquisiciones en la biología evolutiva desde la formación la célula eucariota (Margullis 1967; Cooper 2000) hasta la colonización terrestre de las plantas superiores (Wang et al. 2009). Estas relaciones están ampliamente extendidas y se suelen dar entre organismos evolutivamente distantes (Dimijian 2000), como por ejemplo los hongos micorrizógenos y las plantas y la muy singular simbiosis liquénica.

La definición más ampliamente aceptada por la comunidad científica de liquen es "una asociación mutualista estable y obligada ecológicamente, entre un componente fúngico (el micobionte) "exhabitante" y una población "inhabitante" de algas unicelulares o cianobacterias localizadas extracelularmente (el fotobionte)" (Honnegger & Hawkswoth 1994). Esta asociación forma una estructura estable y compleja a nivel morfológico, la cual se ha considerado desde un punto de vista ecológico como un pequeño ecosistema. Esta consideración está basada en la capacidad del talo (cuerpo) liquénico para integrar los dos principales procesos metabólicos (autotrofismo y heterotrofismo), así como la capacidad de albergar una gran diversidad de organismos que viven asociados al mismo talo: bacterias (Grube et al. 2009), hongos suplementarios liquenícolas (Hawksworth 2003) o endosimbiontes (Spiribille et al. 2016). El resultado de la interacción entre todos los componentes es conocido como "holobionte" y presenta una morfología, una ecología y una fisiología específica y propia.

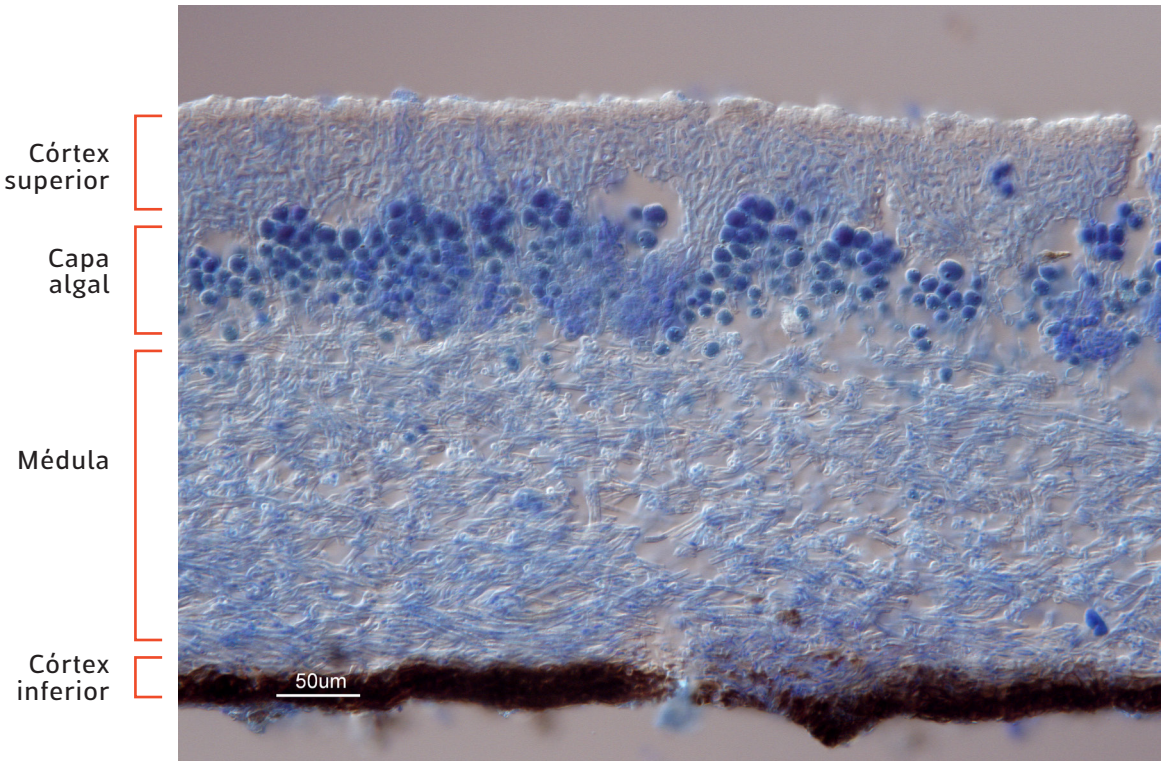


Fig I1 - Corte transversal del talo liquénico de *Physconia thorstenii.*

La liquenización representa uno de los estilos de vida más frecuente dentro de los hongos con más de 19.400 especies aceptadas, comprendiendo el 20% de todos los hongos conocidos (Hawksworth 2001; Nash 2008). Este proceso ha ocurrido de manera independiente en diferentes linajes del Reino Fungi por convergencia evolutiva (Gargas et al. 1995), tratándose así de un grupo polifilético, donde la mayor parte de los micobiontes pertenecen a la división Ascomycota, siendo la subclase Lecanoromycetidae la más diversa con más de 10.000 especies descritas subdivididas en ocho órdenes (Kraichak et al. 2018); la práctica totalidad de las especies de la subclase son líquenes.

Líquenes Parmeliaceos y aspectos de su biología evolutiva

La familia Parmeliaceae, perteneciente al orden Lecanorales, es la más amplia de los hongos liquenizados, e incluye alrededor de 2.700 especies (un 10% del total de especies de líquenes conocidas). Se encuentra ampliamente distribuida por todo el planeta, incluyendo ambientes extremos como la Antártida o determinados desiertos del sur de África (Cannon & Kirk 2007). La mayor parte de las especies forman talos con morfologías o bien foliáceas (laminares) con frecuencia provistas de rizinas o rizoides, o bien fruticulosas (microarbustivas) erectas o más comúnmente péndulas; todo el linaje presenta un tipo definido de ascocarpo o cuerpo fructífero que posee una morfología peculiar dentro de los apotecios lecanorinos (Ferencova 2012).

Fig I.2 - Diferentes biotipos de Parmeliaceos. SYSTEMOL

Usnea florida (Uttar Pradesh, India)

Parmelina carporrhizans (Extremadura, España)

Evernia prunastri (Madrid, España)

Parmelina tiliacea (Portugal)

Históricamente la familia Parmeliaceae fue objeto de controversia en relación a los criterios sistemáticos y concepto de género (Nimis 1998; Elix 1993) que surgieron de dos corrientes científicas. Una de las corrientes se basaba en la diferencia de caracteres vegetativos como la morfología o los metabolitos secundarios, y la otra se apoyaba en los caracteres sexuales, que siendo los más consistentes, muchas veces eran inaccesibles por no estar siempre presentes en las muestras y en las especies en estudio. Los resultados aportados por ambas corrientes reflejaban grandes discordancias taxonómicas y nomenclaturales en todas las categorías sistemáticas y particularmente en los géneros. Asimismo, el reconocimiento y delimitación de las especies era muy controvertido debido a que la valoración de los caracteres era discutible y el número de caracteres disponibles resultaba siempre escaso (Crespo et al. 2010; Lumbsch & Leavitt 2011; Crespo et al. 2011).

Con objeto de resolver el debate sistemático planteado en una familia de líquenes tan común y abundante como los parmeliáceas, en cuanto se hacen fácilmente disponibles los caracteres moleculares, esta es una de las familias que primero se estudia (Crespo & Cubero 1999; Mattsson & Wedin 1999). Como consecuencia de la aplicación de esta nueva herramienta se accede al cambio conceptual, paradigmático, que conduce a la Taxonomía Integrada o Integrativa (Henning Steinicke ed. German National Academy Leopoldina, 2014). La nueva metodología, asume que la filogenia molecular proporciona el esquema evolutivo sobre el que apoyar las categorías taxonómicas. Las investigaciones sobre parmeliáceas se convierten en referentes, para el trabajo en sistemática de líquenes y hongos como una rama de facto de la biología evolutiva. Las categorías infra-familiares y señaladamente los rangos de género y especie, se redefinen y se fundamentan sobre los criterios evolutivos y los caracteres fenotípicos se estudian y se revisan sobre las pautas básicas de cada linaje "natural". Las herramientas moleculares y las metodologías de estadísticas de análisis de datos, que van mejorándose continuamente, facilitan el establecimiento de intervalos numéricos en las categorías taxonómicas (Del Prado et al. 2011; Del Prado et al. 2013; Divakar et al. 2019) y las dataciones paleohistóricas de los linajes (Amo et al. 2012).

Así, muchas de las hipótesis tradicionalmente aceptadas fueron verificadas, aunque a la vez, se generó un gran conflicto con muchas de las hipótesis basadas en los caracteres fenotípicos anteriormente mencionados en particular con los caracteres metabólicos (Feuerer & Hawksworth 2007; Lumbsch & Huhndorf 2010).

Desde las primeras aproximaciones filogenéticas con escasos especímenes y uno o dos marcadores moleculares, se va mostrando que la familia es un linaje monofilético (Crespo & Cubero 1999; Mattsson & Wedin 1999; Blanco et al. 2004). En el trabajo de Crespo et al. 2007, los autores, sin discutir las relaciones filogenéticas de la familia como tal, proponen una sistemática interna de la familia basada en 4 genes marcadores; se incluye a su vez un gran número de especies de diferentes tipos morfológicos (biotipos) y distintas procedencias biogeográficas y ecológicas. Los resultados ponen de manifiesto la monofilia de los principales grupos morfológicos (alectorioide, cetrarioide, hipogimnoide, letarioide, parmelioide y psiloparmelioide) que se consolidan como verdaderos linajes filogenéticos. Simultáneamente, mediante el uso de varios marcadores se establecieron que los enfoques basados en caracteres fenotípcos subestimaban enormemente el número de especies existente, dando lugar al concepto de especies crípticas (Crespo et al. 2011).

En posteriores estudios se realizó un gran esfuerzo para establecer las relaciones filogenéticas no resueltas de esta familia mediante el uso de varios marcadores moleculares dando como resultado filogenias de Parmelioides bastante consistentes (Crespo et al. 2010; Divakar et al. 2012). Sin embargo, a pesar de todos estos esfuerzos numerosas relaciones evolutivas de esta familia no quedaron resueltas, ya que cierto número nodos internos carecían de suficiente apoyo (Divakar et al. 2015).

De forma experimental se aprecia que el uso de un número limitado de marcadores moleculares puede suponer un problema cuando se pretende establecer relaciones evolutivas complejas (Divakar et al. 2015). La reconstrucción de las relaciones evolutivas de linajes que han experimentado una acelerada diversificación, como es el caso de la familia Parmeliaceae, puede ser difícil debido a que los rápidos eventos de especiación no permiten una acumulación significativa de señal filogenética (Townsend 2007). Además, el impacto de las duplicaciones de genes, la transferencia horizontal, el reparto incompleto de linajes (incomplete lineages sorting) o sesgos en la selección de los marcadores filogenéticos pueden dar lugar a resultados discrepantes (Katz & Grant 2015; Burki et al. 2016; Heiss et al. 2018). Afortunadamente, y gracias a los avances mencionados en las herramientas de producción de datos genómicos de los últimos años se ha hecho posible una rápida

acumulación de genomas de gran cantidad de organismos. Esto, junto con la mejora de los métodos computacionales, ha facilitado desarrollar nuevos procedimientos que han permitido generar filogenias mucho más consistentes (Lemmon & Lemmon 2013).

Más allá que los marcadores genéticos de primera generación, los datos genómicos han permitido tener una nueva visión, más sólida y precisa, de las relaciones evolutivas gracias a la capacidad de obtener cientos de marcadores independientes distribuidos a lo largo de los genomas (Philippe et al. 2005; Meuseman et al. 2010; Misof et al. 2014; Spatafora et al. 2017). Los genes de una sola copia son considerados útiles y adecuados para inferir relaciones filogenéticas previamente no resueltas en diversos linajes de Eucariotas (Shen et al. 2013; Zeng et al. 2014; Ren et al. 2016). Algunas características de estos genes, como la singularidad y la alta conservación de secuencias a través de los distintos linajes, permiten que puedan ser amplificados y secuenciados fácilmente (Levin et al. 2009; Salas-Leiva et al. 2014). La secuenciación de genomas completos ha facilitado la identificación de estos genes de una sola copia de forma masiva y han sido ampliamente usados como marcadores filogenéticos para reconstruir relaciones evolutivas complejas en hongos, especialmente en Ascomycota (Liu et al. 2009; Medina et al. 2011; Ebersberger et al. 2012; Choi & Kim 2017; Menardo et al. 2017; Teixeira et al. 2017).

Aun así, y pesar de estos avances, el número de datos y estudios genómicos en líquenes es pobre (Leavitt et al. 2016; Grewe et al. 2017), si se compara con la progresión del conocimiento filogenético y evolutivo en otros grandes grupos de organismos, incluyendo otros hongos. Esta situación de relativo retraso, es probablemente debida a la dificultad de obtener cultivos aposimbióticos a partir de los cuales secuenciar los genomas completos de los hongos asociados en tan compleja e interesante simbiosis. La secuenciación metagenómica se ha establecido como una herramienta útil para la obtención de secuencias de genomas completos de calidad en organismos mixtos y simbiontes (revisado en Quince et al. 2017). Gracias a esta técnica es posible secuenciar todos los organismos que componen el talo liquénico y posteriormente separar las secuencias pertenecientes a cada uno de estos organismos mediante el empleo de herramientas bioinformáticas. Estes procedimeinto ha resultado de enorme utilidad para disponer de gran número de datos genéticos (sin duda el linaje liquénico del que se dispone de mayor información son las

parmeliáceas), lo que ha permitido tener un modelo para profundizar, no solo en la sistemática y filogenia de los ascomicotas sino también en los mecanismos evolutivos, pautas y expresión. La dificultad de cultivo in vitro es un hándicap que ha ralentizado históricamente la investigación de la biología de los co-simbiontes, y desentrañar realmente el particular proceso simbiótico que es la liquenización.

2.2 REPRODUCCIÓN EN LÍQUENES

Los hongos liquenizados muestran los modos reproductivos propios de hongos Ascomycotas: reproducción asexual o vegetativa, que con frecuencia es muy importante, especializada y eficiente, y reproducción sexual. En los líquenes, generalmente, el micobionte es el único que presenta reproducción sexual, quedando el alga restringida a la reproducción asexual (Honegger & Zippler 2007). La reproducción sexual en líquenes lleva implícita la necesidad del micobiointe de encontrar en el medio un fotobionte adecuado para establecer de nuevo la relación simbiótica. Cabe recordar que la simbiosis liquénica se comporta como una simbiosis obligada, ya que el micobionte no es viable en vida libre, no simbiótica. Para solventar esto, los líquenes han desarrollado sistemas de reproducción vegetativa, que resultan en la propagación, relativamente rápida y eficaz, de ambos organismos de forma conjunta (Dal Grande et al. 2012), en detrimento de la variabilidad genética conseguida a través de la reproducción sexual.

Reproducción asexual en líquenes

Los líquenes han desarrollado diferentes estructuras de reproducción vegetativas que les confieren indudables ventajas para la colonización de nuevos medios ya que aseguran la presencia de ambos simbiontes. Estos propágulos vegetativos son importantes y se singularizan en los diferentes

linajes de parmeliáceas, y otras grandes familias, sin que se haya podido establecer los mecanismos genéticos que conducen a su expresión. Tales estructuras son muy distintas en cuanto a morfología y ontogenia, siendo los más comunes los soredios e isidios.

Los soredios son estructuras vegetativas compuestas por hifas que envuelven un escaso número de células algales formando minúsculos glóbulos sin protección cortical. Estos pueden desarrollarse o bien en áreas concretas que pierden la protección cortical formando lo que se denominan soralios, o bien estar distribuidos de forma imprecisa, ocupando amplias áreas de la superficie del talo.

Los isidios, en cambio, presentan una estructura corticada generalmente bien definida, como evaginaciones continuas del talo, visibles individualmente con lupa de mano, y que pueden presentar morfologías muy variadas; la forma evaginada, por lo general quebradiza en condiciones ambientales, hace que estos propágulos se puedan desprender mediante mecanismos físicos (o biológicos: se especula con ácaros o insectos), permitiendo así la dispersión a distancias largas (Dal Grande 2014, b). La morfología y localización de estas estructuras de reproducción vegetativa han sido ampliamente empleadas en la taxonomía y sistemática de este grupo, ya que son caracteres generalmente constantes dentro de los diferentes clados (Printzen 2010). Como se ha dicho, las pautas de expresión de tales caracteres están por dilucidar.

Parmelia saxatilis (isidios)

Parmotrema reticulatum (soredios)

Parmelina quercina (apotecios)

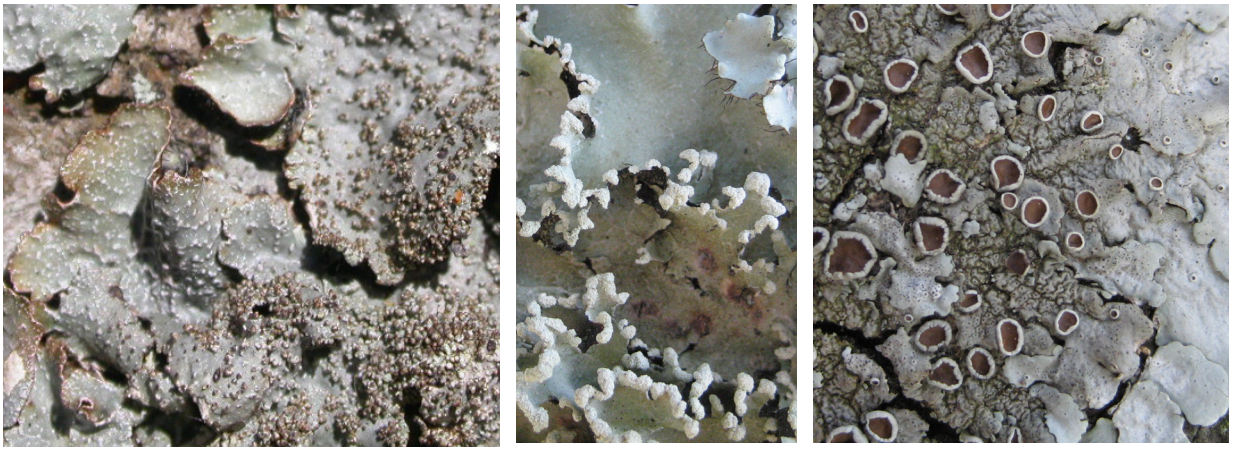


Fig I.3 - Tradicionalmente, los modos y estructuras reproductivas (asexual: isidios/ soredios, y sexual: apotecios) han sido frecuentemente usados para diferenciar especies en Parmeliaceae y líquenes en general. SYSTEMOL.

Reproducción sexual en líquenes

La reproducción sexual en hongos liquenizados implica la formación de ascósporas generadas a partir del proceso meiótico llevado a cabo en el ascoma o cuerpo fructífero de los ascomicotas (Ferencova 2012). Estas estructuras están compuestas por hifas haploides (generalmente periféricas conformando un tejido de protección con la misma estructura que el córtex talino) e hifas ascógenas dicarióticas, de las cuales se originan los ascos y las paráfisis conformando el himenio. Los ascos son órganos esporíferos donde, después de la cariogamia, se origina una célula zigótica diploide a partir de las cuales se originan las ascosporas haploides mediante meiosis. La estructura, anatomía y mecanismo de dehiscencia de los ascos y la morfología, dimensiones y número de las ascósporas, también han sido ampliamente usados como caracteres de gran importancia en la sistemática y taxonomía de líquenes debido a su gran variabilidad y especificidad a diferentes grupos de hongos liquenizados (Printzen 2010, Grube & Wedin 2016 a la luz de los datos proporcionados por la Taxonomía Integrativa).

El proceso de la reproducción sexual de los hongos pertenecientes a la subdivisión Pezizomycotina está controlado por una región genómica denominada locus MAT ("Mating-Type"). Esta región incluye dos alelos, *mat1-1* y *mat1-2*, los cuales debido a su alta divergencia de secuencia son considerados genes idiomorfos (Metzenberg & Glass 1990). Cada uno de estos genes codifica para dos factores de transcripción que son claves para la regulación de otros genes implicados en la reproducción sexual. Concretamente, *mat1-1* codifica para una proteína con un domino (Alfa-Domain) de unión a ADN y *mat1-2* para una proteína con un dominio HMG-Box (High Mobility Group). Dependiendo de la presencia de uno o ambos genes en un mismo genoma podemos encontrar: micelios homotálicos (autofértiles) con ambos genes en el mismo genoma, los cuales pueden completar el ciclo sexual sin necesidad de encontrar a otro individuo compatible, o micelios heterotálicos (no autofértiles), que presentan solo uno de los genes en el mismo genoma y necesitan encontrar un individuo compatible para que se dé la reproducción sexual.

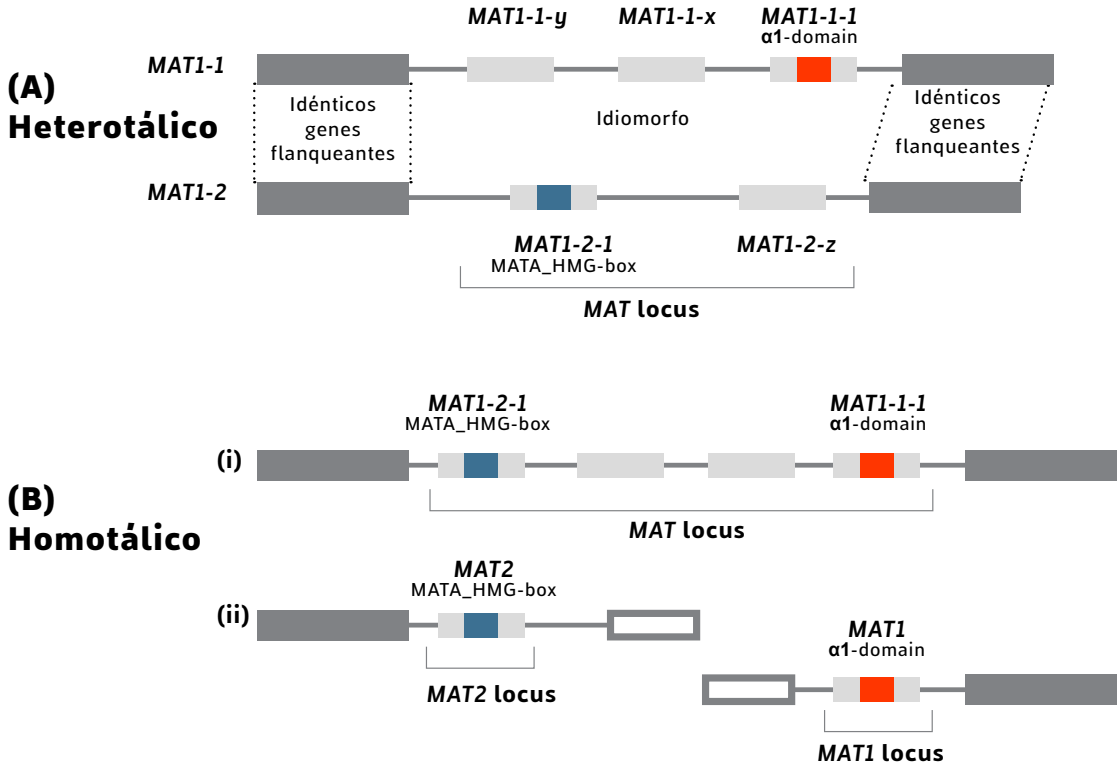


Fig. I.4 - Arquitectura típica del locus MAT en especies heterotálicas (A) y homotálicas (B) de Pezizomycotina. En homotálicos: (i) Misma región genómica (ii) Distinta región genómica. Esquema basado en el original del Capítulo 14 de Hyphal Tip Growth in Filamentous Fungi.

Estudios relativamente recientes (Murtagh et al. 2000) parecían confirmar que los líquenes Lecanoromycetes eran organismos homotálicos, lo cual era consistente debido a que ello podía conferir ventajas en el proceso de la liquenización ya que ofrecía menores riesgos; además este tipo de reproducción se veía muy pertinente en la preservación de líneas genéticas estables y altamente adaptadas a ambientes extremos (Murtagh et al. 2000). La homotalia se infería a partir de estudios donde se emplearon herramientas moleculares disponibles que eran relativamente indirectas. Así, se evaluaba la heterogeneidad de las progenies de las esporas usando técnicas RADPD-PCR (Seymour et al. 2005) o mediante la captura y secuenciación por PCR de regiones parciales del locus MAT (Honeger et al. 2004; Scherrer et al. 2005). Sin embargo, gracias al desarrollo de las técnicas que permiten la secuenciación masiva y también a la reciente disponibilidad de herramientas bioinformáticas, se ha podido abordar el análisis de las regiones genómicas de una forma más directa y por ello, más precisa.

2.3 METABOLISMO SECUNDARIO EN LÍQUENES

Los líquenes son conocidos por producir un alto número de compuestos, la mayoría de ellos únicos (Yamamoto et al. 2015). La familia de los parmeliáceos presenta una gran diversidad de estos metabolitos y muchos de ellos han sido investigados con fines taxonómicos y también prácticos (Fernandez-Moriano et al. 2014; Rankovic et al. 2015). Estos compuestos derivados del metabolismo secundario son conocidos como "sustancias liquénicas", y representan un conjunto de moléculas extraordinariamente heterogéneas de un peso molecular relativamente bajo (Turk et al. 2003). Generalmente, suelen ser biosintetizados por el micobionte (Huneck 1999) y bioacumulados en forma de cristales ácidos en el córtex superior (ej. ácido úsnico, atranorina o melaninas fúngicas) o en la médula (ej. ácido fisódico o ácido protocetrárico). Estos compuestos juegan roles importantes en la biología de los líquenes como antioxidantes, protectores frente a un amplio espectro de microorganismos, protección frente radiación UV, desecación, sin embargo, la función biológica de la gran parte de estas sustancias está aún por elucidar. (Elix 1996; Huneck & Yoshimura 1996; Cocchieto et al. 2002, Legouin et al. 2017, Fernández-Moriano et al. 2016).

Las sustancias liquénicas suelen clasificarse dependiendo de su estructura química: compuestos fenólicos (derivados del orcinol o beta-orcinol), dibenzofuranos (Ej. ácido úsnico), dépsidos (Ej. ácido barbático), depsidonas (Ej. ácido salacínico), depsonas (Ej. ácido picroliquénico), lactonas (Ej. ácido nefrosterínico), quinonas (Ej. parietina) o derivados del ácido pulvínico (Ej ácido vulpínico).

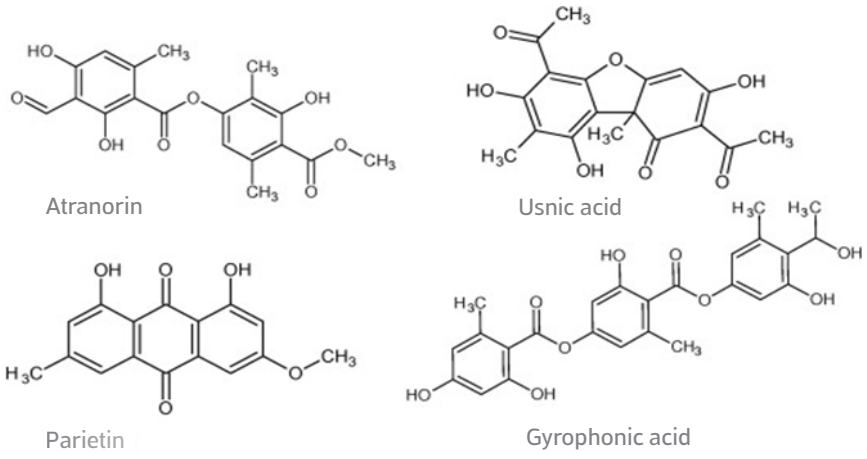


Fig I5 - Ejemplos de compuestos liquénicos.
Imagen extraída de Lichen Secondary Metabolites as DNA-interacting agents.

Se han identificado tres rutas principales para la producción de estos compuestos (Bustie & Grube 2005):

1) Ruta de condensación Acetato-Malonato: de donde se derivan la mayor parte de los compuestos liquénicos, principalmente dépsidos, depsidonas, xantonas, debenzofuranos y antraquinonas.

2) Ruta del Ácido Shikímico: los compuestos sintetizados en esta ruta son principalmente el ácido pulvínico, terpenilquinonas y aminoácidos fenólicos.

3) Ruta del Ácido Mevalónico: principal vía de síntesis de terpenos y otros compuestos isoprenoídicos.

La ruta de condensación de acetato-malonato es la principal ruta en la producción de sustancias liquénicas, la cual se da mediante la polimerización de policétidos derivados del acetil-CoA. Estas reacciones de polimerización son catalizadas por policétido sintasas (PKS), familia de proteínas con multidominios enzimáticos que puede dividirse en varios tipos dependiendo de su estructura primaria y mecanismo catalítico (Cox & Simpson 2009). Además de las PKS, la producción de los metabolitos secundarios requiere otras proteínas con funciones diversas como la regulación de la transcripción de los genes del clúster o la codificación para transportadores encargados de la secreción del metabolito (Brown et al. 2012). Los genes codificantes de estas proteínas están generalmente localizados en una misma región genómica formando clústeres biosintéticos, lo cual se ha descrito como un factor evolutivo importante, ya que facilitaría la expresión coordinada de genes para la producción de los metabolitos (Keller et al. 2005; Osbourn 2010; Khaldi & Wolfe 2011).

Los genomas de hongos filamentosos (Division Ascomycota) contienen decenas de clústeres bioisintéticos implicados en la producción de compuestos secundarios. Como se ha mencionado anteriormente, estos compuestos, generalmente, están implicados en las interacciones con el ambiente (producción de toxinas, fotoprotección, patogénesis, etc.), y, asociados a nichos ecológicos específicos. Esto podría reflejar la existencia de patrones entre la ecología de un organismo y su metabolismo secundario. Sin embargo, desde una perspectiva evolutiva, los metabolitos secundarios son un complejo

rompecabezas debido a que la distribución taxonómica de los mismos, que por lo general, no se ajusta a los modelos evolutivos estrictos. Varios estudios han demostrado que especies pertenecientes a un mismo linaje pueden exhibir un numero distinto de clústeres bioisintéticos, a pesar de compartir una alta homología y sintenia en otras regiones genómicas (Chang et al. 2005; Schumacher et al. 2013; Wiemann et al. 2013; Chiara et al. 2015).

En los últimos años, y gracias a los estudios de genómica comparativa, se ha puesto de manifiesto cómo los procesos de selección han influido en la diversidad de estos clústeres de bioisíntesis. Por ejemplo, estudios como el de Keller et al. 2005 o Fischer et al. 2015 demostraron como la transferencia horizontal de genes (HGT) es un mecanismo por el cual los clústeres biosintéticos se intercambian entre especies distantes (por ejemplo, Hongos-Plantas o Hongos-Bacterias). Aunque la HGT ha sido frecuentemente usada para explicar las distribuciones y patrones evolutivos de metabolitos secundarios, también se han evaluado las implicaciones de los complejos procesos evolutivos de ganancia y perdida de genes (Yu et al. 2004; Patron et al. 2007; Proctor et al. 2008; Bushley & Turgeon 2010; Slot et al. 2011).

A pesar de la importancia de algunos metabolitos secundarios como factores clave en la radiación adaptativa de líquenes (Solhaug & Gauslaa 1996; Divakar et al. 2013; Legouin et al. 2017), así como una importante fuente de compuestos activos, se han llevado a cabo pocos estudios de genómica sobre ellos. El primer clúster biosintético identificado en líquenes fue el del ácido grayánico en *Cladonia grayi* por Armaleo et al. 2011, caracterizado mediante la combinación de diferentes técnicas moleculares y análisis filogenéticos. Posteriormente, Abdel-Hameed et al. 2016 identificaron las dos enzimas principales del clúster de biosíntesis del ácido úsnico y de la 6-Hydroxymelleina en *Cladonia uncialis*. Recientemente, un estudio de genómica comparativa llevado a cabo por Calchera et al. 2019 ha desvelado la gran riqueza y la alta diversidad de clústeres bioisintéticos encontrados en *Evernia prunastri* y *Pseudevernia furfuracea*. Por lo tanto, abordar desde una perspectiva genómica el estudio de las regiones implicadas en la producción de metabolitos secundarios, ayuda a descifrar el potencial que tiene las diferentes especies en la producción de estos compuestos y a su vez, elucidar los mecanismos evolutivos que actúan sobre ellos.

3.OBJETIVOS

3.OBJETIVES

3

3.1 JUSTIFICACIÓN

El desarrollo de las técnicas de secuenciación ha permitido la obtención de genomas de infinidad de organismos no modelo sin la necesidad de llevar a cabo los frágiles y altamente costosos cultivos, posibilitando abordar la biología orgnísmica desde diferentes perspectivas. Los líquenes, y en concreto la familia Parmeliaceae, han sido objeto de estudio debido a su especial relación simbiótica, así como por su capacidad de colonizar nichos ecológicos extremos y ser una nueva fuente de compuestos con interés farmacológico.

En esta tesis se pretende contribuir al conocimiento de la biología evolutiva de líquenes de la clase Lecanoromycetes con especial atención a la familia Parmeliaceae. Por un lado, se pretende resolver las relaciones evolutivas internas aún no resueltas de la familia Parmeliaceae (Divakar et al. 2015) usando datos genómicos y a la vez, proporcionar una metodología para analizar datos metagenómicos de líquenes simbiontes. Esta primera parte se lleva a cabo tomando de referencia estudios como el de Male et al. 2014, donde mediante el uso de metagenomas se resolvieron complejas relaciones filogenéticas. Por otro lado, la importancia de la reproducción sexual en los hongos liquenizados y la posible implicación en su éxito evolutivo da cabida a explorar desde una perspectiva de genómica comparativa las regiones determinantes de las estrategias reproductivas. Por último, se llevará a cabo un estudio acerca de la diversidad de los clústeres de biosínteisis en líquenes, así como análisis de los procesos evolutivos que permiten explicar la distribución del clúster de biosíntesis del ácido úsnico en Lecanoromycetes.

3.2 OBJETIVOS

Objetivos Generales

• Evaluar la secuencición metagenómica como herramienta útil en la recuperación de los genomas de micobionte de líquenes de la familia Parmeliaceae.

• Establecer las relaciones evolutivas aún no resueltas de los líquenes de la familia Parmeliaceae, y evaluar a su vez, el uso de genes unicopia para la resolución de las mismas.

• Mejorar el conocimiento sobre las regiones genómicas determinantes de la reproducción sexual en hongos liquenizados.

• Determinar las bases genéticas e identificar los procesos evolutivos que influyen en la producción de metabolitos secundarios en líquenes.

Objetivos Específicos

• Secuenciar y ensamblar el holobionte de 40 especies representativas de los seis mayores grupos de la familia Parmeliaceae (Capítulo 1).

• Recuperar los genomas del micobionte de cada una de las especies (Capítulo 1).

• Llevar a cabo la predicción y anotación de genes de los genomas del micobionte de 40 especies (Capítulo 1).

• Evaluar la integridad del genoma (Capítulo 2).

• Testar la utilidad de metagenomas para la obtención de genes codificantes de una sola copia (Capítulo 2).

• Establecer relaciones filogenéticas internas de la familia Parmeliaceae mediante una aproximación filogenómica y coalescencia (Capítulo 2).

- Caracterizar y analizar la arquitectura genética del locus MAT en genomas de distintas especies de Lecanoromycetes que representan un amplio rango de formas de crecimiento y tipos de reproducción. (Capítulo 3).

- Determinar el tipo de estrategia reproductiva, homotalismo vs.heterotalismo en Lecanoromycetes. (Capítulo 3).

- Explorar la diversidad de clústeres de genes implicados en la biosíntesis de metabolitos secundarios en diferentes especies de Lecanomycetes (Capítulo 4).

- Determinar las relaciones evolutivas de las PKS no reductoras implicadas en la síntesis de sustancias liquénicas (Capítulo 4).

- Identificar y analizar la arquitectura genética del clúster biosintético del ácido úsnico (Capítulo 4).

- Análisis evolutivo del clúster biosintético del ácido úsnico y determinación de la presencia o ausencia en especies productoras o no de dicho compuesto (Capítulo 4).

4.CAPÍTULO 1

SECUENCIACIÓN, ENSAMBLAJE Y ASIGNACIÓN TAXONÓMICA DE METAGENOMAS DE LÍQUENES

4.CHAPTER 1

SEQUENCING, ASSEMBLY AND TAXONOMIC ASSIGNMENT OF LICHEN METAGENOMES

4

4.1 INTRODUCTION

The advent of DNA sequencing has advanced our understanding on biodiversity research (reviewed in Bik et al. 2012). The analyses of single and multicous datasets in phylogenetic frameworks are frequently used for biodiversity assessment. However, these belong to a small portion of a genome and may bias the accurate species assessment. Therefore a larger portion of genome or whole genome is crucial for a more complete biodiversity assessment. More recently, genomic and metagenomic data are being used for biodiversity research. For example, metagenomic data have been shown to be useful for accurate species assessment (including cryptic) especially in mixed sample or multi-lineage assemblages of organisms ('holobionts') such as metazoan parasites (reviewed in Bass et al. 2015), viral pathogens (reviewed in Bibby, 2013) and other microbial communities (Louca et al. 2016). Furthermore, the genomic and metagenomic data has been and is being frequently used for comparative genomic, biosynthetic gene cluster studies as well as advancing evolutionary biology research (Khaldi et al. 2010; Alföldi and Lindblad-Toh 2013; Quince et al. 2017; Schuelke et al. 2017). The metagenomic approach has been shown to be an alternative method to directly access the uncultured genome for natural products research and for the discovery of novel, bioactive substances in symbiotic organisms (reviewed in Brady et al. 2009).

Lichens represent iconic examples of multi-lineage assemblages of organisms. Lichens form obligate symbiotic association between fungus (exhabitant or host) and microalgae, including cyanobacteria (reviewed in Hawksworth, 2015). In addition to the fungal and photosynthetic partners, a wide variety of non-photosynthetic bacteria, additional fungi as enodlichenic and lichenicolous fungi are regularly found in the lichen symbiosis (reviewed in Hawksworth 2015; Grube and Wedin 2016; Lumbsch and Rikkinen 2017).

The appearance of cryptic species is a common phenomenon in lichen forming fungi and fungi in general (see e.g. Crespo and Lumbsch 2010; Lumbsch and Leavitt 2011; Leavitt et al. 2016; Grube et al. 2017). Cryptic biodiversity is an essential component of biodiversity that will be considerably affected by global climate change in the next decades (Bálint et al. 2011). While the molecular studies based on single and multicolus dataset have revolutionized our understanding of species delimitations in lichen-forming fungi, the application of small portion of genome may obscure the accurate cryptic biodiversity assessment.

Here we sequenced metagenomes of holobiont of 40 lichen species (Parmeliaceae, Ascomycota) to advance different questions related with the biology of lichen-forming fungi as phylogenetic relationship, sexual reproduction and diversity of biosynthetic gene cluster implicated in the production of lichen compounds. The species included, cover a wide range of growth forms, different reproductive strategies, lichen substances production and represent the six major groups of Parmeliaceae family (Divakar et al. 2015; Brodo and Hawksworth 1977). Most of the species included are wide spread in Europe, Asia, Africa, America or Antarctica, and they have been frequently used in bio-monitoring studies especially monitoring forest health (Velmala & Myllys, 2011).

4.2 MATERIALS AND METHODS

DNA extraction and sequencing

Total genomic DNA of 40 specimens of Parmeliaceae species (see Table 1.1.) was extracted from apothecia or thalli devoid of any visible damage, areas with black dot, contamination or dissimilar colors of the entire thallus using the Quick-DNA TM Fungal/Bacterial Miniprep Kit (Zymo Research, Irvine, CA) following the manufacturers' instructions. DNA concentration was calculated using the Qubit dsDNA BR Assay kit (Thermo FisherScientific, San Diego, CA),

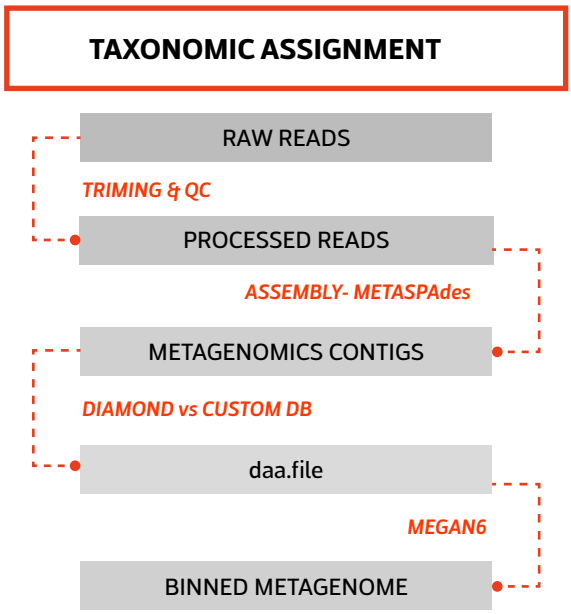
resulting in a total yield between 1 and 4 µg DNA. Paired-end libraries (250-bp) were built either using TrueSeq or Nextera XT DNAlibrary preparation kits (Illumina, San Diego, CA) for most of the samples. However, in addition, two pair-end libraries with average insert length of 3 and 8 kb were built for one sample of *Alectoria sarmentosa*. Sequencing of Nextera XT libraries was carried out by the University of Illinois at Chicago Research Resource Center (Chicago, IL) on Illumina's NextSeq platform and TruSeq libraries were sequenced on Illumina MiSeq platform at the Pritzker Laboratory for Molecular Systematics and Evolution at The Field Museum, Chicago, IL, USA. Raw sequences were downloaded from an Illumina BaseSpace application and quality trimmed and filtered using Trimmomatic-0.36 (Bolger et al. 2014). Bases were trimmed when the average quality of 5-base sliding windows was below 20 and bases at the start and end of reads had a quality below 3 and 10, respectively. Subsequently, all trimmed reads shorter than 36 bp were filtered out.

Genome assembly and taxonomy assignment

Trimmed paired-end reads were assembled using MetaSPAdes (Nurk et al. 2017) using default parameters and checking the suitability of k-mer (K21, K33, K55 and K77). To filter the assembly from potential contaminants and to extract contigs of fungal origin, assemblies were subjected to BLASTX searches using DIAMOND (Buchfink et al. 2015) against a custom database comprising the protein sets of Archaea, Bacteria, Eukaryota, and Viruses of the NCBI non-redundant database (downloaded in August 2016), in addition, 150 complete fungal genomes and 20 algal genomes were added from JGI. Four additional unpublished Parmeliaceae genomes generated from axenic cultures from species within Parmeliaceae [*Cetraria islandica, Parmelina carporrhizans,* unpublished; *Evernia prunastri* and *Pseudevernia furfuracea* (Meiser et al. 2017)] were used as reference genomes for taxonomy assignment. The results of the DIAMOND search were then used as input for MEGAN6 (Huson et al. 2016) for taxonomic assignment (parameters: min-support = 1, min-score = 50, top-hit =10%, no low complexity filtering). Contigs that were assigned as Parmeliaceae were extracted and genome statistics were generated using QUAST v. 4.3 (Gurevichet al. 2013), as shown in Table 1.2 and Figure 1.1. We estimated the coverage of the lichen-forming fungus in each metagenome by mapping the filtered Illumina reads against the respective Parmeliaceae-binned contigs.

Assessing genome completeness of Lichen forming fungi

Evaluation of the completeness of fungal genome assemblies was carried out using the BUSCO v3 pipeline (Simao et al. 2015). This program is useful to detect genome duplication or two genomes in a same sample due to high number of duplicated genes. Every lichenized fungal genome was assessed using 3156 single-copy genes present in at least 90% of the 50 reference genomes in OrthoDB v9 (Waterhouse et al. 2013). BUSCO uses reciprocal best hit (RBH), creating a Hidden Markov Model (HMM) profile using the protein sequences of 50 reference genomes for each single-copy gene. Every HMM profile generated was then used as query in tBLASTn searches against each genome to find the putative genomic region. An AUGUSTUS (Stanke et al. 2004) prediction was performed for each of the genomic regions. Predicted genes were then aligned to the HMM profiles of the BUSCO genes, and only the gene with the highest bit-score was kept. Finally, an additional whole genome gene prediction was conducted with MAKER2 (Holt and Yandell 2011), using Augustus v3.22 (Stanke et al. 2006) and GeneMarkES (Ter-Hovhannisyan et al. 2008).



Taxonomic assignment pipeline

4.3 RESULTS

Sequencing and Metagenome Assembly

The sequencing of lichen metagenomes yielded a different number of short-reads depending of library preparation and number of samples per run. The highest number of reads were found in the sequencing of *Punctelia borreri* 2 (20,672,010 short-reads) and *Hypogymnia subphysodes* (19,446,871 short-reads) metagenomes and the lowest one in *Usnea antarctica* (1,145,486 short-reads) and *Bulbothrix sensibilis* (2,923,755 short-reads). The total metagenome length varied from 331,725,591 bp in *Relicina intertexta* to the lowest metagenome length of 50,717,285 bp in *Oropogon secalonicus* 1. The number of contigs metagenome and N50 was very fluctuant and these were not correlated with total metagenome length (see Table 1.2). This could be highly influenced by the number of small bacterial contigs and other organisms associated to lichen thallus.

Lichen Metagenome Scaffold Diversity

The metagenome of Parmeliacea species was composed in all cases for three main groups of cellular organisms: Bacteria, Opistokonta and Viridiplantae (See Fig 1.1). The major proportion of assigned scaffolds was for the Kingdom Bacteria, representing in the most part of lichen species with more than 60% of scaffold assigned, except *Arctoparmelia centrifuga* (37,3%), *Parmotrema schelpei* (17,2%) and *Usnea strigosa* (25,3%). Inside this kingdom, the groups containing the most number of scaffolds were Proteobacteria, representing around 60% in the most of species, being Alphaproteobacteria the most abundant group. Other groups also contained a high number of assigned scaffolds as Acidobacteria, which varied from 50,4% in *Menegazzia cincinnata* to 3,3% in *Protousnea magellanica*, or Terrabacteria group, where *Protousnea magellanica* and *Parmelia saxatilis* showed the high proportion with a 45,2% and 43,3%, respectively (Table 1.3).

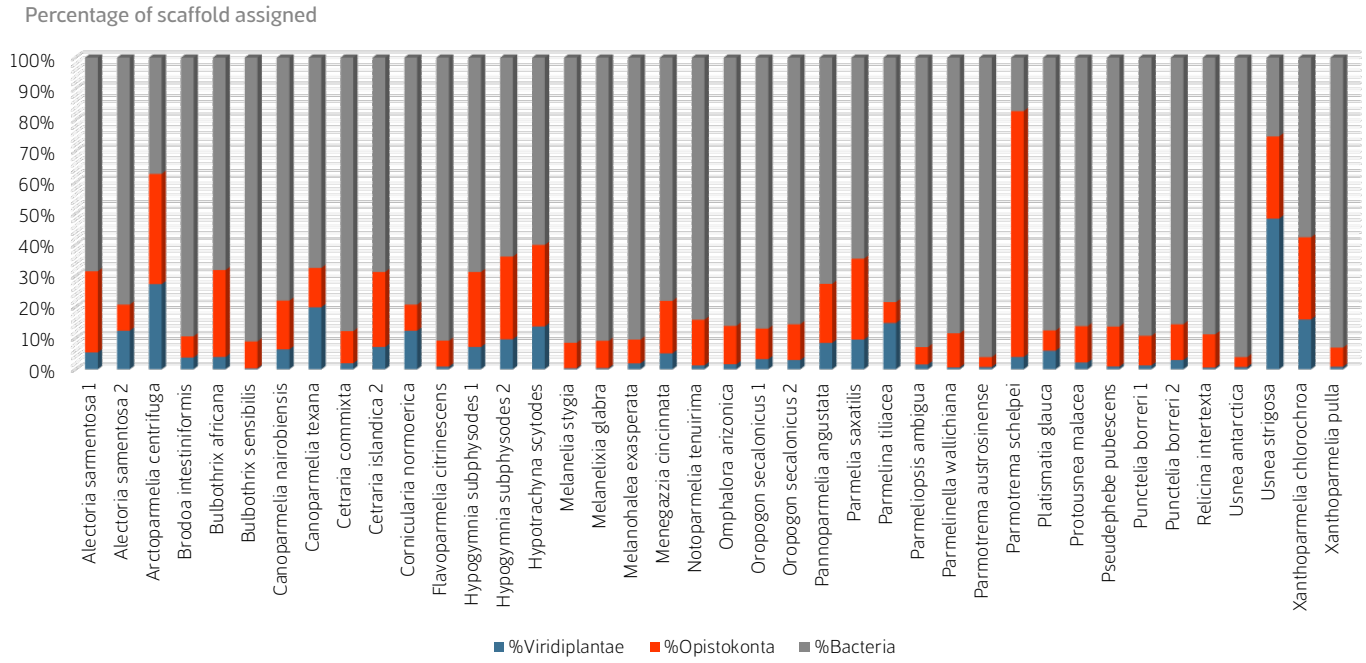


Figure 1.1

Percentage of assigned scaffolds to main cellular organisms groups after taxonomic assignment of lichen metagenomes.

The proportion of Opistokonta scaffold varied from 27% in *Arctoparmelia centrifuga* to 3,2% in *Parmotrema austrosinense*. In all cases the majority of assigned scaffolds to Opistokonta belong to Order Lecanoromycetes (Ascomycota) that will be further analyzed in detail using metric of assemblies and genome completeness. Regarding Viridiplantae scaffolds proportion, in all species the 98% of scaffolds were assigned to the Division Chlorophyta; Class Trebouxiophyceae, where the species with high proportion were *Usnea strigosa* and *Canoparmelia texana* with a 48,3% and 19,9% respectively. On the other side, the lowest proportions were found in *Bulbothrix sensibilis* and *Melanelia stygia* with 0,3% of scaffolds.

Assembly Metrics and Completeness of lichen-forming fungi genomes

The genome assembly metrics for the four specimens represented by axenic culture – *Cetraria islandica*, *Evernia prunastri*, *Parmelina carporrhizans* and *Pseudevernia furfuracea* – and used as part of the reference genome data-base used in the taxonomic assignment, as well as the assembly statistics for the 40 binned metagenomes included in this study are reported in Table 1.2. The length of the assembled mycobiont genomes ranged from 24.1 (*Usnea antarctica*) to 47.6 (*Pannoparmelia angustata*) Mbp, with a mean value of 32.8 Mbp and a %GC content between 47.3 (*Arctoparmelia centrifuga*) to 56.1% (*Oropogon seacalonicus*), with a mean value of 49.2%. Average coverage of the reconstructed mycobint genomes varied from 9.8 (*Bulbothrix sensibilis*) to 95.3-fold (*Hypogymnia subphysodes*) (mean = 38.6-fold) (see Tables 1.4 A & B).

The average number of predicted genes conducted with MAKER2 was 8353. The species *Usnea antarctica* and *Bulbothrix sensibilis* showed the lowest values with 5713 and 6139, respectively. These two species also have the lowest genome completeness, based on the 3156 single copy genes of Pezizomycotina, both of them around the 70% of genome recovered from metaassemblies. The specie *Pannoparmelia angustata* showed the highest number of predicted genes with 15223 and a genome completeness of 86.9% and low number of duplicated BUSCO genes. Other species with a high value of predicted genes were *Parmotrema austrosinense* and *Hypogymnia subphysodes* (2) with 12962 and 12063, respectively. Regarding the highest genome completeness, *Alectoria sarmentosa* (2) showed a 96,3% (Figure 1.2) with a total of 3041 completed BUSCO genes, this values is very similar to the genome completeness of genomes from axenic culture where the number of predicted genes varies from 15364 in *Cetraria islandica* 2 to 8235 in *Gyalolechia flavorubescens* (see Table 1.4 B).

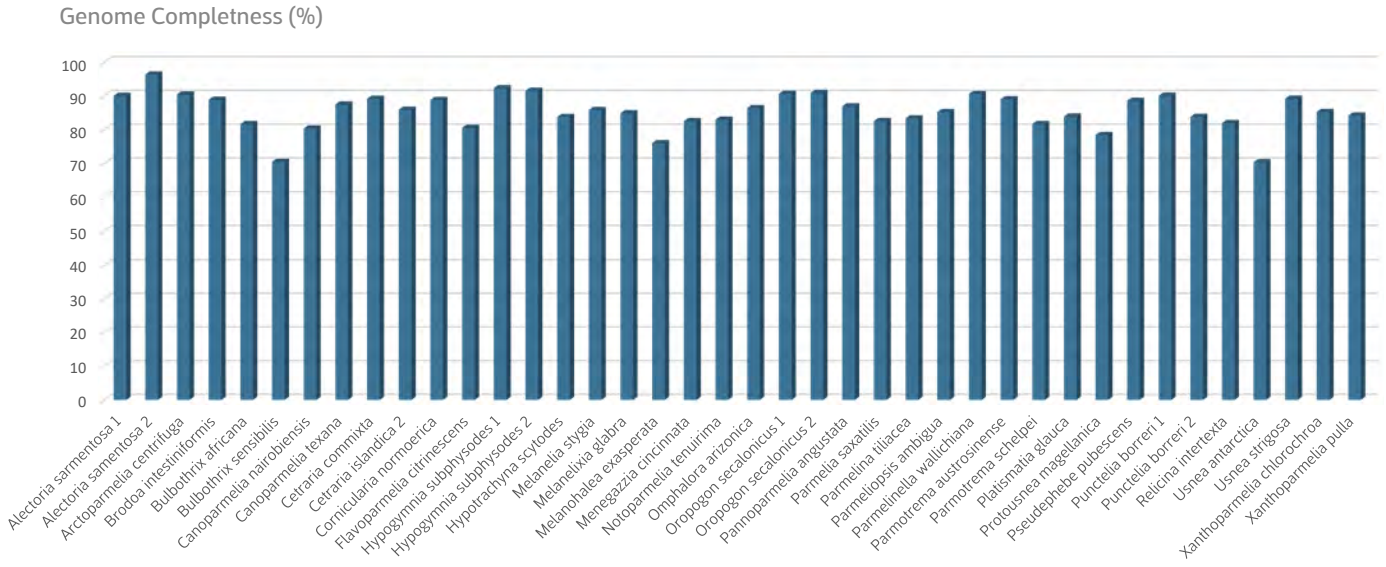


Figure 1.2

Genome completness based on single copy BUSCO genes (Pezizomytotina dataset)

4.4 DISCUSSION

Genome skimming consists of using a single sequencing library layout to sequence to shallow depth individual species' genomes (Elgar et al. 1999). This approach has been successfully used in phylogeographic and phylogenomic studies at different taxonomic levels, from intraspecific (i.e., ultra-barcoding) to family-wide analyses (Male et al. 2014; Dodsworth, 2015; Weitemier et al. 2015; Denver et al. 2016; Gardner et al. 2016; Grandjean et al. 2017). Particularly when working with multi-species assemblages of non-model organisms, this method represents a cost-effective approach to survey the genomes of the individual species in the mixture (Greshake et al. 2016; Quince et al. 2017). The skimming approach is increasingly used in lichens, including multispecies fungal–algal symbioses harboring complex fungal and bacterial microbiomes to tackle genome-wide phylogenetics (Leavitt et al. 2016c; Grewe et al. 2017; Pizarro et al. 2018) and population genetics (Dal Grande et al. 2017), as well as biotechnological aspects of the lichen symbiosis (Kampa et al. 2013).

Recently, it has been shown that metagenome skimming is a valid approach to retrieve almost the entire gene space of lichen-forming fungi from whole lichen thalli (Meiser et al. 2017). The genome completeness estimates of most of the reconstructed fungal genomes in our study are in line with reports from a previous study (Meiser et al. 2017), i.e. 85–90% (Fig. 1.2). Further, the numbers of predicted genes in almost all binned genomes are concordant to other recently sequenced genomes of lichen-forming fungi from culture (e.g., *Gyalolechia flavorubescens*: 9,695 genes; Park et al. 2013. *Cladonia macilenta*: 7,322; Park et al. 2013. *Endocarpon pusillum*: 9,285; Wang et al. 2014. and *Ramalina intermedia*: 8,871; Wang et al. 2018) (Table 1.4). While the scaffolds proportions of Bacteria and Viridiplantae (Table 1.3) were highly similar to those reported in previous studies (Grube et al. 2015; Cernava et al. 2017; Dal Grande et al. 2018), a further investigation will be needed in order to explore diversity of microorganisms associated to lichen thallus.

4.5 CONCLUSION

We show that this approach is a suitable method for recovering the lichen forming fungi part of holobiont. The sizes of assembled mycobiont genomes were ranged from 24.1 (*Usnea antarctica*) to 47.6 (*Pannoparmelia angustata*) Mbp, with a mean value of 32.8 Mbp. The genome completeness was ranged between 70% (*Bulbothrix sensibilis* and *Usnea antarctica*) to 96,3% (*Alectoria sarmentosa 2*). In addition, this study highlights the need of further investigation to explore microbial diversity associated with Parmeliaceae lichen thallus. In the next chapters the binned genomes will be used for answering different questions related with evolutionary biology of lichen - forming fungi as the extraction of genome-wide phylogenetic markers for resolving the deep evolutionary relationship inside of the family Parmeliaceae, as well as comparative genomic analysis to deciphering the genetic mechanisms behind of sexual reproduction or the discover the diversity of putative secondary metabolites.

Table 1.1
Samples used in the study including voucher information and/ or accession numbers.

Genus and species	Family	Genomic DNA source	Voucher/Accession number	Location
Alectoria sarmentosa 1	Parmeliaceae	Apothecia	Wedin 9564 (S, MAF-Lich)	Sweden: Jämtland, Wedin 13.09.2014
Alectoria samentosa 2	Parmeliaceae	Thalli	21536, (MAF-Lich)	Norway: Telemark, Boluda 20.08.2015
Arctoparmelia centrifuga	Parmeliaceae	Apothecia	Wedin 9557 (S, MAF-Lich)	Sweden: Jämtland, Wedin 13.09.2014
Brodoa intestiniformis	Parmeliaceae	Apothecia	Wedin 9560 (S, MAF-Lich)	Sweden: Jämtland, Wedin 13.09.2014
Bulbothrix africana	Parmeliaceae	Apothecia	Kirika 4819 & Lumbsch (F)	Kenya: Western Province, Kirika & Lumbsch 2015
Bulbothrix sensibilis	Parmeliaceae	Apothecia	Kirika 3860 (EA, MAF-Lich)	Kenya: Eastern Province, Kirika & Lumbsch 12.03.2014
Canoparmelia nairobiensis	Parmeliaceae	Apothecia	Kirika 4423 (EA, MAF-Lich)	Kenya: Central Province, Kirika 10.04.2014
Canoparmelia texana	Parmeliaceae	Thalli	Kirika 5254 (EA, F)	Kenya: Rift Valley Province, Kirika & Divakar 25.05.2017
Cetraria commixta	Parmeliaceae	Apothecia	21548, (MAF-Lich)	Portugal: Beira Alta, Crespo et al. 11.06.2014
Cetraria islandica 1	Parmeliaceae	Apothecia	Wedin 9562 (S, MAF-Lich)	Sweden: Jämtland, Wedin 13.09.2014
Cornicularia normoerica	Parmeliaceae	Apothecia	21549, (MAF-Lich)	Portugal: Beira Alta, Crespo et al. 11. 06. 2014
Flavoparmelia citrinescens	Parmeliaceae	Apothecia	19333, (MAF-Lich)	Chile: Araucania, Elvebakk 21.12.2013
Hypogymnia subphysodes 1	Parmeliaceae	Thalli	Lumbsch 20921a (F)	Australia: Western Australia, Lumbsch & Huang 20921a 26.03.2018
Hypogymnia subphysodes 2	Parmeliaceae	Thalli	Lumbsch 20926a (F)	Australia: Western Australia, Lumbsch & Huang 20926a 26.03.2018
Hypotrachyna scytodes	Parmeliaceae	Apothecia	Kirika 4236 (EA, MAF-Lich)	Kenya: Western Province, Kirika 29.03.2014
Melanelia stygia	Parmeliaceae	Apothecia	Wedin 9555 (S, MAF-Lich)	Sweden: Jämtland, Wedin 13.09.2014
Melanelixia glabra	Parmeliaceae	Apothecia	21550, (MAF-Lich)	Spain: Madrid, Divakar & Crespo 20.11.2014
Melanohalea exasperata	Parmeliaceae	Apothecia	21551, (MAF-Lich)	Portugal: Beira Alta, Crespo et al. 6998ay, 11.06.2014
Menegazzia cincinnata	Parmeliaceae	Apothecia	21552, (MAF-Lich)	Chile: IX Region, Villagra 2013
Notoparmelia tenuirima	Parmeliaceae	Apothecia	21554, (MAF-Lich)	Australia: Victoria, Crespo, Gavilan & Divakar 6925A, 22.07.2011
Omphalora arizonica	Parmeliaceae	Apothecia	55768, (BRY-C)	USA: New Mexico, Leavitt & St. Clair 20128
Oropogon secalonicus 1	Parmeliaceae	Thalli	8728, (KUN)	China: Yunnan Province, Wang et al. 12-38331
Oropogon secalonicus 2	Parmeliaceae	Thalli	14429, (HMAS-L)	China: Yunnan Province, Wei
Pannoparmelia angustata	Parmeliaceae	Thalli	Lumbsch 2166 (F)	Australia: Tasmania, Widhelm, Lumbsch & Grewe 2166 26.03.2018
Parmelia saxatilis	Parmeliaceae	Apothecia	Wedin 9551 (S, MAF-Lich)	Sweden: Hälsingland, Wedin 12.09.2014
Parmelina tiliacea	Parmeliaceae	Apothecia	21556, (MAF-Lich)	Spain: Madrid, Divakar & Crespo 20.11.2014
Parmeliopsis ambigua	Parmeliaceae	Apothecia	Wedin 9561 (S, MAF-Lich)	Sweden: Jämtland, Wedin 13.09.2014
Parmelinella wallichiana	Parmeliaceae	Thalli	Kirika 5235 (EA, F)	Kenya: Rift Valley Province, Kirika & Divakar 25.5.2017
Parmotrema austrosinense	Parmeliaceae	Thalli	Kirika 5247 (EA, F)	Kenya: Rift Valley Province, Kirika & Divakar 25.5.2017
Parmotrema schelpei	Parmeliaceae	Apothecia	Kirika 3963 (EA, MAF-Lich)	Kenya: Coast Provience, Kirika & Lumbsch 15.03.2014
Platismatia glauca	Parmeliaceae	Apothecia	21558, (MAF-Lich)	Spain: Caceres, Divakar et al. 18.07.2013
Protousnea malacea	Parmeliaceae	Apothecia	19345, (MAF-Lich)	Chile: VIII Region del Bio-Bio, Elvebakk 16.01.2014
Pseudephebe pubescens	Parmeliaceae	Apothecia	Wedin 9559 (S, MAF-Lich)	Sweden: Jämtland, Wedin 13.09.2014
Punctelia borreri 1	Parmeliaceae	Apothecia	21555, (MAF-Lich)	Australia: Victoria, Crespo, Gavilan & Divakar 6922h , 20.07.2011
Punctelia borreri 2	Parmeliaceae	Thalli	Kirika 5257 (EA, F)	Kenya: Rift Valley Province, Kirika & Divakar 25.5.2017
Relicina intertexta	Parmeliaceae	Apothecia	21559, (MAF-Lich)	Thailand: Nakhon Ratchasima Province, Divakar, Boluda & Bouprong 09.08.2014
Usnea antarctica	Parmeliaceae	Thalli	EL0409 (F)	Antarctica: King George Island, Elisa 08.04.2016
Usnea strigosa	Parmeliaceae	Thalli	TW2605 (F)	USA: Arkansas, Widhelm 19.05.2016
Xanthoparmelia chlorochroa	Parmeliaceae	Apothecia	9866, (BRY-C)	USA: Nevada, Leavitt & St. Clair
Xanthoparmelia pulla	Parmeliaceae	Apothecia	21557, (MAF-Lich)	Spain: Madrid, Divakar & Crespo 20.11.2014
Cetraria islandica 2	Parmeliaceae	Axenic mycobiont culture	17201, (MAF-Lich)	Spain: Segovia, Crespo& Rico 12.2015
Cladonia grayi	Cladoniaceae	Axenic mycobiont culture	JGI/Cgr/DA2myc/ss v2.0	
Cladonia macilenta	Cladoniaceae	Axenic mycobiont culture	KoLRI003786/AUPP01000000	China: Yunnan Province
Cladonia metacorallifera	Cladoniaceae	Axenic mycobiont culture	KoLRI002260/AXCT00000000	South Korea: Gangwon-do
Evernia prunastri	Parmeliaceae	Axenic mycobiont culture	Schmitt (SBiK-F), C0001/NKYR00000000	Spain: Madrid, Dal Grande & Divakar 06.2012
Gyalolechia flavorubescens	Teloschistaceae	Axenic mycobiont culture	KoLRI002931/AUPK01000000.	South Korea: Muju
Lasallia hispanica	Umbilicariaceae	Axenic mycobiont culture	FR-0265086; Schmitt (SBiK-F C0002)	Spain: Ávila, Dal Grande & Divakar 25.07.2014
Parmelina carporrhizans	Parmeliaceae	Axenic mycobiont culture	19191, (MAF-Lich)	Spain: Ávila, Crespo et al. 10.2012
Pseudevernia furfuracea	Parmeliaceae	Axenic mycobiont culture	AKPM0122M/NKYQ00000000	Slovenia. Olivetoric acid chemotype, Yoshimura, 17.10.1996
Rhizoplaca melanophthalma	Lecanoraceae	Axenic mycobiont culture	LMCC0506 (GZU)	USA.: Colorado, Fernandez-Mendoza 18.07.2009
Xanthoria parietina	Teloschistaceae	Axenic mycobiont culture	JGI/46-1-SA22 v1.1	

Table 1.2

Whole-metagenome assembly metrics

Species	Number of reads	Total Length	GC%	Total Contigs	N50
Alectoria sarmentosa 1	13658918	121041647	47.49	89892	7022
Alectoria samentosa 2	13798851	122359974	46.35	342891	86321
Arctoparmelia centrifuga	15848677	252502022	54.39	520022	3676
Brodoa intestiniformis	7463461	147495048	51.74	89068	2522
Bulbothrix africana	5135277	102384724	49.7	20176	11598
Bulbothrix sensibilis	2923755	122017941	51.16	307133	5671
Canoparmelia nairobiensis	5125947	211896626	54.59	592877	2730
Canoparmelia texana	14748096	113686614	49.71	219123	3543
Cetraria commixta	5718965	140138668	48.54	334946	7427
Cetraria islandica 2	5598856	116151959	49.46	186155	2729
Cornicularia normoerica	8223289	202266657	49.4	578067	2574
Flavoparmelia citrinescens	6265625	258826468	52.18	817412	3166
Hypogymnia subphysodes 1	19446871	167119643	54.38	367044	17867
Hypogymnia subphysodes 2	17839449	182700872	58.89	1011122	2312
Hypotrachyna scytodes	4788408	142533158	49.67	360350	5238
Melanelia stygia	4799432	203987769	53.56	479883	1339
Melanelixia glabra	3830633	82194787	51.56	172451	13230
Melanohalea exasperata	5726706	142049358	52.76	341171	2814
Menegazzia cincinnata	5507102	225683514	53.3	654432	2330
Notoparmelia tenuirima	6581834	316520788	57.24	934903	1447
Omphalora arizonica	7342868	133854498	53.12	283011	4547
Oropogon secalonicus 1	8178754	50717285	54.82	180207	7800
Oropogon secalonicus 2	12144821	90663783	52.17	417552	2178
Pannoparmelia angustata	17842460	112509529	55.47	444571	2343
Parmelia saxatilis	5181447	173648292	55.75	431336	3579
Parmelina tiliacea	4575261	73633639	48.6	140740	15090
Parmeliopsis ambigua	18793305	165776902	52.56	421122	3937
Parmelinella wallichiana	5356063	160954090	51.58	340931	1722
Parmotrema austrosinense	16193386	135468817	51.5	151753	3800
Parmotrema schelpei	4956613	207602242	56.62	524671	2702
Platismatia glauca	6136910	198403014	60.42	404483	4099
Protousnea magellanica	4537711	61313715	47.58	107042	10254
Pseudephebe pubescens	8995576	193041864	51.73	337924	2772
Punctelia borreri 1	4978384	79197312	49.53	151902	11729
Punctelia borreri 2	20672010	170442767	58.08	325561	1361
Relicina intertexta	10437739	331725591	54.93	920983	1421
Usnea antarctica	1145486	149788489	47.82	563272	1694
Usnea strigosa	13744510	107743148	49.48	124587	2660
Xanthoparmelia chlorochroa	5657981	137240983	50.8	167997	2784
Xanthoparmelia pulla	5211756	161577513	52.6	510813	5154

Table 1.3
Bacteria scaffolds diversity

Species	Acido %	FCB%	Proteo%	PVC%	Spiro%	Terra%
Alectoria sarmentosa 1	31,04	1,72	62,96	1,39	0,17	2,72
Alectoria samentosa 2	38,23	1,78	50,21	0,80	0,14	8,84
Arctoparmelia centrifuga	17,84	0,64	46,11	14,06	0,03	21,31
Brodoa intestiniformis	25,13	0,90	58,02	4,94	0,03	10,98
Bulbothrix africana	25,13	0,90	58,02	4,94	0,03	10,98
Bulbothrix sensibilis	27,60	5,90	30,82	2,41	0,03	33,24
Canoparmelia nairobiensis	26,31	1,36	50,75	5,23	0,02	16,34
Canoparmelia texana	42,32	6,85	39,69	2,23	0,02	8,89
Cetraria commixta	39,37	2,71	54,27	1,80	0,05	1,79
Cetraria islandica 2	41,17	0,36	53,45	2,13	0,03	2,85
Cornicularia normoerica	37,55	1,18	46,32	2,05	0,03	12,86
Flavoparmelia citrinescens	4,43	21,88	35,85	1,78	0,02	36,03
Hypogymnia subphysodes 1	22,33	1,20	58,79	0,36	0,14	17,17
Hypogymnia subphysodes 2	9,54	21,63	27,23	2,50	0,04	39,06
Hypotrachyna scytodes	32,54	2,39	49,15	4,09	0,04	11,80
Melanelia stygia	29,51	2,71	27,25	1,40	0,03	39,11
Melanelixia glabra	22,64	10,07	52,14	2,27	0,02	12,85
Melanohalea exasperata	30,12	13,87	38,56	3,91	0,02	13,52
Menegazzia cincinnata	50,48	2,39	9,60	6,83	0,05	30,65
Notoparmelia tenuirima	19,55	1,51	36,67	0,99	0,02	41,27
Omphalora arizonica	27,00	6,27	60,22	1,04	0,02	5,45
Oropogon secalonicus 1	19,97	5,42	66,73	0,40	0,21	7,28
Oropogon secalonicus 2	22,84	6,74	63,87	0,56	0,35	5,63
Pannoparmelia angustata	38,02	0,76	50,02	0,74	0,08	10,39
Parmelia saxatilis	24,14	1,62	28,46	2,46	0,02	43,29
Parmelina tiliacea	15,17	26,89	37,96	1,79	0,04	18,15
Parmeliopsis ambigua	27,57	3,13	56,38	5,23	0,03	7,66
Parmelinella wallichiana	38,02	0,76	50,02	0,74	0,08	10,39
Parmotrema austrosinense	26,03	0,59	64,40	0,39	0,00	8,59
Parmotrema schelpei	32,42	0,74	35,21	1,14	0,02	30,47
Platismatia glauca	23,88	0,66	59,25	5,79	0,03	10,40
Protousnea malacea	3,35	0,64	50,64	0,19	0,00	45,17
Pseudephebe pubescens	28,94	0,80	46,34	3,48	0,04	20,40
Punctelia borreri 1	5,05	20,02	31,96	1,96	0,02	40,98
Punctelia borreri 2	35,07	5,60	46,68	4,30	0,02	8,34
Relicina intertexta	32,06	0,70	47,06	4,33	0,03	15,81
Usnea antarctica	15,62	5,03	36,81	36,06	0,03	6,45
Usnea strigosa	6,33	0,00	86,75	0,43	0,00	6,48
Xanthoparmelia chlorochroa	4,12	24,00	48,30	0,59	0,00	22,99
Xanthoparmelia pulla	22,41	11,38	39,34	2,10	0,04	24,72

* **Acido**=Acidobacteria, **FCB**= Sphingobacteria, **Proteo**= Proteobacteria, **PVC**=Planctobacteria, **Spiro**= Spirochaete, **Terra**= Terrabacteria

Table 1.4-A
Binned Metagenome assembly metrics

Species	Total Length	GC%	Contigs	N50	Coverage	Parmeliacea length	Number of predicted genes	BUSCO Genes	% Completeness
Alectoria sarmentosa 1	33052253	48.35	3342	25327	54	27,30651294	7854	2842	90,05
Alectoria samentosa 2	46540876	40.34	1788	92863	49	33,13457389	9695	3041	96,36
Arctoparmelia centrifuga	40459659	47.29	6931	23666	74	16,02349901	7565	2855	90,46
Brodoa intestiniformis	38270146	47.46	11324	16037	24	25,9467328	7292	2805	88,88
Bulbothrix africana	30662347	47.95	3502	17258	32	29,9481659	6594	2577	81,65
Bulbothrix sensibilis	26829917	48.68	5085	9396	98	21,98850167	6139	2225	70,50
Canoparmelia nairobiensis	27201108	49.27	4073	11657	19	12,8369708	7353	2536	80,35
Canoparmelia texana	36952011	49.03	10351	16147	75	32,50339657	10209	2761	87,48
Cetraria commixta	33545404	48.11	7219	15855	31	23,93729331	7414	2814	89,16
Cetraria islandica 2	34044480	48.57	5060	24037	33	29,31029342	7692	2712	85,93
Cornicularia normoerica	30301608	47.72	13048	19303	32	14,98101983	7208	2803	88,81
Flavoparmelia citrinescens	26342035	49.24	4828	12180	18	10,17748888	7016	2543	80,58
Hypogymnia subphysodes 1	44273933	49.66	31631	36657	85	26,49235733	7639	2913	92,30
Hypogymnia subphysodes 2	51700170	50.34	45304	20279	79	28,29771387	12063	2889	91,54
Hypotrachyna scytodes	28847286	49.00	17991	14076	24	20,23900011	7730	2644	83,78
Melanelia stygia	31606081	49.12	6915	21268	18	15,49410592	7594	2709	85,84
Melanelixia glabra	29221822	49.84	3214	17361	24	35,55191645	7446	2680	84,92
Melanohalea exasperata	29551578	49.77	10174	8923	29	20,80373922	7040	2399	76,01
Menegazzia cincinnata	32469692	49.12	26921	79423	13	14,387268	7231	2605	82,54
Notoparmelia tenuirima	29178179	47.87	4734	14279	17	9,218408429	6204	2619	82,98
Omphalora arizonica	28699306	48.76	3711	16491	43	21,44067359	7850	2727	86,41
Oropogon secalonicus 1	37000476	56.09	194774	14638	87	72,95437049	6751	2861	90,65
Oropogon secalonicus 2	31633776	50.14	11109	14922	40	34,89130384	10334	2869	90,91
Pannoparmelia angustata	47561874	52.42	60516	9925	17	42,27364066	15223	2743	86,91
Parmelia saxatilis	30564937	49.62	5762	16605	23	17,60163411	7364	2606	82,57
Parmelina tiliacea	31229044	49.04	3579	17640	29	42,41138211	8719	2631	83,37
Parmeliopsis ambigua	36254821	48.76	22351	17440	74	21,86964563	12331	2690	85,23
Parmelinella wallichiana	32514193	48.27	6866	21266	28	20,20091133	11033	2859	90,59
Parmotrema austrosinense	43516517	49.46	9175	14288	74	32,12290324	12962	2810	89,04
Parmotrema schelpei	29016443	48.51	7361	13723	15	13,97694106	7395	2577	81,65
Platismatia glauca	28681065	51.1	3607	17398	20	14,45596235	9106	2648	83,90
Protousnea magellanica	30922013	48.5	6278	11553	20	50,43245708	7592	2476	78,45
Pseudephebe pubescens	29125554	48.72	3646	19539	50	15,08768792	8212	2795	88,56
Punctelia borreri 1	37682510	49.91	16328	17333	33	47,58054162	8247	2843	90,08
Punctelia borreri 2	31384725	48.42	4528	15250	84	18,41364439	7439	2645	83,81
Relicina intertexta	31836904	47.79	15979	12463	37	9,597361453	7099	2587	81,97
Usnea antarctica	24119594	48.79	1100	23562	21	16,10243495	5713	2221	70,37
Usnea strigosa	37557709	48.92	8736	14477	65	34,85855917	8247	2815	89,20
Xanthoparmelia chlorochroa	28326562	49.27	3078	18494	26	20,64001684	6713	2691	85,27
Xanthoparmelia pulla	27802724	49.11	7345	14590	24	17,20705034	7818	2657	84,19

Table 1.4-B

Genome assembly metrics

SPECIES	length	GC	Contig	N50	Predicted Genes
Cladonia grayi	34622149	44,44	414	243412	11389
Cladonia macilenta	36862809	44,77	66	1469036	9237
Cladonia metacoralifera	36682060	44,91	30	1591850	9296
Cetraria islandica	51316390	44,41	3472	95207	15364
Evernia prunastri	40346456	48,97	227	264454	10984
Gyalolechia flavorubescens	34468235	41,89	36	1693300	8235
Lasallia hispanica	41207996	51,28	1619	145035	11973
Parmelina carporhizans	34751387	48,78	11765	9802	9434
Pseudevernia furfuracea	37795223	47,86	46	1178799	9211
Rhizoplaca melanophthalma	33087415	48,8	1430	49682	9223
Xanthoria parietina	31900637	49,89	39	1731186	10818

5.CAPÍTULO 2

FILOGENÓMICA DE HONGOS SIMBIONTES (FAMILIA PARMELIACEAE)

5.CHAPTER 2

PHYLOGENOMIC OF SYMBIONT FUNGI (FAMILY PARMELIACEAE)

5

5.1 INTRODUCTION

Progress in sequencing technology has facilitated the efficient generation of genome-scale DNA sequence data and has revolutionized biological research in a wide range of fields. In phylogenetic systematics, genomic data has provided unprecedented insight into evolutionary relationships with the potential to generate thousands of independent phylogenetic markers distributed across genomes (Eisen and Fraser 2003; Murphy et al. 2004; Delsuc et al. 2005; Philippe et al. 2005; Meusemann et al. 2010; Torruella et al. 2012; Misof et al. 2014; Spatafora et al. 2017). Types of loci used in phylogenomic analyses range from restriction site-associated DNA sequencing [RADseq; (Andrews et al. 2016)] to target enrichment to capture known protein-coding regions distributed across the genome (Bragg et al. 2016) or ultraconserved elements (UCEs) (Faircloth et al. 2012). Single-copy genes, or single-copy gene families, are valuable molecular markers for inferring relationships of previously unresolved lineages across Eukaryota (Shenet al. 2013; Salas-Leiva et al. 2014; Zeng et al. 2014; Ren et al. 2016). Generally single-copy genes are easily amplified and sequenced, because they occur only once in the genome and they are highly conserved across species. Whole genome sequencing can facilitate the identification of single-copy genes as potential phylogenetic markers. Two decades after the first fungal genome was sequenced, genomes from over 800 fungal species are now available (Spatafora et al. 2017). These data have been widely used to address phylogenetic relationships in fungi, especially Ascomycota (Robbertse et al. 2006; Liu et al. 2009; Wang et al. 2009; Medina et al. 2011; Ebersberger et al. 2012; Ekanayake et al. 2013; Zheng et al. 2013; Payen et al. 2014; Hettiarachchige et al. 2015; Luo et al. 2015; Lachance et al. 2016; Whiston and Taylor 2016; Choi and Kim 2017; Menardo et al. 2017; Spatafora et al. 2017; Teixeira et al. 2017). However, the number of phylogenomic studies addressing evolutionary relationships in lichen-forming

fungi, which represent an important part of ascomycete diversity (Ott and Lumbsch 2001; Nash 2008; Jaklitsch et al. 2016), is still low and limited to resolving relationships among closely related species (Leavitt et al. 2016c; Grewe et al. 2017). Here we generated a phylogenomic data set to address evolutionary relationships among major clades in the sub-family Parmelioideae of Parmeliaceae, which represents one of the most diverse clades of lichen-forming fungi (Kraichak et al. 2015; Divakar et al. 2017; Kraichak et al.2017). Parmeliaceae includes almost 2800 currently accepted species (Jaklitsch et al.2016), but the number of species is probably higher due to the frequent presence of cryptic species in this group (Kroken and Taylor 2001; Crespo and Perez-Ortega 2009; Divakar et al. 2010b; Lumbsch and Leavitt 2011; Molina et al. 2011a, b; Núñez-Zapata et al. 2011; Amo de Paz et al. 2012; Leavitt et al. 2012, 2016a, b; Altermann et al. 2014; Singh et al. 2015; Alors et al. 2016; Del-Prado et al. 2016). Estimates suggest that the family originated during the Cretaceous with increase diversification rates after the Cretaceous-Paleo-gene (K-Pg) boundary and diversification within genera predominantly during the Miocene, resulting in the current hyper-diversity of the family (Amo de Paz et al. 2011; Kraichak et al. 2015; Huang et al. 2017). Parmeliaceae occurs in all terrestrial ecosystems and all continents, including Antarctica, but has its highest species diversity in tropical montane areas and temperate to sub-tropical regions with Mediterranean-type climate (Crespo et al. 2010b; Thell et al. 2012). Members in the family grow on various substrates, including tree bark, wood, rocks, and soil. Species of Parmeliaceae have been widely used as bioindicators for air pollution and ecological health (Hawksworth and Rose 1970, 1976; Crespo et al. 1999).Additional human uses of Parmeliaceae include dyes, perfumes and medicinal remedies with pharmacological potential in folk medicines (Gomez-Serranillos et al. 2014). The circumscription of the family Parmeliaceae, its sister-group relation, and the classification into genera within the family has been controversial and much discussed (as reviewed in (Elix 1993; Crespo et al. 2010b; Thell et al. 2012; Singh et al. 2013). However, phylogenetic studies (e.g., Mattsson and Wedin 1999; Wedin et al. 1999; Arup et al. 2007; Amo de Paz et al. 2010a, b, 2011; Crespo et al. 2010a, b; Divakar et al. 2010a; 2012, 2013, 2015, 2017; Thell et al. 2012; Singh et al. 2013; Kirika et al. 2015, 2016a, b) resulted in a general consensus on the circumscription of the family, and the sister-group relationship with Gypsoplacaceae (Singh et al. 2013). More recently, two subfamilies have been proposed—Protoparmelioideae and Parmelioideae (Divakar et al. 2017), including a revised circumscription of

genera based on strongly supported monophyletic groups of comparable ages (Divakar et al. 2017). Within the large subfamily Parmelioideae, which includes the bulk of species in Parmeliaceae, a number of strongly supported monophyletic clades have been identified in previous studies (Crespo et al. 2007b; Divakar et al. 2015). These were given preliminary clade names derived from a characteristic genus, e.g., the cetrarioid clade including the genus Cetraria Ach. or the parmelioid clade including the genus Parmelia Ach. A recent multi-gene study based on six loci and 293 operational taxonomic units (OTU) recognized seven strongly supported major clades within Parmelioideae, most of which were also identified in a previous study (Crespo et al.2007): (1) Alectorioid [incl. Alectoria Ach., Bryocaulon Kärnefelt, Bryoria Brodo & D. Hawksw., Nodobryoria R.S. Common & Brodo, Pseudephebe M. Choisy], (2) Anzioid [incl. Anzia Garovagl., Pannoparmelia (Müll. Arg.) Darbishire, Protousnea (Motyka) H. Krog, Raesaenenia D. Hawksw., Boluda & H. Lindgr. – recently (Divakar et al. 2017) included in Protousnea], (3) Cetrarioid [incl. five genera following (Divakar et al. 2017), namely Cetraria, Dactylina Nyl., Esslingeriana Hale & M.J.Lai, Melanelia Esslinger, Nephromopsis Müll.Arg.], (4) Hypogymnioid [incl. Arctoparmelia Hale, Brodoa Goward, Hypogymnia (Nyl.) Nyl., Pseudevernia Zopf], (5) Parmelioid [incl. c. 25 genera following (Divakar et al. 2017)], (6) Psiloparmelioid [incl. Everniopsis Nyl., Psiloparmelia Hale], and (7) Usneoid [incl. Cornicularia (Schreb.) Hoffm., Dolichousnea (Ohmura) Articus, Eumitria Stirt., Usnea Adans.]. While these monophyletic clades with more than one genus received strong support in multi-gene studies, relationships among the major clades largely remained unsupported. The aim of this study was to use phylogenomic data to resolve evolutionary relationships among major clades in the diverse subfamily Parmelioideae. Our specific objectives were twofold: (1) identify genome-wide single-copy protein coding genes and infer relationships among the major clades within Parmeliaceae and (2) assess if phylogenomic datasets can effectively be obtained from metagenomic sequences (Huson et al. 2009; Mitra et al. 2009) to elucidate evolutionary relationships in symbiotic fungi. For this purpose, we used a low-coverage shotgun sequencing (i.e., metagenome skimming) approach to recover genome-wide phylogenetic markers for 46 lichen-forming fungi, representing six of the seven major clades of the subfamily Parmelioideae identified previously (Divakar et al. 2015, 2017).

5.2 MATERIALS AND METHODS

Taxon Sampling

A total of 51 lichen-forming fungal species were included in this study (Table 1.1, Chapter 1). We selected 46 samples representing six of the seven previously identified major clades in subfamily Parmelioideae of Parmeliaceae (Divakar et al. 2015, 2017). The psiloparmelioid clade was not included, because no fresh material was available. We included genome sequences from seven additional species belonging to other families in Lecanorales (*Cladonia grayi*, *C. macilenta*, *C. metacorallifera* and *Rhizoplaca melanopthalma*), and the orders Teloschistales (*Gyalolechia flavorubescens* and *Xanthoria parietina*) and Umbilicariales (*Lasallia hispanica*; Dal Grande et al. 2018), which also belong to the class Lecanoromycetes, as outgroups.

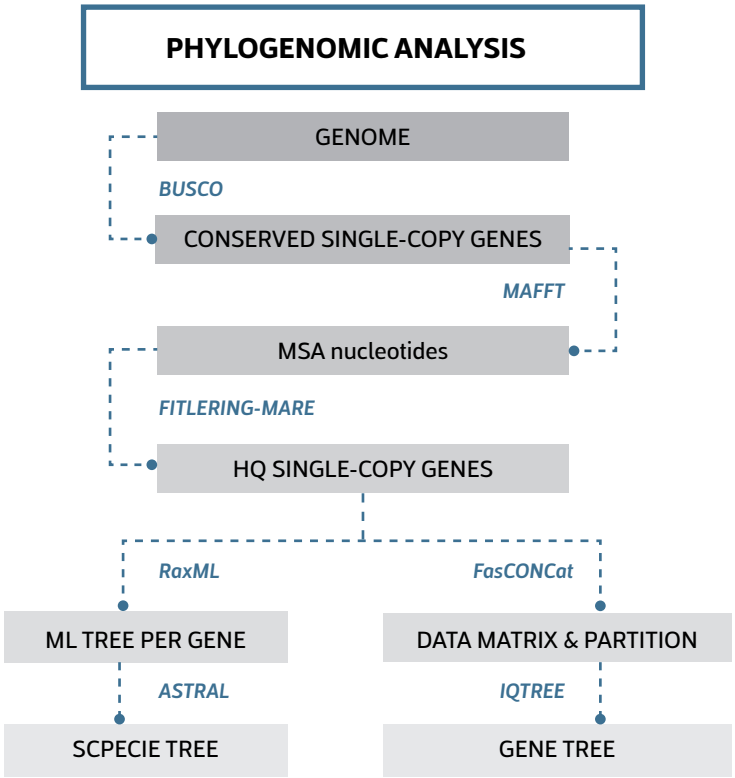
Orthologs Identiffication

The complete BUSCO single-copy genes predicted in each genome were extracted (First Chapter); duplicated and fragmentary genes were removed using an in-house Perlscript. A supermatrix was created concatenating every BUSCO single-copy gene using FASconCAT.pl. (Misof et al. 2013). Every BUSCO single-copy gene recovered from each genome was aligned using MAFFT L-INS-i (–localpair –maxiterate1000). In order to reduce the effects of missing data, individual gene alignments with more than 20% of missing data were removed. A supermatrix was produced concatenating every BUSCO single-copy gene using FASconCAT.pl. Ambiguous regions within the alignment were removed using Gblocks v. 0.91b (Castresana2000) implementing default parameters with the exception of using the option 'with-hal' gaps and a 'minimum length of a block' of 5.

Phylogenomic analyses

Evolutionary relationships were estimated both from a supermatrix comprising the concatenated single-locus alignments under maximum likelihood (ML) and using species tree interferences implementing the multispecies coalescent model (Degnan and Rosenberg 2006). ML analyses of the partitioned-marker supermatrix were run using IQTree v.1.5.5 with standard model selection (Nguyen et al. 2015), implementing the GTR+G+I model using ModelFinder (Chernomor et al. 2016). For each analysis, 1000 bootstrap replicates were calculated using the fast bootstrapping option implemented in IQTree. Trees were drawn with FigTree v. 1.4.2 (Rambaut 2009). Based on previous results (Miadlikowska et al. 2014), the resulting tree was rooted with *Lasallia hispanica*.

Since phylogenetic inferences from concatenated data-sets may differ from species tree approaches (Knowles2009), we inferred species-trees for Parmeliaceae using the summary coalescent approach ASTRAL-II (Mirarab and Warnow 2015). We used ASTRAL-II with multi-locus bootstrapping (MLBS) option. Nucleotide substitution models were inferred for each locus using the program PartitionFinder v1.1.1 (Lanfear et al. 2012) with Akaike Information Criterion model selection.



Pylogenomic analysis pipeline

5.3 RESULTS AND DISCUSSION

Genomic data and genome assembly

The percentage of complete BUSCO genes among 51 genomes varied from 70.51 to 98.4% of 3156 Pezizomycotina BUSCO genes, with an average of 84.14% (Fig2.1). From the total of 3156 BUSCO genes, between 2213 and 3015 could be recovered for each individual genome. Only two of the 51 lichen forming fungi had 80% of 3156 Pezizomycotina BUSCO genes, with the assemblies of *Bulbothrix sensibilis* (2223 genes) and *Usnea antarctica* (2213 genes) representing the lowest values of completeness with more than 900 BUSCO genes ''missing'', ''fragmented'' or ''duplicated'' (Figure 2.1). Out of 3156 BUSCO genes, 2556 were present in 80% of taxa and we reused for the construction of individual gene trees or included into a concatenated super matrix (Table 2.2).

Phylogenomic Dataset

The complete and Gblocks-filtered BUSCO data sets were deposited in FigShare: (https://figshare.com/s/0f50e0d44f8bec621d10). The raw BUSCO matrix included 2556 single-copy genes, including introns and small portions of upstream and downstream regions, with a median protein length ranging from 77 to 4912 aa (mean = 629±432.6aa).

The raw alignment had a total size of 10.04 Mbp. Informative and coding sites represented 34.8 and 48.7% of the total length, respectively. After filtering the alignment with Gblocks, the supermatrix consisted of 4,506,888 sites, including 59.4% informative and 94.1% coding sites. The number of unambiguous nucleotide positions in each dataset, variable sites and the best fitting models of evolution selected in PartitionFinder, and the average bootstrap tree length, average branch length, variance in branch length, slope, R-Square, percentage of constant sites and number of taxa in each data set are summarized in Table 2.3 (Stored in: https://figshare.com/s/0f50e0d44f8bec621d10).

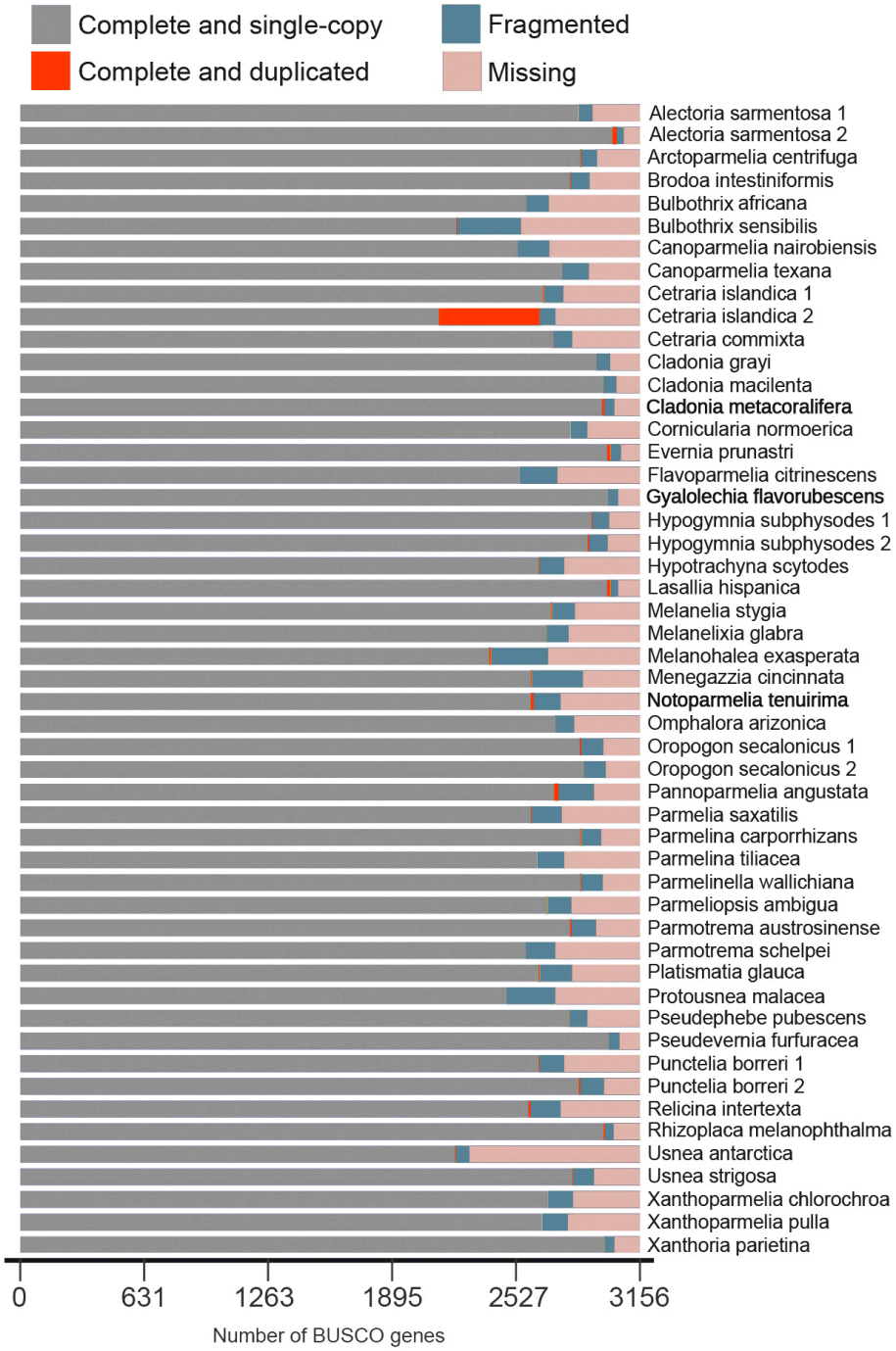


Figure 2.1

Genomic quality assessment of the 51 genomes of lichen-forming fungi used in this study. The plot shows the number ofBUSCO single-copy genes, which are "complete", "missing","duplicated", and "fragmented" in each genome

Phylogenomic inferences

Phylogenies inferred from the concatenated phylogenomic dataset and the species tree evaluations recovered highly similar topologies, with few exceptions (Fig. 2.2 a, b). We only consider clades receiving bootstrap support of at least 90% and multilocus-bootstrap of at least 0.98 as strongly supported and focus on these well-supported relationships here. The monophyly of the six major clades of Parmeliaceae represented in this study was strongly supported, with each major clade recovered with 100% bootstrap support and 1.0 multilocus-bootstrap (MLBP). Relationships among these major clades generally received strong support in our phylogenomic analyses (Fig. 2.2). A comparison of phylogenetic relations and node support between a previously published multi-gene study and the current phylogenomic study is depicted in Table 2.1. The sister-groups of the usneoid and alectorioid clades were unresolved in a previously published multi-gene study, and the genus *Menegazzia* had a well-sup-ported sister-group relationship with *Coelopogon* (not included in this study). However, this clade had no strongly supported close relative. Our study was not entirely conclusive in this regard. While in the species tree analysis *Menegazzia* grouped within the Alectorioid clade (Fig. 2.2b), in the concatenated phylogenomic dataset analysis the genus *Menegazzia* represented the earliest diverging group in the family, followed by the alectorioid clade (Fig. 2.2a). In both phylogenomic analyses, the alectorioid clade was strongly supported as sister-group to all remaining clades. The remaining clades were divided into two major sister-groups, (i) *Evernia* + cetrarioid + hypogymnioid + usneoid + anzioid clades, and (ii) *Oropogon* + *Platismatia* + *Omphalora* + parmelioid clade. In the first major group, the usneoid clade was strongly supported (93% bootstrap support) as sister-group to the anzioid clade. This relation was, however, not supported in the species tree inference (0.55 MLBP). The anzioid and usneoid clades formed a strongly supported (93% bootstrap support, 1.0 MLBP) sister-group relationship with the cetrarioid + hypogymnioid group. Within this latter group, the sister-group relationship of *Evernia* with the cetrarioid clade, which lacked support in a previous study (Divakaret al. 2015), received strong support with both methods here (Fig.2.2). The clade including Evernia + the cetrarioid clade formed a strongly supported sister-group relationship with the hypogymnioid clade. Both the cetrarioid and the hypogymnioid clades have their centre of

distribution in the northern Hemisphere (Goward 1986; Hale 1986; Thell et al. 2009, 2012; Miadlikowska et al. 2011; Nelsen et al. 2011; Randlane et al. 2013; Egan 2016). In the second major group, the position of *Omphalora arizonica* was not clearly established. In the concatenated phylogenomic approach *Omphalora arizonica* was strongly supported as sister to the entire parmelioid clade, which represents the most speciose clade in the family. *Oropogon* and *Platismatia* were sister to the parmelioid + *Omphalora* group (Fig. 2.2a). In the species tree, instead, O. arizonica grouped without support (0.53 MLBP) with the *Oropogon* + *Platismatia* group, which formed a strongly supported sister-group relationship with the entire parmelioid clade (1.00 MLBP; Fig. 2.2b). Within the parmelioid clade, the position of *Parmeliopsis ambigua* deserves further investigation. Our taxon sampling is insufficient to discuss phylogenetic relationships within major clades since only two to four species were included, with the exception of the parmelioid clade. Interestingly, the relationships of previously recognized lineages within the parmelioid clade (Blanco et al. 2006; Crespo et al. 2010b) lacked support in our phylogenomic analyses. Whether this is due to a rapid radiation at the origin of the clade or can be addressed using a larger taxon sampling remains to be investigated.

Sister-group relations between major clades	The six genes phylogeny (Divakar et al. 2015)	The 2556 single-copy protein-coding genes phylogeny (current study)
Alectorioid and all remaining clades	Unresolved	Strongly supported
Parmelioid and Oropogon+Platismatia	Unsupported	Strongly supported
Anzioid and Usneoid	Unresolved	Weakly supported
Cetrarioid and Hypogymnioid	Unresolved	Strongly supported
Cetrarioid and Evernia	Unsupported	Strongly supported
Cetrarioid+Evernia and Hypogymnioid	Unresolved	Strongly supported
Cetrarioid+Evernia+Hypogymnioid and Anzioid+Usneoid	Unresolved	Strongly supported
Cetrarioid+Evernia+Hypogymnioid+ Anzioid+Usneoid and Parmelioid+ Oropogon+Platismatia+Omphalora	Unresolved	Strongly supported

Table 2.1.
Comparing phylogenetic relations between six genes phylogeny (Divakar et al. 2015) and 2556 single-copy protein-coding genes phylogeny (current study), among the major clades of subfamily Parmelioideae

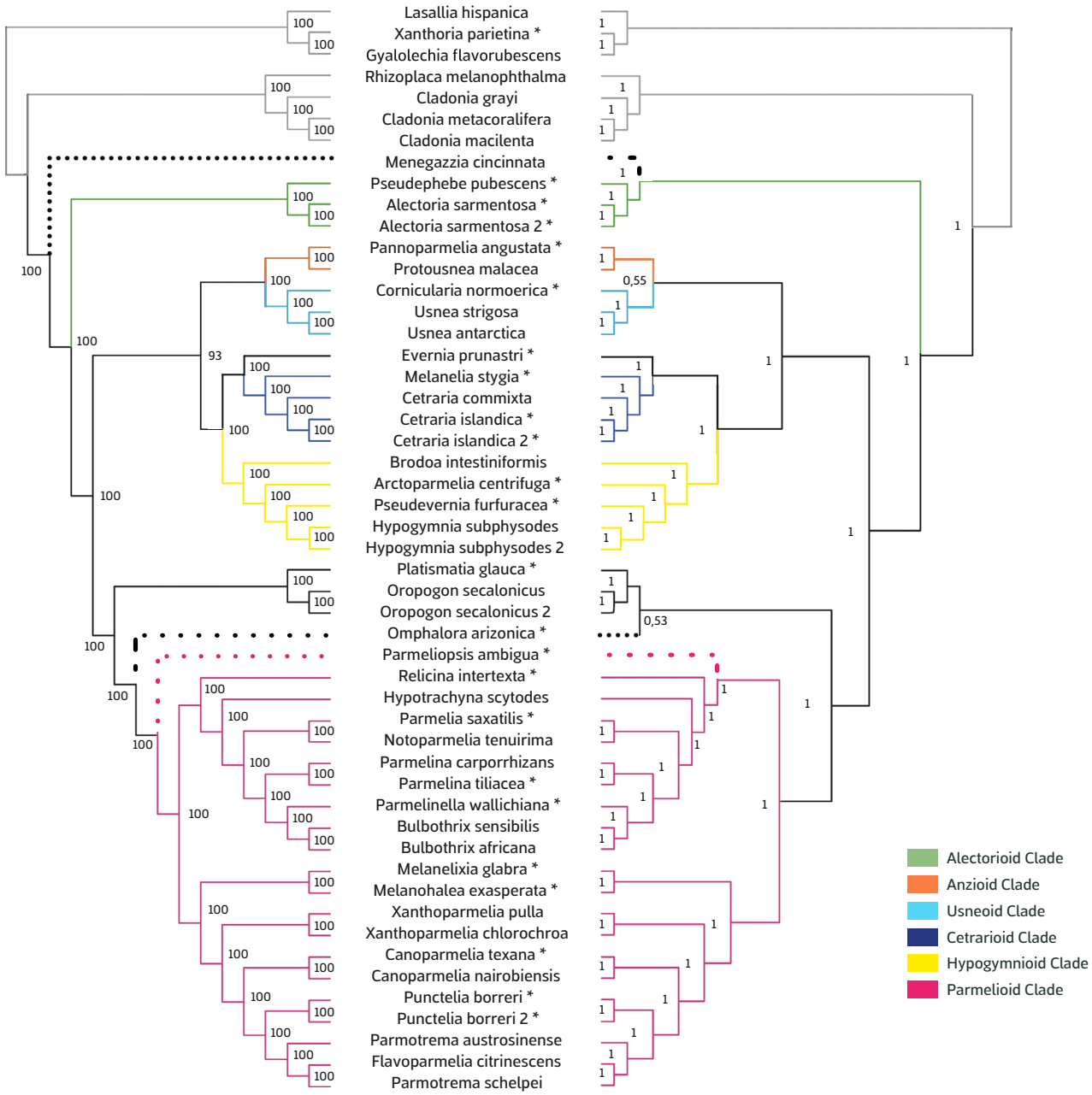


Figure 2.2

Evolutionary relationships of six major clades (represented by differently coloured branches) in the subfamily Parmelioideae (Parmeliaceae) based on aconcatenated Gblocks alignment of 2556 single-copy protein-coding genes. The tree on the left (a) is a ML tree with ML bootstrap values (1000 bootstrap pseudoreplicates) from a IQTree analysis added at nodes (Concatenation-based phylogeny). The tree on the right (b) is an ASTRAL species tree based on the ML inferred individual gene trees from the 51 taxa of the Parmelioideae (Coalescence-based phylogeny). Dot lines show conflicts between coalescence-based phylogeny and concatenation-based phylogeny. The type species of genera are marked with an asterisk.

Utility of metagenome skimming for phylogenomic analyses

The approach is suitable for extracting genome-wide phylogenetic markers. The genome completeness estimates of most of the reconstructed fungal genomes in our study are in line with reports from a previous study (Meiser et al. 2017)(Fig. 2.1). Furthermore, here we showed that the implemented metagenomic approach was able to recover a common set of single-copy protein coding genes in species with varying coverage depths. Per genome missing loci represented on average only 6.2 (±5.6) % of the total gene set used for phylogenomic inference. Even in reconstructed genomes with low coverage (e.g., 17-fold for *Pannoparmelia angustata*, *Flavoparmelia citrinescens, Notoparmelia tenuirima*) the percentage of missing loci did not exceed 7.2%, suggesting that the single-copy fraction of the fungal nuclear genome was readily recovered even at low sequencing depths. Further, an evaluation of the minimum number of targeted markers required to obtain a robust phylogeny is recommended for phylogenomics study [e.g. Rokas et al. 2003]. In a recent study, Ai and Kang (2015) developed a robust approach (including random sampling of genes starting from 25 with increments of 25, until the percentage of correct trees reached 100), based on 11 species with 830 single-copy nuclear genes to infer the need of minimum numbers of both random and selected genes to obtain a robust phylogeny. Given the large sizes of many of our subset data sets (i.e., computational limitations) and a desire to employ the same methodology when generating nodal support for all concatenated datasets analyses of the randomly sampled genes, we had little choice to use this methodology. Nonetheless, the single gene trees of the selected 2556 single-copy genes had an average bootstrap above 80% (Table 2.3), indicating the out-performance of these markers to resolve phylogenetic relations in Lecanoromycetes and in Ascomycota in general. Gene selection strategies based on single-gene tree performance are strongly recommended in phylogenomic analyses (Ai and Kang 2015). The comprehensive data set of 2556 homologous single-copy genes created in this study is a valuable genomic resource. It can, e.g., be used to identify new phylogenetic markers, extract optimal sets of phylogenetic markers, re-evaluate the resolution power of existing markers, or to identify the minimum number of loci or variable sites necessary to reconstruct well-resolved and highly supported phylogenies in

the Lecanoromycetes (and potentially other ascomycete classes). Furthermore, it may spur investigations into the evolution of functional genes and gene families in lichen-forming fungi. It is the first phylogenomic study addressing evolutionary relationships at the family level in lichen-forming fungi, focusing on the hyperdiverse family, Parmeliaceae (Kraichak et al. 2015). Extended taxon sampling of metagenomic data and additional phylogenomic approaches, such as RADseq and target enrichment of ultraconserved elements provide promising avenues for resolving problematic phylogenetic relationships in this species-rich fungal subfamily.

Table 2.2
Number of single-copy BUSCO genes recovered for each specimen used in the alignment

Species	Complet and single-copy BUSCOs (S)	S* genes used in the alignment (out of 2556 total)
Alectoria sarmentosa 1	2842	2478
Alectoria samentosa 2	3015	2524
Arctoparmelia centrifuga	2849	2497
Brodoa intestiniformis	2800	2484
Bulbothrix africana	2574	2356
Bulbothrix sensibilis	2223	2086
Canoparmelia nairobiensis	2535	2361
Canoparmelia texana	2758	2481
Cetraria commixta	2130	1829
Cetraria islandica 2	2712	2450
Cornicularia normoerica	2802	2451
Flavoparmelia citrinescens	2543	2372
Hypogymnia subphysodes 1	2908	2520
Hypogymnia subphysodes 2	2889	2504
Hypotrachyna scytodes	2639	2412
Melanelia stygia	2701	2448
Melanelixia glabra	2680	2436
Melanohalea exasperata	2387	2216
Menegazzia cincinnata	2599	2345
Notoparmelia tenuirima	2602	2393
Omphalora arizonica	2725	2448
Oropogon secalonicus 1	2847	2478
Oropogon secalonicus 2	2869	2493
Pannoparmelia angustata	2717	2396
Parmelia saxatilis	2602	2409
Parmelina tiliacea	2631	2411
Parmeliopsis ambigua	2851	2498
Parmelinella wallichiana	2683	2444
Parmotrema austrosinense	2799	2497
Parmotrema schelpei	2576	2379
Platismatia glauca	2641	2394
Protousnea magellanica	2476	2256
Pseudephebe pubescens	2792	2471
Punctelia borreri 1	2641	2425
Punctelia borreri 2	2843	2517
Relicina intertexta	2587	2370
Usnea antarctica	2213	1971
Usnea strigosa	2809	2497
Xanthoparmelia chlorochroa	2690	2456
Xanthoparmelia pulla	2656	2428

6.CAPÍTULO 3

ANÁLISIS DE GENOMAS COMPLETOS DESVELAN HETEROTALISMO EN LECANOROMYCETES

6.CHAPTER 3

GENOME-WIDE ANALYSIS UNCOVER HETEROTHALLISM IN LECANOROMYCETES

6

6.1 INTRODUCTION

Sexual reproduction in filamentous fungi is controlled by genes of the mating-type locus (MAT1) (Coppin et al. 1997; Kronstad and Staben 1997). This locus comprises two highly dissimilar allelic variants, the MAT1-1 and MAT1-2 idiomorphs (Metzenberg and Glass 1990). These variants encode highly divergent proteins: a region encoding an **α**1 domain characterizes the core *mat1-1* gene, while a MATA_HMG (high-mobility group)-box characterizes the core *mat1-2* gene (Turgeon and Yoder 2000). The transcription factors of the MATA_HMG domain are involved in sexual development and have been proposed to be the ancestral fungal sex determinant in fungi (Idnurm et al. 2008; Lee et al. 2008). The molecular function of the **α**-box is still unclear, although evidence suggests that it may act as transcriptional coactivator (Hagen et al. 1993).

Fungal mating systems can be classified based on the genic content of the MAT1 locus as, in general, sexual reproduction requires the expression of genes from both MAT1 idiomorphs (Ni et al. 2011; Dyer et al. 2016). Individuals of heterothallic (out-breeding) species possess genes from only one of the two idiomorphs. Individuals of heterothallic species are thus obligately out-crossing as they require a compatible partner for sexual reproduction to occur (Dyer 2008).

On the other hand, homothallism is an umbrella term that describes a variety of distinct mechanisms that collectively allow for single individuals to be self-fertile and may be classified as primary and secondary homothallism (Wilson et al. 2015). Primary (i.e., genetic) homothallic species possess genes of both MAT1-1 and MAT1-2 idiomorphs within a single genome. Secondary homothallism refers to other mechanisms that allow for homothallic behavior such as: 1) uni-/bidirectional mating-type switching when individuals of one or

both mating types are able to reversibly (or irreversibly) switch to the opposite mating type forming a mixed, functionally heterothallic colony; and 2) unisexuality when individuals of the same mating type are able to undergo sexual reproduction regardless of the absence of a compatible mating partner. Self-fertility can also be achieved via pseudohomothallism, when opposite mating-type nuclei are packed within a single spore which produces, upon germination, a heterokaryotic, self-fertile mycelium (Whitehouse 1949: Olive 1958; Nelson 1996; Yun et al. 1999; Whittle et al. 2011). All these different systems of secondary homothallism allow, from one side, the preservation of homothallic mating under conditions in which the compatible mating partner is absent or not easily accessible, while retaining the ability to outcross.

Lichen-forming fungi have two alternative reproductive strategies: asexual reproduction and sexual ascospore-producing reproduction. Asexual reproductive systems generally result in the simultaneous propagation of fungal and photosynthetic symbionts (Dal Grande et al. 2012, but see Wornik and Grube 2010), either in granules of varying size containing algal cells and fungal hyphae (soredia or isidia), which are easily detached outgrowths from the lichen thallus. Although asexual reproduction codisperses the fungal and photosynthetic partners, exclusive asexuality has often been interpreted as an evolutionary dead end (Normark et al. 2003). Sexual reproduction decouples the symbionts and the fungus must find a suitable photosynthetic partner for the establishment of a new lichen thallus. While the morphological underpinnings of the reproductive modes in lichens have been dissected in detail (Büdel and Scheidegger 2008), the genetic basis of sexual reproduction in lichens remain, however, largely unknown because of the failure to induce sexuality in vitro (Murtagh et al. 2000).

Primary homothallism is widespread among filamentous ascomycetes, where it is derived from heterothallic ancestors via genetic capture (Beukeboom and Perrin 2014). Self-fertility via homothallism has been proposed to be a prevalent characteristic of lichen-forming fungi (Murtagh et al. 2000), which represent about half of the known ascomycetes (reviewed in Hawksworth 2015).

Primary homothallism has been unequivocally demonstrated for only one lichen-forming fungus of the class Eurotiomycetes, *Endocarpon pusillum*, based

on results of whole-genome analysis (Wang et al. 2014). On the other hand, for the Lecanoromycetes, the largest class of lichenized fungi, reports of homothallism have mostly been based on indirect evidence, such as genetic uniformity based on RAPD-PCR fingerprinting of ascospores from the same ascomata (Murtagh et al. 2000; Seymour et al. 2005; Honegger et al. 2007). Obligate, behavioral homothallism has been reported for only two species of the order Teloschistales, *Xanthoria elegans*, and *X. parietina* (Scherrer et al. 2005). In the first case, both mating types were detected in all haploid, single spore isolates, although the exact mating-locus architecture has not been recovered. The latter, instead, represents the firstly reported example of a unisexual lichen species: although genetically heterothallic having lost MAT1-1, descendants of meiosis displayed no segregation at the mating locus. Results based on MAT1 sequencing via PCR amplification and population genetic data have unequivocally demonstrated genetic and behavioral heterothallism for several Lecanoromycetes species from different families representing various reproductive strategies (Ludwig et al. 2017; Tripp et al. 2017; Dal Grande et al. 2018), such as from predominantly sexually reproducing (e.g., *Parmelina carporrhizans*, Parmeliaceae) (Alors et al. 2017) to predominantly vegetatively reproducing species (*Lobaria pulmonaria*, Lobariaceae) (Singh et al. 2012, 2015). It has been hypothesized that homothallism might be widespread among lichen-forming fungi, especially in the order Lecanorales (Murtagh et al. 2000). It is still unclear, however, whether this reproductive mode represents the ancestral or derived state in lichenized ascomycetes.

In this study, we tested the hypothesis of widespread, ancestral primary homothallism in lichen-forming fungi using genomic data. For this purpose, we gathered whole-genome sequence data and characterized the MAT1 locus of a set of genomes of lichen-forming fungi representing a wide range of growth forms and reproductive strategies, with particular focus on the Lecanoromycetes, the largest clade of lichenized fungi. Our findings will contribute to the understanding of the regulation of reproductive processes and the evolution of the mating locus in the Lecanoromycetes. This will further contribute to our understanding of the mechanisms behind the accelerated diversification of this important and diverse group of symbiotic fungi.

6.2 MATERIALS AND METHODS

Taxon Sampling

A total of 41 lichen-forming fungal species were included in this study (Table 3.1). We included genomes of species belonging to different classes and orders of lichen-forming fungi. In the class Lecanoromycetes, we analyzed genomes of species belonging to the order Teloschistales (*Xanthoria parietina* and *Gyalolechia flavorubescens*), Umbilicariales (*Umbilicaria pustulata*, *Umbilicaria hispanica* and *Umbilicaria muehlenbergii*), Ostropales (*Graphis scripta*), Peltigerales (*Leptogium austroamericanun*), and Lecanorales (family Cladoniaceae: *Cladonia grayi*, *C. macilenta*, *C. metacorallifera*; Lecanoraceae: *Rhizoplaca melanophthalma*; Icmadophilaceae: *Dibaeis baeomyces*; Parmeliaceae: 27 species representing six of its seven major clades; see Table 3.1). Additionally, we included two species belonging to the sister class Eurotiomycetes (*Endocarpon pusillum)* and Arthoniomycetes (*Arthonia rubrocinta*).

Trimming and Assembly of New Genomes

The same procedure described in the chapter 1 for Parmeliaceae genomes was carried out for genomes retrieved from NCBI, that is, *Arthonia rubrocincta* (PRJNA256244) and *Graphis scripta* (PRJNA256475), and the metagenomes of *Leptogium austroamericanum* (PRJNA256476) and *Dibaeis baeomyces* (PRJNA256246). Trimmed paired-end reads were assembled using SPAdes or MetaSPAdes (Nurk et al. 2017), depending on the type of data, using default parameters.

Ortholog Identification and Tree Reconstruction

To infer the phylogenetic placement of the class Lecaronomycetes within the Pezizomycotina, we selected 53 genomes representing the major groups of this subphylum (Table 3.2). Orthologs genes were recovered using the CEGMA pipeline (Parra et al. 2007). Every genome was explored using a data set containing 458 proteins of Core Eukaryote Genes. The complete CEGMA genes predicted in each genome were extracted and aligned using MAFFT L-INS-i (Standley 2013). A supermatrix was created by concatenating all alignments using FASconCAT.pl (Kück and Longo 2014). Then, in order to optimize information content and data saturation we used MARE (Misof et al. 2013) with iterative steps of gene exclusion, resulting in an optimal subset of 81 genes. Evolutionary relationships were inferred from this subset using ML analysis as implemented in IQTree v1.5.5 with standard model selection (Nguyen et al. 2015). For each analysis, 1,000 bootstrap replicates were calculated using fast bootstrapping option. The resulting tree was drawn using FigTree v 1.3.1 (Rambaut 2009).

We followed a similar procedure to infer the phylogenetic relationships among 39 lichen-forming fungi belonging to the class Lecanoromycetes. For this purpose, we extended the orthologs gene set to 3,156 single-copy genes of Pezizomycotina as implemented in BUSCO v3 (Simão et al. 2015), resulting in a final matrix of 735 genes. All subsequent analyses were carried out as outlined earlier. The genomes of *Endocarpon pusillum* (Eurotiomycetes) and *Arthonia rubrocinta* (Arthoniomycetes) were used as outgroup.

Mating-Type Locus Identification

In order to identify the mating-type locus in every genome, we first selected protein sequences of SLA2 of *Xanthoria polycarpa* (CAI59767.1), APN2 of *Xanthoria parietina* (CAI59775.1), Alpha-domain of MAT1-1 of *Xanthoria polycarpa* (CAI59771.1), and HMG-domain of MAT1-2 of *Dufourea flammea* (CAI59780.2)

from Scherrer et al. (2005). These proteins and sequence domains were used as queries in tBLASTn (Altschul et al. 1990) searches against every genome assembly. Scaffolds containing more than one sequence query were retrieved and gene prediction was carried out using MAKER2 (Holt and Yandell 2011). Proteins derived from gene prediction were annotated comparing them with KEGG (Kanehisa et al. 2016) and COG databases (Tatusov 2000). Only scaffolds with complete mating-type loci, that is, containing complete anchoring genes (SLA2, APN2) were included in further analyses.

Comparison between Two Different Mating-Type Loci

Two loci of opposite mating type from *Umbilicaria pustulata* (Dal Grande et al. 2017; Dal Grande et al. 2018) were aligned using LASTZ (Harris 2007). Sequence conservation and visualization were carried out using Zpicture (Ovcharenko et al. 2004). Regions with >90% of ECR similarity were retrieved. We further extracted the intergenic regions between mating-type genes and the flanking genes of the two loci using BEDtools (Quinlan and Hall 2010). Intergenic flanking regions were aligned with nucleotide sequence of the opposite mating-type locus using MAFFT.

6.3 RESULTS AND DISCUSSION

While a solid foundation on morphological and anatomical understanding of the reproductive modes in lichens have been presented in great detail (Büdel and Scheidegger 2008), the genetic basis of sexual reproduction in lichens remain largely unknown. Here, we analyzed the mating-type locus from 41 genomes representing 4 classes and 9 orders of lichenized fungi. Our results show the loss of primary homothallism in the Lecanoromycetes, the largest group of lichenized fungi (Fig. 3.1). In all 39 Lecanoromycete genomes, we found the same organization of the mating locus, with a single MAT1 core gene, *mat1-1* or *mat1-2*, flanked by the highly conserved cytoskeleton protein (*sla2*)

and DNA lyase (*apn1*) genes (Fig. 3. 2 and Table 3.1). This is in accord with studies reporting heterothallic organization for several species in this group. Sequences flanking the core MAT1 genes were rather conserved and the transition between similar/dissimilar regions in both idiomorphs was abrupt (Fig. 3.2). This is similar to the domain organization found in the heterothallic ascomycete *Cochliobolus heterostrophus* (Turgeon et al. 1993). We found a novel gene between *mat1-1* and *sla2* in every species containing a MAT1-1 idiomorph. Within MAT1-2 loci, a different auxiliary gene was detected between *mat1-2* genes and *apn2* in every MAT1-2 species, with the exception of *Graphis scripta*. Preliminary phylogenetic analysis suggests that these genes might be lichen-specific (Fig. 3.5): the auxiliary *mat1-1* genes of Lecanoromycetes clustered together forming an unsupported monophyletic group with an unsupported sister relationship to the *mat1-1-4* gene cluster, the latter commonly found in Eurotiomycetes, for example in *Paracoccioides brasiliensis* (Desjardins et al.

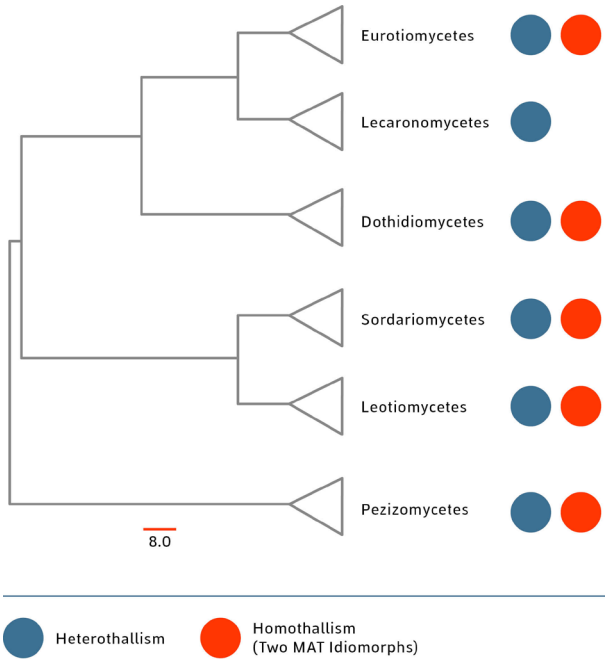


Figure 3.1

Phylogenetic placement of the class Lecanoromycetes within the Pezizomycotina. This is a phylogenetic tree from an IQTree analysis based on a concatenated alignment of 81 CEGMA genes. The tree includes 53 species representing the major groups of Pezizomycotina. The nodes were collapsed at the class rank for clarity of presentation. On the right, gray and white circles represent heterothallic and homothallic organization, respectively. All nodes received maximum ML bootstrap support values (100%).

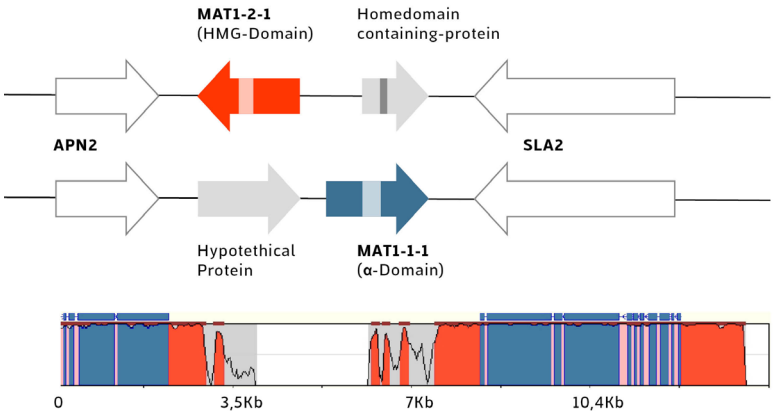
2011); similarly, the auxiliary *mat1-2* genes also formed an unsupported monophyletic group clustering all Lecanoromycete species, except *Xanthoria parietina*, which clustered with *Aspergillus mat1-2-4*. Sequence analysis showed high divergence among Lecanoromycetes species; in addition, a different number of introns were found depending on the species (Table 3.3). Interestingly, in some species, for example, *Cetraria islandica*, *Cetraria commixta*,

and *Umbilicaria pustulata* (Dal Grande et al. 2018), the *mat1-2* auxiliary genes displayed a conserved homeodomain leucine-zipper domain (pfam11569, Fig.3.4). The characterization of these new genes warrants further study.

Evolutionary transitions between homothallism and heterothallism have commonly occurred in both directions throughout the fungal kingdom (Lin and Heitman 2007). These transitions between inbreeding and outcrossing are likely a response to biological and/or environmental cues that favor one or the other strategy. It has been suggested that homothallism represents the primary reproductive mode of lichen-forming fungi (Murtagh et al. 2000). The key argument in support of this hypothesis was that, based on initial analyses of

Figure 3.2

Top: Schematic representation of the organization of the two types of mating alleles in Lasallia pustulata (Umbilicariaceae, Lecanoromycetes). Bottom: Whole-locus alignment of the two mating-type alleles using LASTZ. Red blocks indicate highly conserved regions between the loci; coding parts are highlighted in blue



mat gene organization in lichens, homothallism was presumed to be widespread. Our findings suggest the opposite, that is, that heterothallism is the prevalent organization of the MAT1 locus across the wide taxonomic diversity of the Lecanoromycetes under study, including several supposedly asexual lichens. Furthermore, our results indicate a highly conserved organization and synteny of the MAT1 locus in lichens. By and large, our results strongly support the hypothesis of an ancestral heterothallic state in lichens. This scenario is thus similar to what has been described for the evolution of breeding systems in other ascomycete genera such as *Cochliobolus* (Yun et al. 1999) and, more recently, *Aspergillus* (de Vries et al. 2017; Ojeda-Lopez et al. 2018).

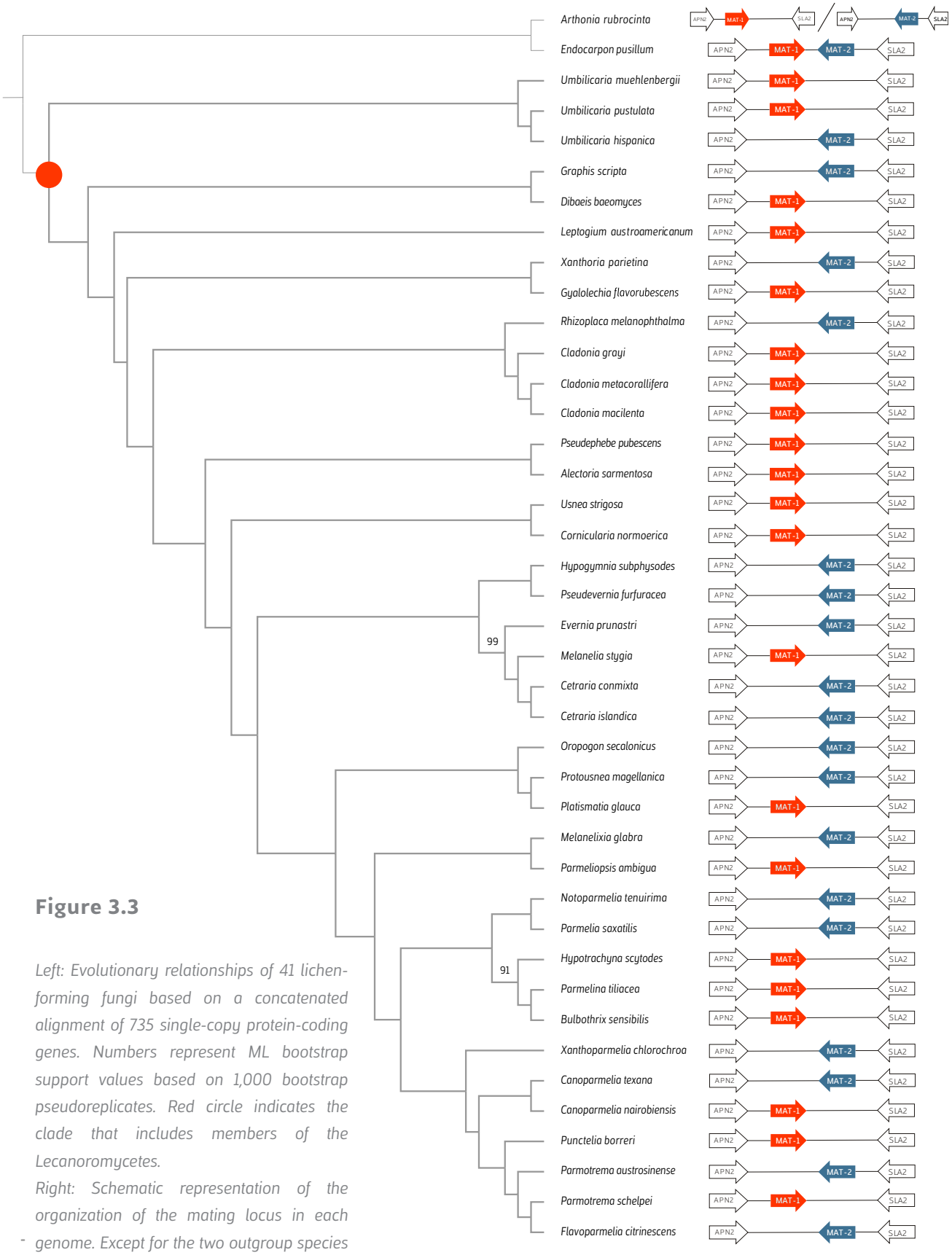


Figure 3.3

Left: Evolutionary relationships of 41 lichen-forming fungi based on a concatenated alignment of 735 single-copy protein-coding genes. Numbers represent ML bootstrap support values based on 1,000 bootstrap pseudoreplicates. Red circle indicates the clade that includes members of the Lecanoromycetes.

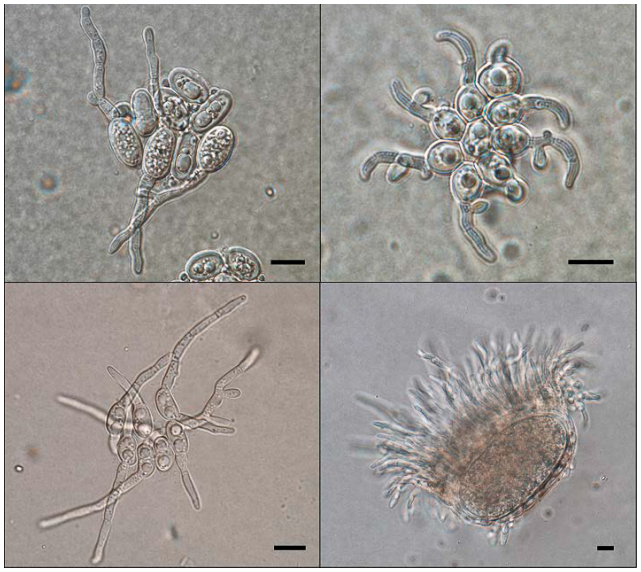
Right: Schematic representation of the organization of the mating locus in each genome. Except for the two outgroup species that showed homothallic organization either on the same (Endocarpon pusillum) or different scaffolds (Arthonia rubrocinta), all remaining genomes (i.e., Lecanoromycetes) displayed heterothallic organization

The lack of primary homothallism in the Lecanoromycetes, the most phenotypically diverse class of lichenized fungi, is somewhat surprising. This may well be influenced by the symbiotic lifestyle of these fungi. On the one hand, from the fungal perspective, heterothallism, or obligate outcrossing, can be considered as a high-risk, high-reward strategy. Some portion of the population may, in fact, not be able to find a compatible mating partner, especially for those species with skewed distributions of the MAT idiomorphs. On the other hand, the progeny of successful matings will have higher genetic diversity (Otto 2008). Compared with homothallic systems, outbreeding fungi may display accelerated adaptive evolution and more efficient elimination of deleterious mutations, thus they might be more able to avoid Muller's Ratchet (Roach and Heitman 2014). This is particularly true in environments with more novel factors (Murtagh et al. 2000). It is thus tempting to speculate that the tendency to engage in more prominent outbreeding might be responsible for the accelerated diversification found in this fungal clade (Gaya et al. 2015; J.P. Huang et al., submitted).

From the perspective of the lichen holobiont, sexual reproduction allows for the possibility of reshuffling of the symbionts to generate novel fungus–alga pairs (Dal Grande et al. 2012). As shown recently, these new associations may be key to expanding a lichen's niche (Rolshausen et al. 2017). On the other hand, the absence of compatible mating partners in the population in case of obligatory outcrossing lichen-forming fungi would comport the risk of being stuck with suboptimal or maladapted photobionts. In this respect, the widespread heterothallism in the Lecanoromycetes would still be advantageous in the presence of mechanisms that would reduce the cost of sex and avoid the problem of mate finding. Results from literature and our own ongoing research seem to support this scenario.

First, there have been several reports of population and seasonal effects on ascospore discharge and germination in this fungal group. In this regard, for example, seasonality was shown to be the regulating factor in *Cladonia furcata* (Jahns et al. 1979) and a few species of the Parmeliaceae (Ruibal, personal communication), thallus size in *Umbilicaria pustulata* (Hestmark 1992), *Xanthoparmelia cumberlandia* (Pringle et al. 2003), and thallus age in *Parmelia sulcata* (Honegger et al. 2007). This would mean that, like in many other fungi, the timing of sexual reproduction in lichens could be adjusted to when the

Figure 3.4



Top: Simultaneous, unipolar germination of spores ejected from one ascus in Melanelixia glabra (Parmeliaceae, Lecanoromycetes; left) and Cetraria sepincola (Parmeliaceae, Lecanoromycetes; right) at 9 and 8 days after ejection, respectively. Bottom: bipolar (Xanthoparmelia stenophylla, Parmeliaceae, Lecanoromycetes; left) and multipolar (Menegazzia cincinnata, Parmeliaceae, Lecanoromycetes; right) spore germination at 7 and 14 days after ejection, respectively. Ascospore isolation and germination followed the method by Molina and Crespo (2000). Scale bars represent 10 µm.

costs are lowest (Lee et al. 2010; Stelzer 2015). The reproductive strategy being selected for a particular species would therefore depend on interactions among many factors, either environmental (e.g., nutrient availability, competition for space and/or photobiont pools) and/or biological (e.g., population structure, thallus age, and size).

Second, data suggest that many species in this group may be secondarily homothallic. Secondary homothallism in the Lecanoromycetes consists of different strategies or a combination of them, such as unisexuality, the formation of heterokaryotic, self-fertile thalli, and pseudohomothallism. Scherrer et al. (2005) demonstrated unisexuality in the invariably fertile *Xanthoria parietina* as, although being genetically heterothallic, all descendants of meiosis contained only *mat1-2*. A unisexual cycle may be essential for lineage survival when conditions are unfavorable for heterosexual mating or compatible mating-type partners are not available. This is, for instance, the case for species of the genus *Cryptococcus* that are able to produce spores only via a unisexual or heterosexual cycle (Billiard et al. 2012; Fu et al. 2015). The presence of unisexual mating indicates that in certain ecological niches (e.g., for ruderal species and pioneer colonizers) there may be strong evolutionary pressure for homothallism to arise as the derived state. Unisexual reproduction utilizes a similar genetic pathway as heterosexual reproduction (Feretzaki and Heitman 2013). As such, unisexually derived meiotic spores carry clear advantages over clones or mitotic spores (conidia) in terms of survival rates,

especially in adverse environmental conditions (Trapero-Casas and Kaiser 2007). Another advantage of the maintenance of sex via unisexuality may be the reduction of the number of transposons in the genome via increased selection (Roach et al. 2014). The formation of heterokaryotic, self-fertile thalli may be achieved via the joint dispersal and germination of ascospores from the same ascus. The joint ejection and germination of ascospores from *Xanthoria* species, typically early colonizers, was reported by Molina and Crespo (2000) and Honegger et al. (2004). The authors reported that, after only two days from ejection, a mucilage of unknown origins would glue the germinating spores together. We observed a similar phenomenon in several species of the Parmeliaceae (fig. 3.4, top). This suggests that lichen thalli of the Lecanoromycetes might be often composed of mycelia of opposing mating types, thus being de facto self-fertile. This would also mean that lichen thalli would comprise multiple mycobiont genomes, thus rendering metagenomic assemblies further challenging (Meiser et al. 2017; Tripp et al. 2017). The fact that we could retrieve a single, complete mating-type locus in all analyzed metagenomes, could be explained by a skewed mating-type ratio in the thallus portion that was used for DNA extraction. In pseudohomothallism, self-fertility is the result of the packaging of opposite mating-type nuclei within a single spore (Wilson et al. 2015). Although pseudohomothallism in lichen-forming fungi has not yet been reported, we observed bi- or multipolar germination of ascospores in members of the Parmeliaceae (Lecanoromycetes), suggesting the presence of multinucleated spores (fig. 3.4, bottom) (Molina and Crespo 2000). Multinucleate ascospores are a common feature in the order Pertusariales (Lecanoromycetes) (Pyatt 1968). Future studies should thus focus on characterizing the ascospore mating type in these species.

6.4 CONCLUSIONS

This is the first broad scale study dissecting the genetic architecture of the mating locus in lichen-forming fungi. We characterized the MAT1 locus in the genomes of several lichen-forming fungal species representing a wide range of growth forms and reproductive strategies (isidia, soralia, and ascospores). Noteworthy, we showed widespread heterothallism in the largest, phenotypically most heterogeneous group of lichen-forming fungi. We hypothesize that this is related to the symbiotic lifestyle of this fungal group. Furthermore, the consistency of this character allows us to speculate that this may be implicated in the accelerated diversification rates found in the Lecanoromycetes (Huang et al. 2019). As such, our study sets the stage for further exploration of the reproductive strategy of lichens, as well as of its evolutionary outcomes.

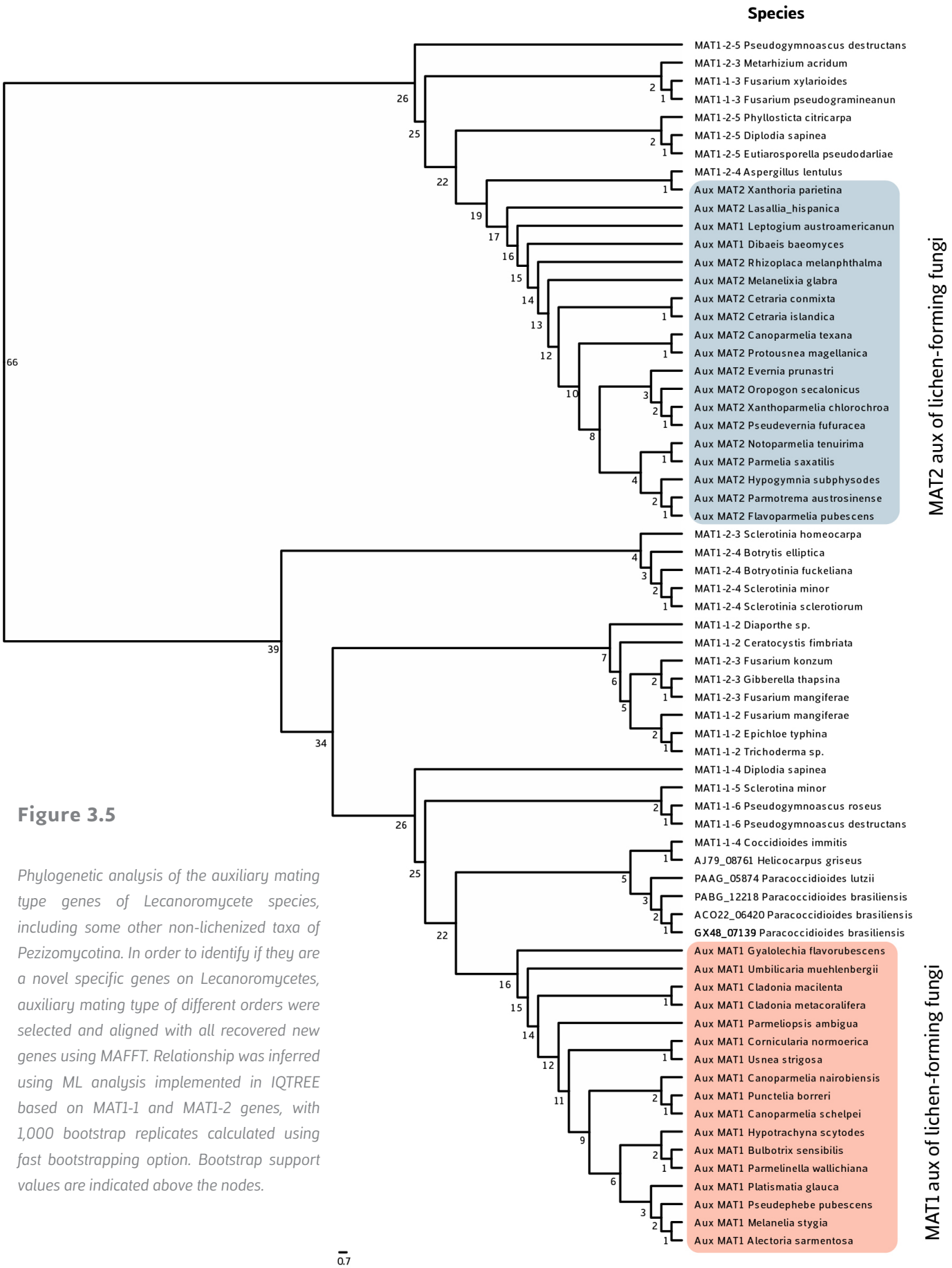


Figure 3.5

Phylogenetic analysis of the auxiliary mating type genes of Lecanoromycete species, including some other non-lichenized taxa of Pezizomycotina. In order to identify if they are a novel specific genes on Lecanoromycetes, auxiliary mating type of different orders were selected and aligned with all recovered new genes using MAFFT. Relationship was inferred using ML analysis implemented in IQTREE based on MAT1-1 and MAT1-2 genes, with 1,000 bootstrap replicates calculated using fast bootstrapping option. Bootstrap support values are indicated above the nodes.

Table 3.1 (part 1)
List of genomes used in this study and their characteristics: reproductive strategy of the species, presence of homeodomain-containing protein in the mating locus, and taxonomy.

Species	Mating Type	Auxiliary Gene MAT1-1	Auxiliary Gene MAT1-2	Reproductive strategy
Alectoria sarmentosa	MAT1-1	present	-	pred. Vegetative
Arthonia rubrocinta	MAT1-1/MAT1-2	absent	-	pred. Sexual
Bulbotrix sensibilis	MAT1-1	present	-	pred. Sexual
Canoparmelia nairobiensis	MAT1-1	present	-	pred. Sexual
Canoparmelia schelpei	MAT1-1	present	-	pred. Sexual
Canoparmelia texana	MAT1-2	-	present	pred. Vegetative
Cetraria conmixta	MAT1-2	-	present	pred. Sexual
Cetraria islandica	MAT1-2	-	present	pred. Sexual
Cladonia grayi	MAT1-1	present	-	pred. Vegetative
Cladonia macilenta	MAT1-1	present	-	pred. Vegetative
Cladonia metacorallifera	MAT1-1	present	-	pred. Sexual
Cornicularia normoerica	MAT1-1	present	-	pred. Sexual
Dibaeis baeomyces	MAT1-1	present	-	pred. Sexual
Endocarpon pusillum	MAT1-1-MAT1-2	-	-	pred. Sexual
Evernia prunastri	MAT1-2	-	present	pred. Vegetative
Flavoparmelia citrinescens	MAT1-2	-	present	pred. Sexual
Graphis scripta	MAT1-2	-	absent	pred. Sexual
Gyalolechia flavorubescens	MAT1-1	present	-	pred. Sexual
Hypogymnia subphysodes	MAT1-2	-	present	pred. Vegetative
Hypotrachyna scytodes	MAT1-1	present	-	pred. Sexual
Lasallia hispanica	MAT1-2	-	present	pred. Sexual
Lasallia pustulata	MAT1-1	present	-	pred. Vegetative
Leptogium austroamericanun	MAT1-1	present	-	pred. Vegetative
Melanelia stygia	MAT1-1	present	-	pred. Sexual
Melanelixia glabra	MAT1-2	-	present	pred. Sexual
Notoparmelia tenuirima	MAT1-2	-	present	pred. Sexual
Oropogon secalonicus	MAT1-2	-	present	pred. Sexual
Parmelia saxatilis	MAT1-2	-	present	pred. Vegetative
Parmelinella wallichiana	MAT1-1	present	-	pred. Vegetative
Parmeliopsis ambigua	MAT1-1	present	-	pred. Vegetative
Parmotrema austrosinense	MAT1-2	-	present	pred. Sexual
Platismatia glauca	MAT1-1	present	-	pred. Vegetative
Protousnea magellanica	MAT1-2	-	present	pred. Sexual
Pseudephebe pubescens	MAT1-1	present	-	pred. Sexual
Pseudevernia furfuracea	MAT1-2	-	present	pred. Vegetative
Punctelia borreri	MAT1-1	present	-	pred. Vegetative
Rhizoplaca melanophthalma	MAT1-2	-	present	pred. Sexual
Umbilicaria muehlenbergii	MAT1-1	present	-	pred. Sexual
Usnea strigosa	MAT1-1	present	-	pred. Sexual
Xanthoparmelia chlorochroa	MAT1-2	-	present	pred. Sexual
Xanthoria parietina	MAT1-2	-	present	pred. Sexual

Table 3.1 (part 2)

scaffold1	scaffold2	class	order	family	Reference
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete MAT1-1	complete MAT1-2	Arthoniomycetes	Arthoniales	Arthoniaceae	McDonald et al. 2013
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Cladoniaceae	Armaleo et al. Unpublished
complete		Lecanoromycetes	Lecanorales	Cladoniaceae	Park et al. 2013
complete		Lecanoromycetes	Lecanorales	Cladoniaceae	Park et al. 2014A
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Icmadophilaceae	McDonald et al. 2013
complete		Eurotiomycetes	Verrucariales	Verrucariaceae	Wang et al. 2014
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	Meiser et al. 2017
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Ostropales	Graphidaceae	
complete		Lecanoromycetes	Teloschistales	Teloschistaceae	McDonald et al. 2013
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Umbilicariales	Umbilicariaceae	Dal Grande et al. 2018
complete		Lecanoromycetes	Umbilicariales	Umbilicariaceae	Dal Grande et al. 2017
complete		Lecanoromycetes	Peltigerales	Collemataceae	McDonald et al. 2013
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	Meiser et al. 2017
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Lecanoraceae	
complete		Lecanoromycetes	Umbilicariales	Umbilicariaceae	Park et al. 2014B
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Lecanorales	Parmeliaceae	
complete		Lecanoromycetes	Teloschistales	Teloschistaceae	Dyer et al. Unpublished

Table 3.2
Accession numbers of genomes used in the phylogeny shown in Fig. 3.1

Species	Acc. Number
Acephala macrosclerotiorum	PRJNA334316
Alternaria alternata	PRJNA371205
Alternaria brassicicola	PRJNA34523
Ascodesmis nigricans	PRJNA247594
Aspergillus clavatus	PRJNA15664
Aspergillus flavus	PRJNA164607
Aspergillus fumigatus	PRJNA18733
Aspergillus wentii	PRJNA170886
Aspergillus zonatus	PRJNA207688
Aureobasidium pullulans	PRJNA207874
Blastomyces dermatitidis	PRJNA41099
Bulgaria inquinans	PRJNA333781
Chaetomium globosum	PRJNA12795
Chalara longipes	PRJNA213334
Cladonia grayi	JGI/Cgr/DA2myc/ss v2.0
Cladonia macilenta	PRJNA210603
Cladonia metacoralifera	PRJNA219240
Coccomyces strobi	PRJNA340577
Cochliobolus heterostrophus	PRJNA83117
Cochliobolus heterostrophusC5	PRJNA42739
Cochliobolus victoriae	PRJNA160941
Cordyceps militaris	PRJNA41129
Corynespora cassiicola	PRJNA234811
Dothistroma septosporum	PRJNA74753
Endocarpon pusillum	PRJNA181958
Evernia prunastri	NKYR00000000
Fusarium graminearum	PRJNA295638
Fusarium oxysporum	PRJNA347190
Gaeumannomyces graminis	PRJNA37931
Glonium stellatum	PRJNA295956
Grosmannia clavigera	PRJNA39837
Lasallia pustulata	PRJNA275184
Loramyces juncicola	PRJNA334368
Melanospora tiffanyae	PRJNA265602
Meliniomyces bicolor	PRJNA196026
Monascus ruber	PRJNA196033
Mycosphaerella fijiensis	PRJNA337612
Mycosphaerella graminicola	PRJNA19047
Neurospora crassa	PRJNA250607
Neurospora discreta matA	PRJNA207861
Neurospora tetrasperma matA	PRJNA65453
Neurospora tetrasperma mata	PRJNA65453
Parmelina carporhrizans	Unpublished
Pseudevernia furfuracea	NKYQ00000000
Pseudographis elatina	JGI/Pseel1/Pseel1
Pyrenochaeta sp	PRJNA255591
Rhizoplaca melanophthalma	LMCC0506 (GZU)
Spathularia flavida	JGI/Spafl1/Spafl1
Thielavia terrestris	PRJNA32847
Trypethelium eluteriae	JGI/Tryvi1/Tryvi1
Umbilicaria muehlenbergii	JFDN00000000
Uncinocarpus reesii	PRJNA15634
Xanthoria parietina	PRJNA62697

Table 3.3
Uniprot entries of auxiliary mating type genes of the species belonging to different orders of Pezizomycotina.

Species	auxiliary MAT gene	Gene length	Number of introns
Alectoria sarmentosa	Aux_MAT1-1	1075	3
Bulbotrix sensibilis	Aux_MAT1-1	1074	3
Canoparmelia nairobiensis	Aux_MAT1-1	1334	5
Canoparmelia schelpei	Aux_MAT1-1	2974	5
Canoparmelia texana	Aux_MAT1-2	696	2
Cetraria commixta	Aux_MAT1-2	771	2
Cetraria islandica	Aux_MAT1-2	771	2
Cladonia grayi	Aux_MAT1-1	1384	4
Cladonia macilenta	Aux_MAT1-1	1247	5
Cladonia metacorallifera	Aux_MAT1-1	1242	5
Cornicularia normoerica	Aux_MAT1-1	1189	4
Dibaeis baeomyces	Aux_MAT1-1	1050	2
Evernia prunastri	Aux_MAT1-2	1617	2
Flavoparmelia citrinescens	Aux_MAT1-2	2520	2
Gyalolechia flavorubescens	Aux_MAT1-1	1316	4
Hypogymnia subphysodes	Aux_MAT1-2	705	2
Hypotrachyna scytodes	Aux_MAT1-1	781	1
Lasallia hispanica	Aux_MAT1-2	837	2
Lasallia pustulata	Aux_MAT1-1	681	2
Leptogium austroamericanun	Aux_MAT1-1	2322	3
Melanelia stygia	Aux_MAT1-1	1203	3
Melanelixia glabra	Aux_MAT1-2	684	2
Notoparmelia tenuirima	Aux_MAT1-2	576	2
Oropogon secalonicus	Aux_MAT1-2	2213	2
Parmelia saxatilis	Aux_MAT1-2	776	2
Parmelinella wallichiana	Aux_MAT1-1	1073	2
Parmeliopsis ambigua	Aux_MAT1-1	1178	4
Parmotrema austrosinense	Aux_MAT1-2	695	2
Platismatia glauca	Aux_MAT1-1	1130	3
Protousnea magellanica	Aux_MAT1-2	777	2
Pseudephebe pubescens	Aux_MAT1-1	1055	3
Pseudevernia furfuracea	Aux_MAT1-2	1617	2
Punctelia borreri	Aux_MAT1-1	1355	6
Rhizoplaca melanophthalma	Aux_MAT1-2	798	2
Umbilicaria muehlenbergii	Aux_MAT1-1	1255	7
Usnea strigosa	Aux_MAT1-1	1198	4
Xanthoparmelia chlorochroa	Aux_MAT1-2	1625	2
Xanthoria parietina	Aux_MAT1-2	1179	2

Table 3.4
List of predicted Auxiliary mating type of Lecanomycetes, their gene length and number of introns.

Species	Gene name	Uniprot Entry
Aspergillus lentulus	MAT1-2-4	S5RDM3
Botryotinia fuckeliana	MAT1-2-4	A0A023PIK2
Botrytis_elliptica	MAT1-2-4	A0A0E3TJ69
Ceratocystis fimbriata	MAT1-1-2	X4YZ77
Coccidioides immitis	MAT1-1-4	A7KPA2
Diaporthe sp.	MAT1-1-2	Q1MX51
Diplodia sapinea	MAT1-1-4	V5NSW1
Diplodia sapinea	MAT1-2-5	V5NSE7
Epichloe typhina	MAT1-1-2	Q5TL91
Eutiarosporella pseudodarliae	MAT1-2-5	A0A2D1GT40
Fusarium konzum	MAT1-2-3	G3G2G0
Fusarium mangiferae	MAT1-1-2	G3G2G2
Fusarium mangiferae	MAT1-2-3	G3G2G5
Fusarium xylarioides	MAT1-1-3	B5WWM1
Fusarium_pseudograminearum	MAT1-1-3	K3VZ10
Gibberella thapsina	MAT1-2-3	G3G2E5
Helicocarpus griseus UAMH5409	AJ79_08761	A0A2B7WQ45
Metarhizium acridum	MAT1-2-3	A0A0P0BZ46
Paracoccidioides brasiliensis	GX48_07139	A0A1E2XVU7
Paracoccidioides brasiliensis	ACO22_06420	A0A1D2J7I2
Paracoccidioides lutzii Pb01	PAAG_05874	C1H533
Paracoccidioides lutzii Pb03	PABG_12218	A0A0A5IEQ3
Phyllosticta citricarpa	MAT1-2-5	A0A1S5QPD1
Pseudogymnoascus destructans	MAT1-1-6	A0A0D3MF52
Pseudogymnoascus destructans	MAT1-2-5	A0A0D3MES0
Pseudogymnoascus roseus	MAT1-1-6	A0A0D3MEY7
Sclerotina_minor	MAT1-1-5	A0A0A0P4J4
Sclerotinia homoeocarpa	MAT1-2-3	H8XXY1
Sclerotinia minor	MAT1-2-4	A0A0A0P2D1
Sclerotinia_sclerotiorum	MAT1-2-4	M1GE12
Trichoderma sp	MAT1-1-2	D2W8Q9

7.CAPÍTULO 4

ANÁLISIS DE CLÚSTERES DE GENES BIOSINTÉTICOS REVELAN PÉRDIDA CORRELATIVA DE GENES CON LA AUSENCIA DE ÁCIDO ÚSNICO EN LÍQUENES

7.CHAPTER 4

GENOME-WIDE ANALYSIS OF BIOSYNTHETIC GENE CLUSTER REVEALS CORRELATED GENE LOSS WITH ABSENCE OF USNIC ACID IN THE LARGEST CLADE OF LICHEN FORMING FUNGI

7

7.1 INTRODUCTION

Lichenized fungi form stable symbiotic relationships with algae and/or cyanobacteria (Crespo et al. 2014; Nash 2008). They form vegetative structures, the so-called lichen thalli, to house the photosynthetically active partner. Lichens are known to produce a large number of secondary metabolites (many of them extrolites) with almost 1000 known substances, the large majority exclusively found in lichen-forming fungi (Calcott et al. 2018; Elix et al. 1984; Hawksworth 1976; Huneck and Yoshimura 1996; Crespo et al. 2010; Lumbsch 1998; Stocker-Wörgötter 2008). Extrolites are deposited extracellularly, mostly in the medullary layer of the lichen thallus or in the cortical layer. Only a small number of substances occur in the cortex. Among those, coupled phenolics, which originate from polyketide pathways, such as depsides, depsidones and usnic acids are found almost exclusively in lichens. In macrolichens, the most common cortical substances are the depside atranorin or usnic acid. The occurrence of these cortical substances is usually constant for clades and has been used to circumscribe genera (Lumbsch 1998) especially in Parmeliaceae, which is the largest family of lichen-forming fungi (Crespo et al. 2010; Elix, 1993; Thell et al. 2012). Usnic acids and cortical substances in general protect the photobiont from solar radiations as UV-B (reviewed in Solhaug and Gauslaa 2011). Within Parmeliaceae, the occurrence of either usnic acid or atranorin is mostly constant for genera but sometimes closely related genera have different cortical chemistry, making the occurrence of these substances over the phylogenetic tree scattered. This pattern suggests several switches in the evolutionary history of these organisms. A previous ancestral character state reconstruction analysis suggested that the ancestor of parmelioid lichens, which is the largest group within Parmeliaceae, contained usnic acid and that this substance has been lost and replaced by atranorin several times independently (Divakar et al. 2013). This study used the presence or absence of secondary metabolites as phenotypical characters, however, did not address genes involved in the biosynthesis of the substances.

Experimental work with lichen-forming fungi is hampered by the fact that they are extremely slow growing; however, the recent availability of genomic data has the potential to demonstrate the underlying genetics for secondary metabolism in lichen-forming fungi (Abdel-Hameed and Sorensen 2018b). Therefore, comparative genomics recently identified that a polyketide synthase gene cluster putatively encodes for the biosynthesis of usnic acid (Abdel-Hameed et al. 2016). Genomic data sets also increased our knowledge of presence and domain architecture of polyketide synthases, which are key for the biosynthesis of fungal phenolics (Calchera et al. 2019).

The biosynthetic pathways that produce secondary metabolites in filamentous fungi - including lichen forming fungi - are typically organized into contiguous gene clusters in the genome, i.e. biosynthetic gene cluster (BGC). These gene clusters contain the chemical backbone synthesis genes, such as nonribosomal peptide synthases (NRPSs) and polyketide synthases (PKSs), tailoring enzymes, transporters and, often, transcription factors that control the expression of the clustered genes (Keller, 2019). Polyketide synthases (PKS) catalyze repetitive condensations of an acetyl-coenzyme A (CoA) starter with malonyl-CoA units. Most of the fungal PKSs are type I PKSs, which consist of a set of ß-ketoacylsynthase (KS), acyl transferase (AT), and acyl carrier protein (ACP) domains. The PKSs include non-reducing (NRPKS), partially reducing (PRPKS), and highly reducing (HRPKS) types depending on the extent of reduction of the polyketides they code for (Bingle et al. 1999; Miao et al. 1998; Nicholson et al. 2001). Based on phylogenetic relationships and domain architecture, the PKSs were further divided into subclades (Kroken et al. 2003) and horizontal gene transfer from bacteria was hypothesized as a mechanism for fungi to gain PKSs (Kroken et al. 2003; Schmitt and Lumbsch 2009). Within the NRPKS eight groups (I, II, III, IV, V, VI, VII and VIII) with known functions has been described in fungi so far (Liu et al. 2015). The genomes of ascomycetes fungi usually contain dozens of biosynthetic gene clusters. These are either species specific or broadly taxonomically distributed and are often very different between species (Khaldi et al. 2010; Keller, 2015). The total numbers of BGC contents can differ widely even between very closely related species (Lind et al. 2015; Lind et al. 2017). While the BGC contents have been largely studied in non-lichenized ascomycetes fungal species, they are poorly known in lichen forming fungal taxa (Calchera et al. 2019).

We use a genomic approach to elucidate the biosynthetic gene content and the evolution of the PKS gene that putatively is central for the biosynthesis of usnic acid. We test whether the presence and absence of usnic acid in a lichen is caused by differential expression of genes or reflects the presence of a gene cluster in the genome of a lichen-forming fungus. We specifically focus on the research questions: 1) do all usnic acid-containing lichen-forming fungi share a homologous gene cluster?, 2) is this gene cluster present in lichen-forming fungi that do not produce usnic acid?, 3) what is the arrangement of the gene cluster and does it change within Parmeliaceae, and 4) how can the presence or absence of the putatively usnic acid producing gene cluster be explained?

7.2 MATERIALS AND METHODS

Taxon Sampling

A total of 46 producers and non-producers of usnic acid of lichen-forming fungal species were included in this study. We included genomes of species belonging to different orders of Lecanoromycetes including Teloschistales (*Xanthoria parietina* and *Gyalolechia flavorubescens*), Umbilicariales (*Lasallia hispanica*), and Lecanorales (family Cladoniaceae: *Cladonia grayi*, *C. macilenta*, *C. metacorallifera*; Lecanoraceae: *Rhizoplaca melanophthalma*; Parmeliaceae: 39 species representing six of its seven major clades; see Table 4.1).

Genome Completeness Assessment and Phylogenomic analysis

BUSCO analysis is a well-known procedure and widely used to evaluate the genome completeness (Simão et al. 2015). Here, we used the Pezizomycotina data set containing 3156 single-copy genes to assess the completeness of our 46 genomes. To infer the phylogenetic relationships between producers and non-producers of usnic acid, every genome was explored following the methodology described in (Pizarro et al. 2018). The complete BUSCO single-copy genes predicted in each genome were extracted and aligned using MAFFT L-INS-i (Standley 2013). A supermatrix was created by concatenating all alignments using FASconCAT.pl (Kück and Longo 2014). Evolutionary relationships were inferred from this subset using maximum likelihood (ML) analysis implemented in IQTree v1.5.5 (Nguyen et al. 2015; Kalyaanamoorthy et al. 2017) with standard model selection and 1,000 bootstrap replicates. The resulting tree was visualized using FigTree 1.3.1 (Rambaut, 2009).

Biosynthetic Gene Cluster Prediction and Usnic Acid Genes Cluster Identification

In order to identify the usnic acid PKS cluster, gene prediction and annotation was performed in every genome using AUGUSTUS (Stanke et al. 2004) and MAKER2 (Holt and Yandell, 2011). In parallel, we used every genome and its respective gene prediction as input of Antismash 4.0 (Blin et al. 2017) for secondary metabolites gene cluster prediction. Afterwards, MPAS of *Cladonia uncialis* (A0A0R8YWJ7) and the orthologous protein PKS4 of *Usnea longissima* (AGI60156) sequences were downloaded from Uniprot data base and HMM profile was created using MAFFT and HMMER hmmbuild (Mistry et al. 2013). These profile and hmmsearch tool were used to search the most similar sequences present in every protein data set predicted from each genome. Only Non Reducing Polyketides Synthases were taken for further analysis.

Full-length NR-PKS and KS Domain Phylogenetic Analysis

We used Hidden Markov Model profiles for the identification of a specific type of domain. The recovered NR-PKS proteins were scanned using HMMER against Pfam (Mitchell et al. 2015) protein domain collection. Then using a bioinformatic approach with BEDtool (getfasta) (Quinlan and Hall 2010), the different KS domains from PKS were identified in the corresponding amino acid sequences. An identical procedure was carried out over an additional curated NR-PKS data set (Table 4.2) of characterized fungal NR-PKSs. Multiple alignments were performed using MAFFT L-INS-I over the amino acid sequence of full-length PKS dataset and KS domain dataset. Evolutionary relationships were inferred using ML analysis implemented in IQtree v1.5.5 with standard model selection. For each analysis, a Neighbour Joining and 1000 bootstrap replicates were calculated using fast bootstrapping option. The resulting phylogenetic trees were rooted with 6 Metylsalicylic PKS (6MS) sequences and drawn using FigTree v1.3.

Synteny Comparison of Usnic Acid Biosynthetic Gene Cluster

Genomes that contain the full genes cluster of usnic acid biosynthesis were selected to synteny comparison. Since some of usnic acid producer species had low gene cluster completeness, only *Cladonia macilenta*, *Alectoria sarmentosa*, *Evernia prunastri* and *Rhizoplaca melanophthalma* were recovered for the analysis. In addition, *Usnea florida* and *Lobaria pulmonaria*, downloaded from JGI were also included in the analysis. Using Orthofinder (Emms and Kelly 2015) we initially compared the homology among different genomes and subsequently drew their relationship.

7.3 RESULTS AND DISCUSSION

Producers and non-producers

Phylogeny inferred from the concatenated dataset of 2556 BUSCO genes showed strongly supported relationships between Lecanoraceae and Cladoniaceae as sister-groups of Parmeliaceae (Fig. 4.1). All nodes in the tree received strong support (100%). Usnic acid-producing species were widely spread along Lecanorales. All families included in this study had representatives of usnic acid-producing species as *Rhizoplaca melanophthalma* (Lecanoraceae), *Cladonia metacorallifera* (Cladoniaceae) and several usnic acid-producing species representing the six major-clades of Parmeliaceae (see Fig 4.1). The orthologous usnic acid biosynthetic gene cluster was absent in all studied species lacking usnic acid (Fig. 4.3). Further, in some cases producer and non-producer species formed strongly supported sister relations such as *Cladonia macilenta* (non-producer) and *C. metacorallifera* (producer), and *Xanthoparmelia chlorochroa* (producer) and *X. pulla* (non-producer), suggesting that the usnic acid biosynthetic gene might have lost in the genome of these closely related non-producer species during the evolution. Similar results have been found in

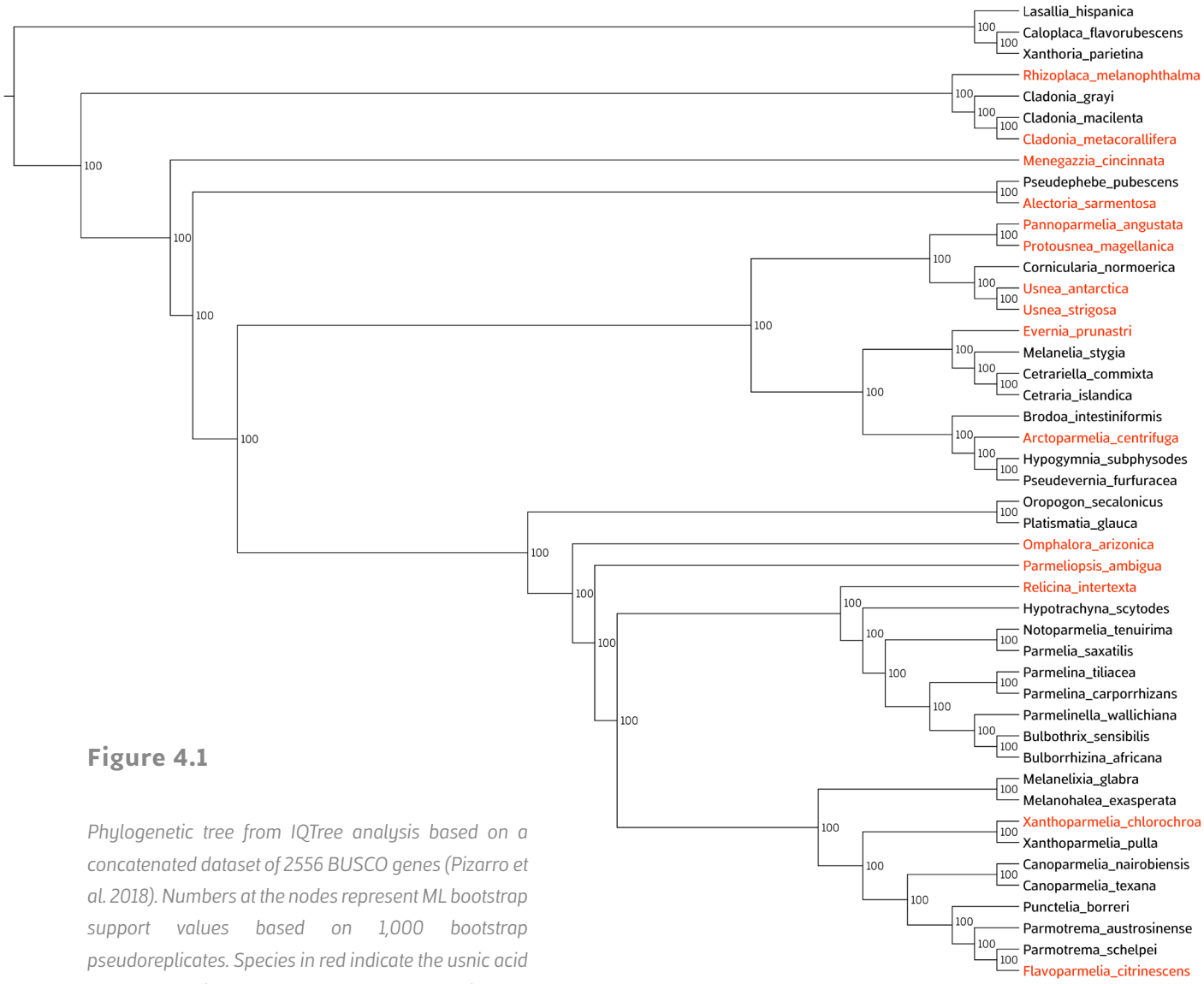


Figure 4.1

Phylogenetic tree from IQTree analysis based on a concatenated dataset of 2556 BUSCO genes (Pizarro et al. 2018). Numbers at the nodes represent ML bootstrap support values based on 1,000 bootstrap pseudoreplicates. Species in red indicate the usnic acid producer species and black non-producer species.

other groups of biosynthetic gene cluster and fungi, for example bikaverin BGC were entirely lost in different *Botrytis* species (Campbell et al. 2013).

In the 2556-gene phylogeny, the species containing usnic acid biosynthetic gene cluster (i.e. producer) do not form one monophyletic cluster. In contrast, the usnic acid producing species are part of a monophyletic group including both usnic acid producers and non-producers, resulting in a pattern of distribution of usnic acid producers across the Lecanorales phylogeny. Such a pattern has been suggested to indicate that the presence of a functional gene is ancestral (Lind et al. 2017). Our results are in accordance with the previous

finding showing usnic acid production as an ancestral state for the parmelioid lichens, the major clade within Parmeliaceae (Divakar et al. 2013). The multiple switches of producer and non-producer species across the phylogenomic tree is suggestive of a repeated loss of the usnic acid biosynthetic gene compared to multiple independent gains by horizontal gene transfer during the evolution of these fungi. Since sequences of the PKSs putatively coding for usnic acid are similar and its phylogeny largely reflects the species phylogeny found in previous studies and the lack of long branches within the phylogeny of the PKS cluster, is evidence that the scattered occurrence of usnic acid in the phylogenetic tree of these lichen-forming fungi cannot be explained by horizontal gene transfer. In addition, the discontinuous distribution of PKSs genes among fungal species is explained by gene loss than horizontal gene transfer (see e.g. (Kroken et al. 2003; Bushley and Turgeon 2010; Campbell et al. 2012; Lind et al. 2017)). This discontinuity of the taxonomic distribution may explain convergent evolution of usnic acid in lichen forming fungi (Stayton, 2015). The evolutionary relations among Lecanorales taxa were highly similar to those published previously (Chapter 2) and therefore are not discussed in detail here.

Biosynthetic Gene Cluster Content

A genome-wide examination highlights that BGC contents are diverse across Parmeliaceae taxa, and provides a road map to the genomic changes essential for the extensive diversity of secondary metabolites in this group of lichen forming fungi. The species with a higher number of biosynthetic gene clusters (BGC) were *Evernia prunastri* with 98 BGC followed by *Pannoparmelia angustata* with 78 BGC and *Parmotrema austrosinense* with 75 BGC. However, the species with lowest number were *Canoparmelia nairobiensis* with 10 BGC, *Bulbothrix sensibilis* with 11 BGC and *Hypotrachyna scytodes* with 13 BGC (Fig. 4.2). While the BGC have been studied for individual species in lichen forming fungi (see e.g. (Armaleo et al. 2011; Bertrand et al. 2018a; Dal Grande et al. 2018b; Calchera et al. 2019), this has not yet been compared at large scale e.g. family level. To our knowledge, this is the first study comparing BGC contents in species belonging to different major clades of the family Parmeliaceae (Fig. 4.2). Within Parmeliaceae, 80 BGC have been reported in *Evernia prunastri* and

51 in *Pseudevernia furfuracea* (Calchera et al. 2019); in Cladoniaceae, *C. uncialis* contained 48 BGC (Bertrand et al. 2018a); in Teloschistaceae, *Gyalolechia flavorubescens* had 13 BGC (Park et al. 2013); and in Umbilicariaceae, *Umbilicaria hispanica* contained 18 BGC (Dal Grande et al. 2018a). While the highest number of BGC has been found in some Parmeliaceae taxa (Fig. 4.2), no correlations between BGC and secondary metabolism or ecology were found. *Evernia prunastri* contain atranorin, usnic acid and evernic acid, and *P. furfuracea* has atranorin, physodic and olivetoric acid (Table 4.1), both species are wide spread in temperate regions. *P. austrosinense* contain atranorin and lecanoric acid (Table 4.1) and are widely distributed in the tropics.

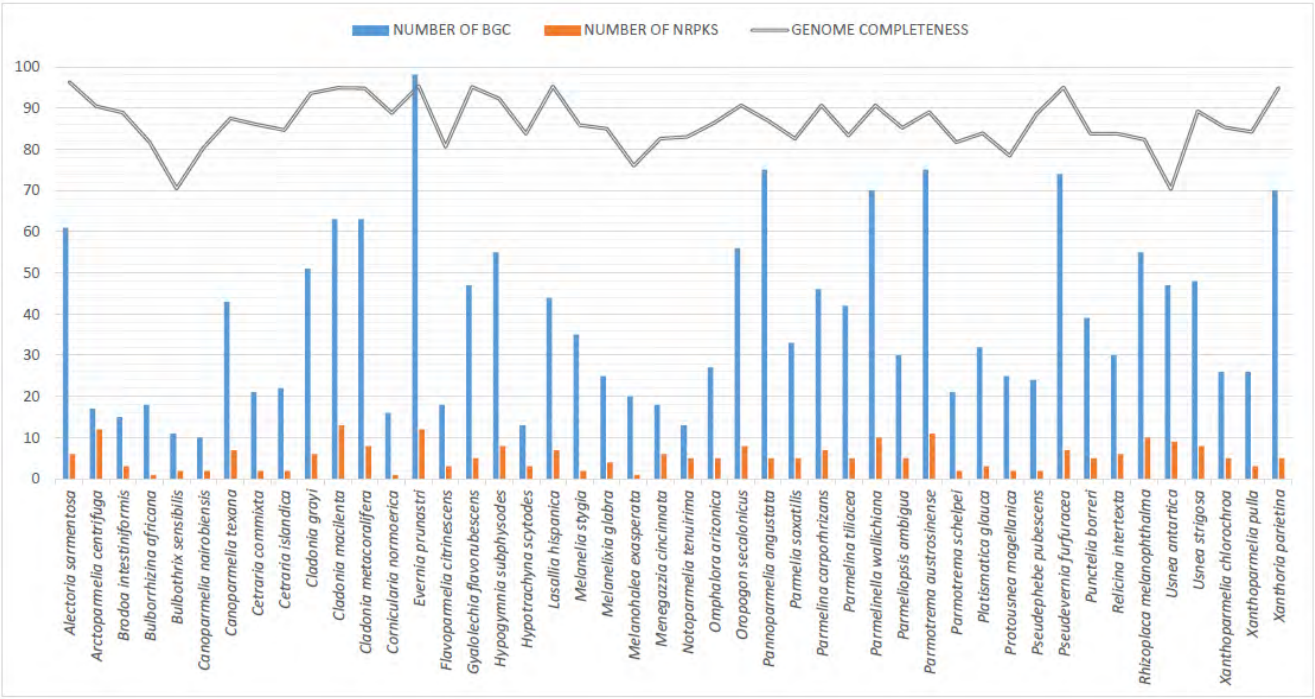


Figure 4.2

Genome completeness and biosynthetic gene cluster assessment in 46 genomes of lichen-forming fungi used for this study. Number of biosynthetic gene cluster (blue bar), Number of Non Reducing PKS (Orange bar), and percentage of Genome Completeness (grey line). The Y-axis represents the values of percentage and the number of biosynthetic gene clusters.

Phylogeny of Non Reducing PKS

A NJ phylogeny of the KS domain and full PKS sequence revealed the evolutionary relationship of biosynthetic gene content and the origins of the usnic acid PKS genes in lichen forming fungi. The KS domain and PKS full-length sequence showed highly similar tree topology (Fig. 4.3) and therefore only PKS phylogeny is discussed in detail. The data matrix of amino acid sequences of PKS domain consists of 356 columns and 167 taxa. All sequences included in the analysis passed the composition chi2-test (P<5%) and the best substitution model was LG+I+G4 according to BIC. The phylogenetic tree inferred from the PKS data matrix was rooted using polyketide 6-methylsalicylic acid sequences (6MS). In the resulted tree, two main groups were found among Non-Reducing Polyketide Synthases (Fig. 4.3). The well-supported topology is concordant with previous reported KS phylogenies (Liu et al. 2015).

The first monophyletic cluster comprised the groups I-V of NR-PKS. The group I included PKS sequences for the biosynthesis of aromatic compound derived from orsellinic acid, such as grayanic acid in *Cladonia grayi* (Armaleo et al. 2011) or mycotoxins, such as zearalenone in *Fusarium gramineanum* (Gaffoor and Trail 2006). The PKS of this group show a domain arrangement of SAT, KS, AT, PT, one or two ACP domains and TE. Group V comprised PKS without the TE domain and is implicated in the production of different mycotoxin, such as desertorin (*Aspergillus nidulans*) or atrochrysone (*Aspergillus fumigatus*) (Lim et al. 2012; Galagan et al. 2005). Groups II-III-IV formed a supported sister-relationship to group V, comprising PKS with TE/CLC domain located on the N-terminal with a similarity in the domain arrangement. Group II contained PKS involved in melanin biosynthesis as PKS1 of *Exophiala dermatiditis* (Feng et al. 2001) and it is characterized by having two ACP domain between PT and TE; in this group many homologous genes to melanin biosynthesis of lichen-forming fungi were found. This indicates that the NR-PKS group II could be responsible for the biosynthesis of melanin in lichen forming fungi. Homology search using a known gene clusters may allow identification of related gene clusters in other fungal genomes (Khaldi et al. 2010; Gardiner and Howlett 2005). However, an additional study will be necessary in order to confirm its function in lichens. Groups III and IV NR-PKS contain proteins that synthesize large polyketides chains as conidial yellow pigment Alb1 (group III) or aflatoxin/sterigmatocystin of *A. nidulans* (Yu & Leonard, 1995) (group IV) (Table 4.2).

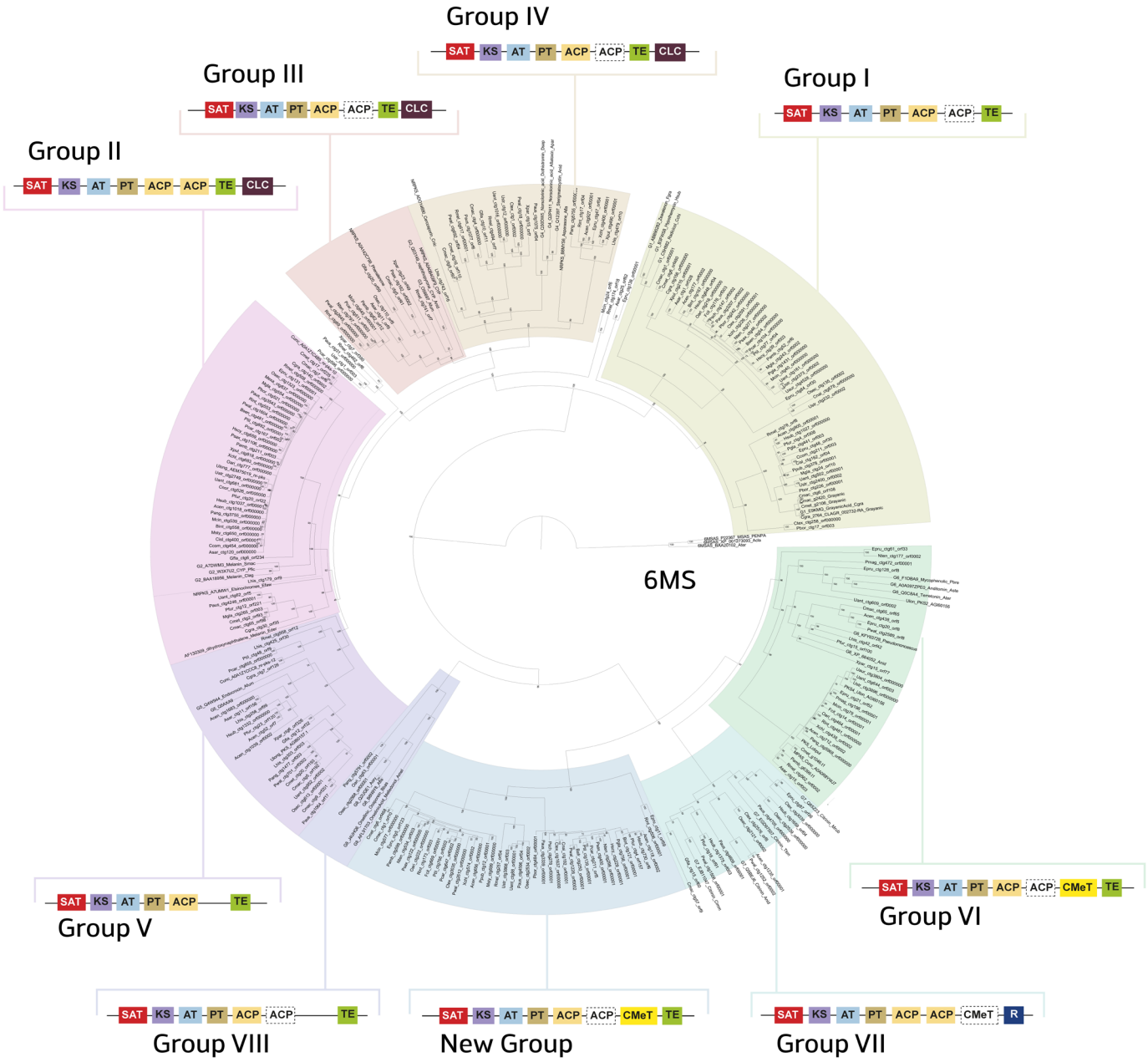


Figure 4.3

Gene tree of NR-PKS dataset inferred by ML analysis in IQtree using 6 Methylsalicylic PKS protein sequences as outgroup. The distinct colors represent the different groups corresponding to the domain arrangement. In each group the domain arrangement of NR-PKS is highlighted with distinct colors.

Abbreviations: SAT = starter unit-ACP transacylase, KS = ketosynthase, AT = acyl transferase, PT = product template, ACP = acyl carrier protein, TE = thioesterase, TE/CLC = thioesterase/Claisen cyclase, CMeT = C-methyltransferase, R = reductase.

The second monophyletic cluster included the groups VI-VII-VIII and an undescribed group (named hereafter "New Group") sister to VI and VII group, which does not contain PKS of known function. As eight NR-PKS groups with known functions have been described in fungi (Liu et al. 2015), our results suggest that this group with unknown function may belong to a new NR-PKS group. This result can serve as valuable entry point to search for functional gene clusters in fungi. However, we refrain to describe this group formally as this is not the main focus of our study and an additional study will be needed. Sequences located on the groups VI, VII and the New Group are largely characterized by having a methylation domain (CMet) between the ACP domain and the N-terminal domain. Group VII, which is responsible for the biosynthesis of citrinin in many fungal species (*Aspergillus nidulans*, *Coccidioides immitis*, Table 4.2), (Gallo et al. 2013; Chiang et al. 2009), is characterized by having a reductase (R) domain placed on the N-terminal. The domain arrangement of this probable new group is similar to group VI. While the new group included only lichen-forming fungi, the Blast analysis found homologous sequences belonging to other groups of fungi with unknown functions. Group VI contained the PKS responsible to usnic acid biosynthesis in lichens, where only producer species were present (Fig. 4.3). We find that all producer species contained PKS sequences homologous to the PKS of usnic acid of *Cladonia uncialis* (MPAS) (Abdel-Hameed et al. 2016). This orthologous gene cluster was absent in all non-producer species. However, in a recent study a PKS gene cluster closely related to usnic acid biosynthetic gene was reported in a non-producer species, *Lobaria pulmonaria* (Calchera et al. 2019), which contains the despidones stictic, norstictic, and constictic acids.

Within group VI a strong phylogenetic structure was observed. Three main strongly supported (>95 bootstrap) monophyletic groups were found concordant with the domain arrangement. The short branches within the phylogeny of Usnic acid PKS gene cluster and concordance in phylogenetic relations (in producer genera e.g. Usnea, Fig. 4.4) between PKS phylogeny and species phylogeny (Pizarro et al. 2018), is indicate that the scattered occurrence of usnic acid across the phylogenetic tree of these lichen-forming fungi cannot be explained by horizontal gene transfer (Kroken et al. 2003; Bushley & Turgeon, 2010; Campbell et al. 2013; Lind et al. 2017).

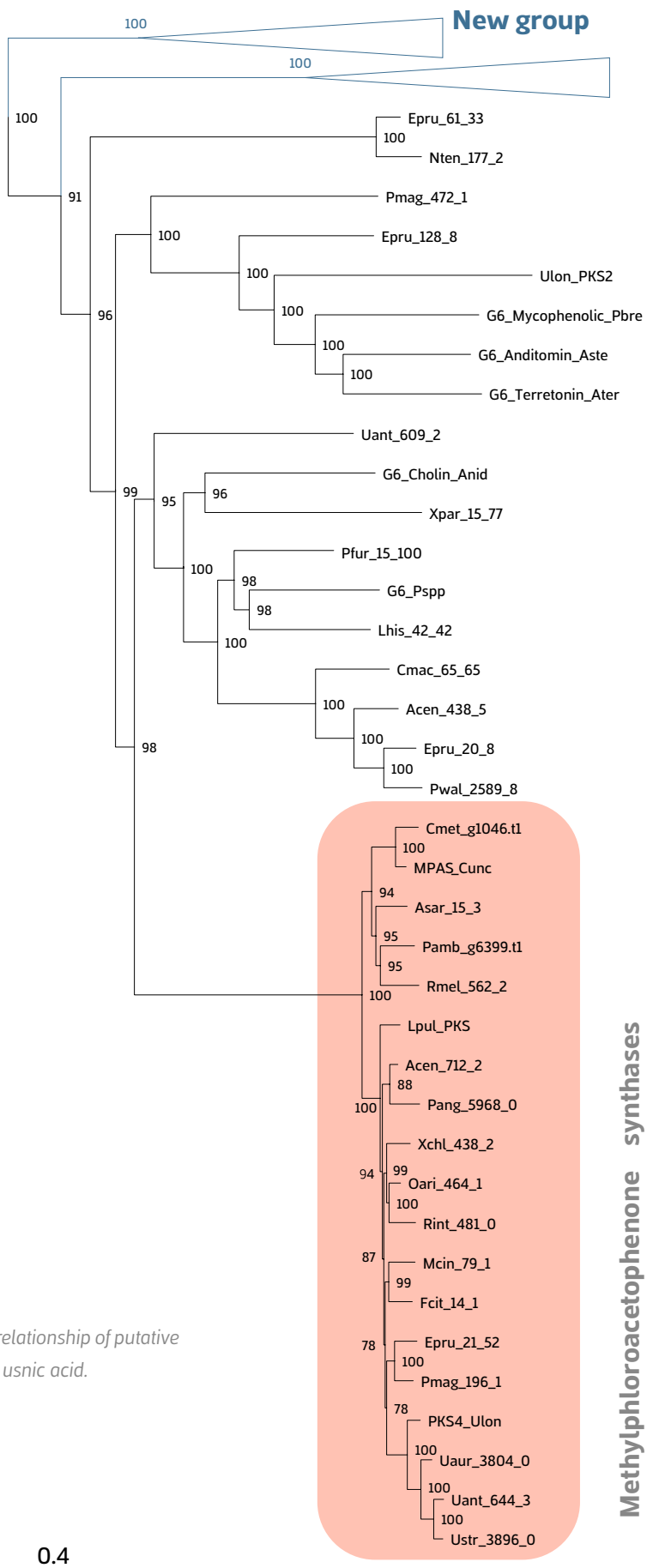


Figure 4.4

Fragment of NR-PKS tree showing the relationship of putative PKS responsible for the biosynthesis of usnic acid.

Gene Arrangement of Putative Usnic Acid Biosynthetic Gene Cluster

The biosynthetic gene cluster implicated in usnic acid production has been compared between six species of Lecanoromycetes. All genomes of producer species shared homologous genes of NR-PKS (as MPAS in *Cladonia uncialis*), CYT450 (MPAO in *Cladonia uncialis*)(Abdel-Hameed et al. 2016) and one or two transcription factors (regulators) (Fig.4.5). The gene arrangement of BGC in *Usnea florida* is very similar to *Evernia prunastri* and *Alectoria sarmentosa*, which all belong to Parmeliaceae. They share homologous genes that are likely involved in the usnic acid biosynthesis since the Lacasse (Jeon et al. 2012) and the putative oxidoreductase gene, which could be implicated in the oxidation of precursors molecules of usnic acid. There are two genes (in red in Fig. 4.5) that contain Cys6-Zn domain specific of many fungal transcription factors (Shelest, 2017) suggesting that they could be regulators of usnic acid production. We found an O-methyl transferase gene flanking the PKS genes in all species, except *Cladonia metacorallifera* and *Rhizoplaca melanophthalma*. This O-methyltransferase gene is common in many BGC and is homologous to the coactivator AflJ, implicated in the biosynthesis of aflatoxin in *Aspergillus parasiticus* (Chang et al. 2003) and to mdpA in *A. nidulans* (Chiang et al. 2009). Flanking the left of the CYT450 is located a FAD/NADP containing protein in all species included with exception of *Lobaria pulmonaria*, which has a serin/threonin kinase. *Evernia prunastri*, *Alectoria sarmentosa*, *Cladonia metacorallifera* and *Rhizoplaca melanophthalma* contain a gene DUF3112 with unknown function, which is in different locations depending on the species. For example, in *Evernia prunastri* it is located between both regulators, whereas in *Alectoria sarmentosa* it is between PKS gene and a regulator. An ABC transporter that putatively would be implicated in the secretion of usnic acid is only found in the *Evernia prunastri* genome.

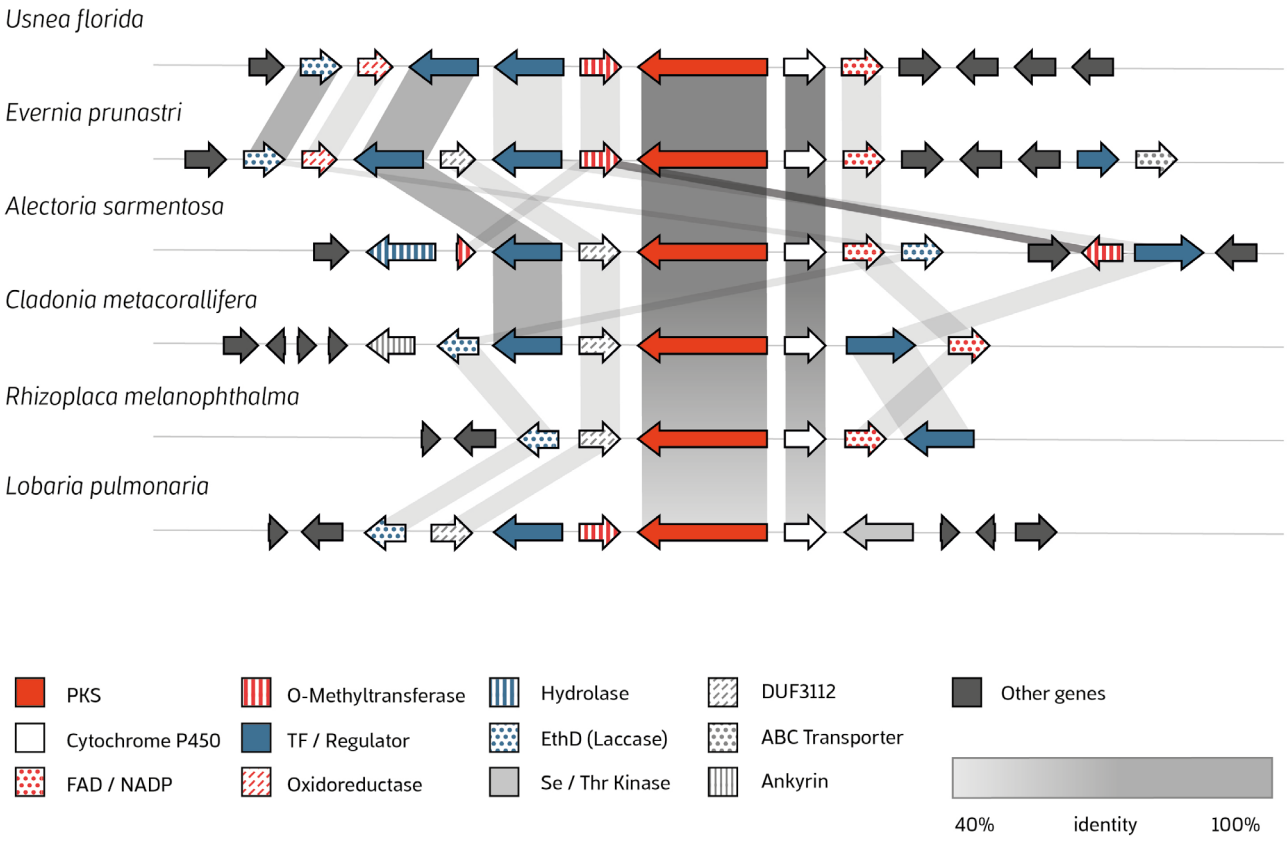


Figure 4.5

Putative usnic acid compound cluster (UA) conservation and synteny. The UA cluster from other lichen-forming fungi resembled to the characterized usnic acid core genes (MPAS and MPAO) from Cladonia uncialis (high percentage of identity in protein-by-protein comparisons of both genes in all species included).

7.4 CONCLUSION

Here we present a first comprehensive study investigating biosynthetic gene contents in the genomes of the largest clade of lichen forming fungi. Comparative genomic analyses shed a light on BGC in this group of fungi. A total of eight NR-PKS groups (I-VIII) with known functions and in addition, one novel NR-PKS group with unknown function in fungi were identified. However, the latter needs additional study. Noteworthy, all species lacking usnic acid also lacked the biosynthetic gene cluster responsible for usnic acid biosynthesis. This domain was homologous to the known usnic acid biosynthetic gene cluster (MPAS) and CYT450 (MPAO) of *Cladonia uncialis* and conserved in all producer species. While the gene arrangement of putative usnic acid biosynthetic gene cluster was highly similar in taxa within the family, they were different among families. Within usnic acid NR-PKS phylogeny, we found strong phylogenetic structure, relatively short branch lengths and three strongly supported main monophyletic groups concordant with domain arrangement. The short branch lengths within the NR-PKS phylogeny, tree topology concordance with species tree of producer taxa and conserved nature of this biosynthetic gene cluster among all usnic acid producer species, suggest that the acquisition of usnic acid biosynthetic genes in the genome of producer species can not be explained by horizontal gene transfer. Rather, we hypothesized that the usnic acid biosynthetic gene cluster might have entirely lost in the non-producer species during the evolution of these fungi. The rapid genomic adaptations may have played important role in the evolution of these fungal species.

Further, our study provides insights into the evolutionary or adaptive success of usnic acid containing taxa to divergent ecological niches. For example, *Xanthoparmelia* comprises over 600 species (of 820 described) with usnic acid, is wide spread in semi-arid regions and have been shown to the adaptive success of this group of fungi in xeric conditions (Lumbsch et al. 2008). Our study allows us to hypothesize that usnic acid may play an important role in the adaptation to divergent ecological niches and open a new research line in this field.

Table 4.1
List of species included in the analysis with number of predicted BGC and secondary metabolites recovered from literature.

Species	Synonym	Usnic Acid	Number of Metabolic Clusters	Compounds
Alectoria sarmentosa	Asar	Producer	61	Usnic acid, alectoronic acid, olivetoric acid
Arctoparmelia centrifuga	Acen	Producer	17	Usnic acid, alectoronic acid, protocetraric acid
Brodoa intestiniformis	Bint	Non-producer	15	Atranorin, prcinol, B-orcinol depsidons
Bulborrhizina africana	Bafr	Non-producer	18	Atranorin, gyrophoric acid, lecanoric acid
Bulbothrix sensibilis	Bsen	Non-producer	11	Atranorin, gyrophoric acid, lecanoric acid
Canoparmelia nairobiensis	Cnai	Non-producer	10	Atranorin, chloroatranorin, Divaricatic/nordivaricatic acid
Canoparmelia texana	Ctex	Non-producer	43	Atranorin, Divariatic acid,stictic acid
Cetraria commixta	Ccom	Non-producer	21	A-collatolic acid
Cetraria islandica	Cisl	Non-producer	22	Fumarprotocetraric acid, protolichesterinic and lichesterinic acid)
Cladonia grayi	Cgra	Non-producer	51	Grayanic/protocetraric/fumarprotocetraric/divaronic/stenophoric acid
Cladonia macilenta	Cmac	Non-producer	63	Barbatic, squamatic,thamnolic, didymic acid (some Usnic acid)
Cladonia metacoralifera	Cmet	Producer	63	Usnic acid, didymic/squamatic/barbatic acid
Cornicularia normoerica	Cnor	Non-producer	16	No compounds
Evernia prunastri	Epru	Producer	98	Atranorin, Usnic acid, Evernic acid
Flavoparmelia citrinescens	Fcit	Producer	18	Usnic acid, Atranorin, protocetraric acid, caperatic acid
Gyalolechia flavorubescens	Gfla	Non-producer	47	Antraquinon
Hypogymnia subphysodes	Hsub	Non-producer	55	Atranorin, physodic acid (Borcinol depsidones –> protocetraric and physodalic)
Hypotrachyna scytodes	Hscy	Non-producer	13	4-O-demethylbarbatic acid | atranorin | barbatic acid | chloroatranorin | echinocarpic acid
Lasallia hispanica	Lhis	Non-producer	44	Gyrophoric acid, papulosin, antraquinon
Melanelia stygia	Msty	Non-producer	35	Fumarprotocetraric acid / caperatic acid
Melanelixia glabra	Mgla	Non-producer	25	Lecanoric acid
Melanohalea exasperata	Mexa	Non-producer	20	Fumarprotocetraric acid / norstictic acid / no compounds
Menegazzia cincinnata	Mcin	Producer	18	Atranorin, depsides and depsidones
Notoparmelia tenuirima	Nten	Non-producer	13	Atranorin, chloratranorin; lobaric acid, salazinic, consalazinic, protocetra-ric and Fumarprotocetraric acids, protolichesterinic acid, echinocarpicacid acid
Omphalora arizonica	Oari	Producer	27	Usnic acid
Oropogon secalonicus	Osec	Non-producer	56	Secalonic Acid
Pannoparmelia angustata	Pang	Producer	75	Usnic acid, divaricatic acid
Parmelia saxatilis	Psax	Non-producer	33	Atranorin, salazilic acid, lobaric and norstictic acid)
Parmelina carporhrizans	Pcar	Non-producer	46	Atranorin, Lecanoric acid
Parmelina tiliacea	Ptil	Non-producer	42	Atranorin, Lecanoric acid
Parmelinella wallichiana	Pwal	Non-producer	70	Atranorin, salazilic acid
Parmeliopsis ambigua	Pamb	Producer	30	Usnic acid, Atranorin, divaricatic acid)
Parmotrema austrosinense	Paus	Non-producer	75	Atranorin, Lecanoric acid
Parmotrema schelpei	Psch	Non-producer	21	Stictic, Norstictic and protocetraric acid
Platismatica glauca	Pgla	Non-producer	32	Atranorin, Caperatic acid
Protousnea magellanica	Pmag	Producer	25	Usnic acid, sekikaic acid
Pseudephebe pubescens	Ppub	Non-producer	24	No compounds
Pseudevernia furfuracea	Pfur	Non-producer	74	Atranorin, physodic/olivetoric/lecanoric acid
Punctelia borreri	Pbor	Non-producer	39	Atranorin, gyrophoric acid
Relicina intertexta	Rint	Producer	30	Usnic acid and protocetraric acid
Rhizoplaca melanophthalma	Rmel	Producer	55	Usnic acid, psoromic/placodiolic/lecanoric/psoromic acid
Usnea antartica	Uant	Producer	47	Usnic acid (Barbatic/diffractaic/squamatic/salazinic/norstictic/stictic)
Usnea strigosa	Ustr	Producer	48	Usnic acid (Barbatic/diffractaic/squamatic/salazinic/norstictic/stictic)
Xanthoparmelia chlorochroa	Xchl	Producer	26	Usnic acid, salazinic acid/norstictic acid
Xanthoparmelia pulla	Xpul	Non-producer	26	Gyrophoric, Stenophoric, Divaricatic acid
Xanthoria parietina	Xpar	Non-producer	70	Parietin

Table 4.2
List of additional sequences included on the PKS gene phylogeny with their corresponding ProteinID, metabolic product when is known and the subgroup of NR-PKS.

Species	Abbreviation	Protein ID	Compound
Cladonia grayi	Cgra	E9KMQ	Grayanic Acid
Fusarium gramineanum	Fgra	ABB90282	Zearalenon
Hypomyces subiculosus	Hsub	B3FWS8	Hypothemycin
Chaetomium chiversii	Cchi	C5H882	Radicicol
Exophiala dermatitidis	Eder	AF130309	Melanin (Dihydroxynaphthalene)
Glomerella lagenarium	Clag	BAA18956	Melanin
Sordaria macrospora	Smac	A7DWM3	Melanin
Pestalotiopsis fici	Pfic	W3X7U2	CYP
Elsinoe fawcettii	Efaw	A7UMW1	Elsinochromes
Cladonia uncialis subsp. *uncialis*	Cunc	A0A1Z1C4B5	Unknown
Usnea longissima	Ulon	AEM75019	Unknown
Penicillium herquei	Pher	A0A142C799	Phenalenone
Metarhizium guizhouense	Mgui	A0A0B4G9F9	CYP anthraquinone
Neosartorya fumigata	Afum	O59897	Conidial Pigment Alb1
Emericella nidulans	Anid	Q03149	Naphthopyrone CYP
Aspergillus flavus	Afla	B8MYS6	Asparasone anthraquinone
Mycosphaerella coffeicola	Cnic	ADO14690	Cercosporin
Dothistroma septosporum	Dsep	Q30DW5	Norsolorinic acid (Dothistromin)
Aspergillus oryzae	Apar	Q2PH11	Norsolorinic acid (Aflatoxin)
Emericella nidulans	Anid	Q12397	Sterigmatocystin
Cladonia uncialis subsp. Uncialis	Cunc	A0A1Z1CCC8	Unknown
Emericella nidulans	Anid	Q5AXA9	
Neosartorya fumigata	Afum	Q4W944	Endocrocin
Usnea longissima	Ulon	AGI60157.1	
Penicillium brevicompactum	Pbre	F1DBA9	Mycophenolic (5-methylorsellinic acid)
Usnea longissima	Ulon	AGI60155	
Emericella variicolor	Aste	A0A097ZPE0	Anditomin (3,5-dimethylorsellinic acid (DMOA))
Aspergillus terreus	Ater	Q0C8A4	Terretonin (3,5-dimethylorsellinic acid (DMOA)
Pseudogymnoascus sp.	Pseu	KFY63728	
Emericella nidulans	Anid	XP_664052	Colchin
Lobaria pulmonaria	Lpul	PKS_Lobpul	Usnic Acid homolog
Usnea longissima	Ulon	AGI60156	Usnic Acid
Cladonia uncialis subsp. Uncialis	Cunc	A0A0R8YWJ7	Usnic Acid
Monascus purpureus	Mrub	Q65Z23	Citrinin
Trichophyton tonsurans	Tton	EGD97507	Citrinin
Coccidioides immitis	Cimm	EJB11047	Citrinin
Emericella nidulans	Anid	Q5BEJ6	Citrinin
Armillaria mellea	Amel	AFL91703	OrselinicAcid_Melledonol
Beauveria bassiana	Bbas	J4UHQ6	Orsellinic_Oosporein
Aspergillus flavus	Afla	B8N9Y8	
Aspergillus oryzae	Aory	Q2UGE1	
Penicillium patulum	Ppat	P22367	Patulin
Aspergillus clavatus	Acla	XP_001273093	Patulin
Aspergillus terreus	Ater	BAA20102	

8.DISCUSIÓN GENERAL

8.GENERAL DISCUSSION

8

DISCUSIÓN GENERAL

La clase Lecanoromycetes forma, con mucho, el grupo más diverso de hongos liquenizados obligados, lo que convierte a esta clase de organismos en un grupo interesante para el estudio de las relaciones simbióticas. Al mismo tiempo, se trata de un grupo difícil de estudiar, sobre todo cuando se usan metodologías que previamente necesitan la obtención in vitro de materiales puros (axénicos) del micobionte, sin liquenizar. Los cultivos son poco productivos, delicados de mantener y muy lentos de crecimiento; por esta razón se procede con metodologías basadas en los análisis metagenómicos que minimicen las dificultades del proceso de cultivo. Hasta la fecha, la mayoría de los estudios llevados a cabo usando herramientas genómicas se han desarrollado en torno al estudio del microbioma bacteriano asociado al talo (Grube et al. 2015; Cervana et al. 2017), cultivo de micobiontes (Wang et al. 2014; SY Park et al. 2013), filogenómica (Leavit et al. 2016; Grewe et al. 2018;) y sobre identificación de metabolitos secundarios (Armaleo et al. 2011; Abdel-Hameed 2016; Calchera et al. 2019).

En esta tesis se propone el uso de metagenomas obtenidos a partir del talo liquénico como una herramienta alternativa a los genomas obtenidos de cultivos axénicos para abordar el estudio de líquenes Lecanoromycetes, con especial atención a la familia Parmeliaceae, la más diversa de hongos liquenizados. Los genomas del micobionte obtenidos a partir de los metagenomas han sido usados para contestar distintas preguntas relacionadas con su biología evolutiva. En primer lugar, se han establecido relaciones filogenéticas de la familia Parmeliaceae que en estudios previos no habían podido resolverse, contrastándose a su vez la utilidad de los metagenomas para la obtención de genes de una sola copia. Esto ha servido como base para abordar otras cuestiones no resueltas en torno a la biología evolutiva de esta exitosa simbiosis como son: los mecanismos genéticos de la reproducción sexual y la implicación de las estrategias reproductivas en el éxito evolutivo de la liquenización como proceso; en este sentido se destaca el haber concluido

en el principal grupo de hongos liquenizantes (Lecanoromicetos) que: (i) la estrategia reproductiva ancestral y generalizada es el heterotalismo, (ii) asimismo se ha llevado a cabo la identificación del clúster completo del ácido úsnico, que es un metabolito exclusivo de hongos liquenizados.

8.1 RECUPERACIÓN DE GENOMAS DE MICOBIONTES A PARTIR DE METAGENOMAS DEL TALO.

La separación de los distintos genomas del metagenoma o "metagenome skimming" se ha usado en diversos organismos eucariotas para llevar a cabo estudios biogeográficos y filogenómicos a diferentes niveles taxonómicos (Male et al. 2014; Denver et al. 2016; Grandjean et al. 2017). En líquenes se ha constatado que el uso de metagenomas a partir de talo es una aproximación válida para recuperar casi todo el genoma del micobiointe (Greshake et al. 2016) mediante la combinación de herramientas bioinformáticas. Disponer de una base de datos extensa, la cual incluya genomas o proteomas de referencia cercanos a la especie de estudio es algo capital cuando se trabaja con organismos "no modelo".

Los resultados obtenidos de en relación a las métricas de los genomas del micobionte de las diferentes especies incluidas en el estudio (Capítulo 1) en comparación con los datos de los metagenomas de *Evenia prunastri* y *Pseudevernia furfurácea* de Meiser et al. 2017 son similares, oscilando entre 85-95% de la integridad de los genomas. De igual forma, la longitud de los genomas y el número de genes obtenidos de cada una de las distintas especies están en línea con la mayoría de los genomas de líquenes recientemente secuenciados a partir de cultivos axénicos (Capítulo 1).

8.2 RECONSTRUCCIÓN DE LAS RELACIONES FILOGENÉTICAS NO RESUELTAS DE PARMELIÁCEOS.

El uso de un alto número de marcadores moleculares, grandes regiones genómicas o genomas completos ha ido creciendo en los últimos años como metodología para resolver relaciones filogenéticas complejas (Allio et al. 2019; Zhang et al. 2019). Los genomas de micobiontes obtenidos del análisis de los metagenomas son útiles para recuperar gran cantidad de genes de una sola copia, los cuales pueden ser usados como marcadores filogenéticos para establecer relaciones evolutivas. En el Capítulo 2 se muestra cómo las filogenias inferidas a partir de datos filogenómicos concatenados, así como el árbol de especies inferido mediante coalescencia muestran topologías muy similares, con pocas excepciones. La monofilia de los seis clados principales de Parmeliaceos incluidos en este estudio está más fuertemente apoyada, además las relaciones entre los clados principales también recibieron un mayor apoyo.

Por ejemplo, en ambos análisis el clado Alectorioide estaba apoyado como grupo hermano a todos los clados restantes, algo novedoso si lo comparamos con el estudio de referencia de Divakar et al. 2015, en el cual la posición del clado Alectorioide no estaba resuelta. Los otros clados a su vez se subdividieron en dos grupos principales bien apoyados. Por un lado, se encontraban los clados Evenioide, Cetrarioide, Hypogymnioide, Usneoide y Anzioide, y por el otro los Parmelioides junto con las especies de los géneros *Oropogon*, *Platismatia* y *Omphalora*. En comparación con el estudio de Divakar et al. 2015, donde la relación entre el clado Evernioide y Cetrarioide carecían de apoyos, en este estudio esta relación está fuertemente apoyada. Sorprendentemente, unas pocas relaciones previamente reconocidas dentro del clado Parmelioide (Crespo et al. 2010) fueron incongruentes entre los dos análisis. Un ejemplo concreto, es la posición discordante de *Parmeliopsis ambigua* entre ambos análisis, en el árbol de genes concatenados esta especie se establece fuera del clado Parmelioide, en cambio en el árbol de especie se posiciona dentro de él. Este resultado merecería una investigación más exhaustiva, incluyendo, por ejemplo, un mayor número de especies del mismo género, lo cual posiblemente ayudaría a establecer de una forma más precisa las relaciones filogenéticas dentro de los clados principales.

Es evidente que el enfoque metagenómico es capaz de recuperar un alto número de genes de una sola copia que son útiles para establecer relaciones evolutivas de linajes complejos como es la familia Parmeliaceae, en relación a estudios donde dichas relaciones evolutivas se infirieron mediante un número de genes limitados. Sin embargo, dado que el conjunto de genes obtenidos en este estudio son un valioso recurso que puede ser usado para identificar nuevos marcadores filogenéticos, resulta recomendable determinar el número mínimo de genes necesarios para obtener una filogenia robusta concordante con árbol de especies obtenido (Rokas et al. 2003). En nuestro caso concretamente, implementar el procedimiento descrito en Ai and Kang, 2015 fuerza a generar un mínimo de diez réplicas de subconjuntos aleatorios de estos genes de forma exponencial hasta alcanzar el total de genes incluidos en el análisis. Posteriormente, en la citada referencia, a partir de los alineamientos concatenados de cada réplica se generan árboles filogenéticos que deben ser comparados con el árbol de especie hasta la obtención de una misma topología que contenga los nodos fuertemente apoyados. En el estudio de Ai and Kang, 2015 se trabajaba con 11 especies y un número de 830 genes. En nuestro estudio, sin embargo, el número de especies (51) y genes (2.556) es muy superior. Por lo tanto, la aplicación de este procedimiento se hace desafortunadamente inviable debido a las limitaciones de tiempo de cómputo, algo que podría realizarse en un futuro cuando se tenga acceso a infraestructuras informáticas de más potencia.

8.3 ESTRATEGIAS REPRODUCTIVAS EN LECANOROMYCETES

La reproducción en hongos liquenizados ha sido estudiada ampliamente a nivel morfológico y anatómico dada la importancia que tienen estos rasgos como caracteres taxonómicos en este como en otros grupos. Sin embargo, las bases genéticas de la reproducción sexual en líquenes se mantenían en gran medida sin estudiar de manera suficientemente convincente. Estudios como el de Murtagh et al. 2000, basado en el análisis de variabilidad de las esporas mediante técnicas moleculares, habían sugerido que el homotalismo representaba el modo reproductivo principal de los hongos liquenizados. Los resultados de esta tesis sugieren lo contrario, es decir, que el heterotalismo es

la organización extendida del locus MAT en hongos pertenecientes a Lecanoromycetes, incluyendo varias especies supuestamente asexuales (Capítulo 3).

El análisis de la estructura génica y la sintenia del locus MAT muestra una alta conservación de secuencias en las regiones flanqueantes al locus (APN1 y SLA2), similar a la organización génica encontrada en otros hongos pertenecientes a Pezizomycotina (Turgeon et al. 1993; Yun et al. 1999; de Vries 2017). Dentro del locus MAT se han identificado por primera vez dos genes auxiliares asociados a cada uno de los ideomorfos, los cuales podrían ser específicos de líquenes. En alguna de las especies estos genes contienen un dominio característico de factores de transcripción, lo cual merecería un estudio más a fondo debido a su posible implicación en la regulación de la reproducción.

La presencia del heterotalismo en todas las especies pertenecientes a este grupo, da cabida a suponer que esto podría estar influenciado por el estilo de vida simbiótico de estos hongos. Como es conocido, esta estrategia reproductiva conlleva un alto riesgo debido a la dificultad de encontrar individuos compatibles, especialmente en poblaciones con un desequilibrio de los genes idomorfos *mat*. Al mismo tiempo, el heterotalismo podría propiciar una adaptación evolutiva acelerada que sería ventajosa en entornos con factores novedosos (Murtagh et al. 2000), ya que proporciona una mayor diversidad genética.

La reproducción sexual en líquenes permite la posibilidad de generar nuevos talos con fotobiontes mejor adaptados, sin embargo, cuando una pequeña población aislada solo contiene un tipo de MAT, los individuos pueden optar por mecanismos de reproducción asexual, reduciendo el coste de la reproducción sexual y evitando el problema de encontrar individuos compatibles. De igual forma, en especies de líquenes muy competitivos, como es el caso de *Xanthoria parietina*, se ha descrito otro tipo de estrategia reproductiva. Esta especie a pesar de tener una estructura génica del gen MAT puramente heterotálica (MAT1-2), su reproducción se da mediante homotalismo secundario de tipo unisexual (Scherrer et al. 2005) en el cual las progenies contienen únicamente un ideomorfo MAT1-2. Las esporas derivadas de la unisexualidad tienen ventajas sobre las estructuras de reproducción vegetativa (soredios/isidios) o esporas mitóticas (conidios) en relación a la tasa de

supervivencia, lo que podría ser una importante ventaja adaptativa, especialmente en condiciones adversas (Trapero-Casas et al. 2007).

8.4 DIVERSIDAD DE CLÚSTERES BIOSINTÉTICOS

Los líquenes pertenecientes a la clase Lecanoromycetes son conocidos por producir una extensa variedad de metabolitos secundarios. En el Capítulo 4, mediante el uso de herramientas para la detección de clústeres de genes biosintéticos (CGB), se han identificado los clústeres implicados en la producción de varios metabolitos. Los resultados muestran que el contenido y diversidad de estos clústeres de genes es variable entre los distintos taxones incluidos, encontrándose especies con un alto número de clústeres como *Evernia prunastri* (98 CGB) o *Pannoparmelia angustata* (78 CGB) y otras con poca diversidad como *Canoparmelia nairobiensis* (10 CGB) o *Bulbothrix sensibilis* (11 CGB). Este hallazgo es congruente con los resultados de la identificación CGB en otras especies secuenciadas a partir de cultivos aposimbioticos, donde el número de estos clústeres no sigue un patrón evolutivo, es decir, especies más cercanas evolutivamente no muestran un número similar de CGB (Park et al. 2013; Bertrand et al. 2018; Dal Grande et al. 2018; Calchera et al. 2019).

Las policétido sintasas (PKS) son las enzimas principales de estos CGB, se dividen en varios grupos dependiendo su estructura y organización de dominios enzimáticos (Yang et al. 2018). Las PKS no redcutoras (NR-PKS) son conocidas por ser responsables de la biosíntesis de gran cantidad de micotoxinas, pigmentos y otras moléculas en distintos linajes de hongos de Pezizomycotina (Huffman et al. 2010; Gallo et al. 2013; Liu et al. 2015). En el Capítulo 4 se evalúa el contenido de estas enzimas en Lecanoromycetes mediante búsqueda por homología (Khaldi et al. 2010) utilizando grupos conocidos de PKS implicados en la producción de diferentes compuestos. Los resultados muestran un árbol dividido en 8 grupos y con topologías concordantes con estudios previos (Liu et al. 2015).

Uno de los principales hallazgos se observa en el grupo VI de las NR-PKS, este grupo alberga PKS implicadas en la síntesis de distintas mycotoxinas conocidas

como la terrotonina (Hamed et al. 2019) o anditomin (Matsuda et al. 2014), así como la PKS de la producción del ácido úsnico, MPAS de *Cladonia uncialis* (Abdel- Hameed et al. 2016). El resultado de este análisis fue sorprendente, ya que solo las especies productoras de ácido úsnico contenían secuencias de PKS homólogas a MPAS. Este descubrimiento pone de manifiesto que la adquisición del clúster del ácido úsnico en Lecanoromycetes no puede ser explicado por transferencia horizontal de genes, sino más bien por una pérdida de dicho clúster en las especies no productoras (Campbell et al. 2013; Lind et al. 2017). En todas las especies productoras fueron encontradas los dos genes enzimáticos (MPAS y MPAO) descritos en Abdel-Hameed et al. 2016. La comparación de la estructura génica del CGB de ácido úsnico desveló una alta conservación en la sintenia de los genes en especies pertenecientes a la misma familia (Parmeliaceae). Todas las especies en las cuales se comparó la estructura génica mostraron la presencia de genes codificantes para factores de transcripción típicos de los clústeres metabólicos. Excepcionalmente, nuestro estudio también localiza una NR-PKS de *Lobaria pulmonaria* homóloga a MPAS, siendo concordante con los resultados encontrados en Calchera et al. 2019. La especie *Lobaria pulmonaria* ha sido considerada históricamente como no productora de ácido úsnico, sin embargo, un estudio llevado a cabo por Cansaran et al. 2008, detectaron ácido úsnico en *Lobaria pulmonaria* mediante técnicas de HPLC.

Otro hallazgo indirecto del análisis fue la identificación de las NR-PKS implicadas en la síntesis de melaninas fúngicas (Feng et al. 2001), sorprendentemente aquí se encontraron muchos genes PKS homólogos de casi todas las especies incluidas, lo cual indica que estas enzimas podrían estar implicadas en la síntesis de melaninas de este grupo. Sin embargo, esto necesitaría una confirmación mediante técnicas de expresión o RNA-Seq en futuros estudios.

Se conoce que la organización de las rutas metabólicas de metabolitos secundarios tiene importantes implicaciones en la adaptación de los hongos (Slot 2017). Por ejemplo, un buen número de estudios de genómica comparativa de hongos han aportado evidencia de la implicación de los procesos evolutivos de ganancia y perdida de genes, así como la transferencia horizontal de genes, en la formación de CGB de metabolitos secundarios en hongos (Khaldi et al. 2008; Slot et al. 2011; Campbell et al. 2013; Wisecaber et al. 2014; Wisecaber

et al. 2015). Los procesos evolutivos de ganancia y perdida de genes suelen darse con más frecuencia, lo cual proporciona una explicación clara del porque muchas de CGB de metabolitos secundarios no muestran patrones de distribución asociados a la taxonomía y pueden variar en gran media entre especies relacionadas evolutivamente o incluso en cepas o variedades de una misma especie (Ferodova et al. 2008; Gibbon et al. 2012; Gibbon et al. 2013; Lind et al. 2017). De igual manera, estos eventos o procesos evolutivos podrían estar vinculados a la ecología fúngica, ya que generalmente suelen ocurrir entre especies que parecen tener nichos ecológicos parecidos (Slot 2017). En el caso del ácido úsnico, a pesar de ser una molécula ampliamente estudiada en relación a la actividad farmacológica o como carácter taxonómico importante, los datos ecológicos que se tienen actualmente son insuficiente para explicar sus patrones de distribución taxonómica.

9.CONCLUSIONES

9. CONCLUSSIONS

9

CONCLUSIONES

1. Se constata que la secuenciación metagenómica es una herramienta alternativa y útil para la obtención del genoma de micobiointes. El porcentaje de integridad de los genomas recuperados es muy similar al obtenido en otros estudios, incluso en algunos casos, a los obtenidos de la secuenciación a partir de cultivos aposimbióticos.

2. Se han establecido gran parte de las relaciones filogenéticas de la familia Parmeliaceae que no estaban resueltas en previos estudios, lo que pone de manifiesto que el uso de los genes codificantes de una sola copia son un elemento útil para esclarecer filogenias difíciles.

3. Se constata que la arquitectura del locus MAT1-2, determinante del tipo de reproducción sexual, es similar en los hongos simbiointes estudiados a la de los demás de hongos pertenecientes a la subdivisión Pezizomycotina, estando flanqueado por los genes *sla2* y *apn1* y conteniendo en algunos casos genes mat auxiliares que podrían ser específicos de líquenes.

4. Se ha descubierto que la estrategia reproductiva más extendida en la clase Lecanomycetes es el heterotalismo, y no el homotalismo como se había propuesto en anteriores estudios.

5. Se confirma que los clústeres de biosíntesis de metabolitos secundarios en hongos Lecanoromycetes son muy diversos y su número variable, incluso en especies relacionadas evolutivamente.

6. Se establece que las enzimas policétido sintasas no reductoras implicadas en la síntesis de sustancias liquénicas se dividen en ocho grupos principales, los cuales comparten una organización específica de dominios enzimáticos.

7. Se ha detectado la presencia del clúster bioisintético del ácido úsnico únicamente en especies productoras de dicho compuesto, algo que solo puede ser explicado por repetidas perdidas de dicho clúster en las especies no productoras.

8. Se ha identificado nuevos genes que podrían estar implicados en la biosíntesis del ácido único, su sintenia u organización génica dentro del clúster está conservada en especies evolutivamente relacionadas.

CONCLUSIONS

1. It is verified that metagenomic sequencing is an alternative and useful approach for obtaining mycobiont genomes part. The genome completeness of binned metagenomes is in concordance with results from other studies, even the ones obtained from aposymbiotic cultures.

2. Previously unresolved relationships among Parmaliaceae have been established with well-support. That highlights the use of single copy genes is useful to clarify this kind of complex evolutionary relations.

3. It is proved that lichenized fungi MAT 1-2 locus synteny is quite similar to the one from other fungi belonging to Pezizomycotina. This locus is flanked by *sla2* and *apn1* genes and it may contain auxiliary MAT genes which could be specific of lichens forming fungi.

4. It is uncovered that in the Lecanoromycetes class heterothallism is wide-spread, instead of what have been established in previous studies about homothallism.

5. It is confirmed that diversity and variability of biosynthetic genes cluster is high, even in closed-related species.

6. It is bolstered that non-reducing poliketydes syntases enzymes involved in lichen substances synthesis are clustered in eight main group which share a specific enzymatic domain arrangement.

7 . It has been found that only usnic acid-producer species contain the biosynthetic gene cluster of usnic acid, discovery which can be explained just by repeating-losses of that cluster in non-producer species.

8. New genes which may be involved in usnic acid biosynthesis have been discovered. Their synteny or genomic organization is quite conserved in closed-related species.

BIBLIOGRAFÍA

BIBLIOGRAPHY

B

BIBLIOGRAFÍA

Abdel-Hameed, M., Bertrand, R. L., Piercey-Normore, M. D. & Sorensen, J. L. Putative identification of the usnic acid biosynthetic gene cluster by de novo whole-genome sequencing of a lichen-forming fungus. Fungal Biol. (2016). doi:10.1016/j.funbio.2015.10.009

Ai, B. & Kang, M. How many genes are needed to resolve phylogenetic incongruence? Evol. Bioinforma. (2015). doi:10.4137/EBO.S26047

Aklitsch, W., Baral, H. O., Lücking, R., Lumbsch, H. T. Syllabus of Plant Families - A. Engler's Syllabus der Pflanzenfamilien Part 1/2: (Schweizerbart Science Publishers, 2016).

Alföldi, J. & Lindblad-Toh, K. Comparative genomics as a tool to understand evolution and disease. Genome Research (2013). doi:10.1101/gr.157503.113

Allio, R. et al. Whole Genome Shotgun Phylogenomics Resolves the Pattern and Timing of Swallowtail Butterfly Evolution. Syst. Biol. (2019). doi:10.1093/sysbio/syz030

Alors, D. et al. Panmixia and dispersal from the Mediterranean Basin to Macaronesian Islands of a macrolichen species. Sci. Rep. 7, (2017).

Alors, D., Lumbsch, H.T., Divakar, P.K., Leavitt, S.D. & Crespo, A. An integrative approach for understanding diversity in the Punctelia rudecta species complex (Parmeliaceae, Ascomycota). Plos One (2016). doi:10.1038/s41598-019-44881-1

Altschul, S. F. et al. Basic local alignment search tool. J. Mol. Biol. 215, 403–410 (1990).

Armaleo, D., Sun, X. & Culberson, C. Insights from the first putative biosynthetic gene cluster for a lichen depside and depsidone. Mycologia (2011). doi:10.3852/10-335

Bálint, M. et al. Cryptic biodiversity loss linked to global climate change. Nat. Clim. Chang. (2011). doi:10.1038/nclimate1191

Bary, A. de. DIE ERSCHEINUNG DER SYMBIOSE. in Die Erscheinung der Symbiose (2019). doi:10.1515/9783111471839-002

Bass, D., Stentiford, G. D., Littlewood, D. T. J. & Hartikainen, H. Diverse Applications of Environmental DNA Methods in Parasitology. Trends in Parasitology (2015). doi:10.1016/j.pt.2015.06.013

Bertrand, R. L., Abdel-Hameed, M. & Sorensen, J. L. Lichen Biosynthetic Gene Clusters Part II: Homology Mapping Suggests a Functional Diversity. J. Nat. Prod. (2018). doi:10.1021/acs.jnatprod.7b00770

Bertrand, R. L., Abdel-Hameed, M. & Sorensen, J. L. Lichen Biosynthetic Gene Clusters. Part I. Genome Sequencing Reveals a Rich Biosynthetic Potential. J. Nat. Prod. (2018). doi:10.1021/acs.jnatprod.7b00769

Beukeboom, L. W. & Perrin, N. The evolution of sex determination. Oxford University Press (2014). doi:10.1093/acprof:oso/9780199657148.001.0001

Bik, H. M. et al. Sequencing our way towards understanding global eukaryotic biodiversity. Trends in Ecology and Evolution (2012). doi:10.1016/j.tree.2011.11.010

Billiard, S., López-Villavicencio, M., Hood, M. E. & Giraud, T. Sex, outcrossing and mating types: Unsolved questions in fungi and beyond. Journal of Evolutionary Biology 25, 1020–1038 (2012).

Bingle, L. E. H., Simpson, T. J. & Lazarus, C. M. Ketosynthase domain probes identify two subclasses of fungal polyketide synthase genes. Fungal Genet. Biol. (1999). doi:10.1006/fgbi.1999.1115

Blanco, O., Crespo, A., Elix, J. A., Hawksworth, D. L. & Lumbsch, H. T. A molecular phylogeny and a new classification of parmelioid lichens containing Xanthoparmelia-type lichenan (Ascomycota: Lecanorales). Taxon (2004). doi:10.2307/4135563

Blin, K. et al. AntiSMASH 4.0 - improvements in chemistry prediction and gene cluster boundary identification. Nucleic Acids Res. (2017). doi:10.1093/nar/gkx319

Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120 (2014).

Boustie, J. & Grube, M. Lichens—a promising source of bioactive secondary metabolites. Plant Genet. Resour. (2005). doi:10.1079/pgr200572

Brady, S. F., Simmons, L., Kim, J. H. & Schmidt, E. W. Metagenomic approaches to natural products from free-living and symbiotic organisms. Natural Product Reports (2009). doi:10.1039/b817078a

Brown, D. W., Butchko, R. A. E., Busman, M. & Proctor, R. H. Identification of gene clusters associated with fusaric acid, fusarin, and perithecial pigment production in Fusarium verticillioides. Fungal Genet. Biol. (2012). doi:10.1016/j.fgb.2012.05.010

Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60 (2014).

Büdel, B. & Scheidegger, C. Thallus morphology and anatomy. Lichen Biol. 42–70 (2008). doi:10.1017/CBO9780511790478.005

Burki, F. et al. Untangling the early diversification of eukaryotes: A phylogenomic study of the evolutionary origins of centrohelida, haptophyta and cryptista. Proc. R. Soc. B Biol. Sci. (2016). doi:10.1098/rspb.2015.2802

Bushley, K. E. & Turgeon, B. G. Phylogenomics reveals subfamilies of fungal nonribosomal peptide synthetases and their evolutionary relationships. BMC Evol. Biol. (2010). doi:10.1186/1471-2148-10-26

Calchera, A., Dal Grande, F., Bode, H. B. & Schmitt, I. Biosynthetic Gene Content of the 'Perfume Lichens' Evernia prunastri and Pseudevernia furfuracea. Molecules (2019). doi:10.3390/molecules24010203

Calcott, M. J., Ackerley, D. F., Knight, A., Keyzers, R. A. & Owen, J. G. Secondary metabolism in the lichen symbiosis. Chemical Society Reviews (2018). doi:10.1039/c7cs00431a

Campbell, M. A., Rokas, A. & Slot, J. C. Horizontal transfer and death of a fungal secondary metabolic gene cluster. Genome Biol. Evol. (2012). doi:10.1093/gbe/evs011

Campbell, M. A., Staats, M., van Kan, J. A. L., Rokas, A. & Slot, J. C. Repeated loss of an anciently horizontally transferred gene cluster in Botrytis . Mycologia (2013). doi:10.3852/12-390

Cannon, P. & Kirk, P. Fungal Families of the World. CAB International. (2007). doi:10/1017/S0024282908009158

Cansaran, D., Aras, S. & Atakol, O. Determination of Usnic Acid Content in Some Lichen Species Found in Anatolia. J. Appl. Biol. Sci. 2, 41–44 (2008).

Cernava, T. et al. Deciphering functional diversification within the lichen microbiota by meta-omics. Microbiome (2017). doi:10.1186/s40168-017-0303-5

Chang, P. K., Horn, B. W. & Dorner, J. W. Sequence breakpoints in the aflatoxin biosynthesis gene cluster and flanking regions in nonaflatoxigenic Aspergillus flavus isolates. Fungal Genet. Biol. (2005). doi:10.1016/j.fgb.2005.07.004

Chang, P.-K. The Aspergillus parasiticus protein AFLJ interacts with the aflatoxin pathway-specific regulator AFLR. Mol. Genet. Genomics (2003). doi:10.1007/s00438-003-0809-3

Chiang, Y. M. et al. A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in aspergillus nidulans. J. Am. Chem. Soc. (2009). doi:10.1021/ja8088185

Chiara, M. et al. Genome sequencing of multiple isolates highlights subtelomeric genomic diversity within Fusarium Fujikuroi. Genome Biol. Evol. (2015). doi:10.1093/gbe/evv198

Choi, J. J. & Kim, S. H. A genome Tree of Life for the Fungi kingdom. Proc. Natl. Acad. Sci. U. S. A. (2017). doi:10.1073/pnas.1711939114

Cocchietto, M., Skert, N., Nimis, P. & Sava, G. A review on usnic acid, an interesting natural compound. Naturwissenschaften (2002). doi:10.1007/s00114-002-0305-3

Cooper, G. M. The Cell: A Molecular Approach. 2nd edition. (Sunderland (MA) Sinauer Associates, 2000).

Coppin, E., Debuchy, R., Arnaise, S. & Picard, M. Mating types and sexual development in filamentous ascomycetes. Microbiol. Mol. Biol. Rev. 61, 411–28 (1997).

Cox, R. J. & Simpson, T. J. Chapter 3 Fungal Type I Polyketide Synthases. Methods in Enzymology (2009). doi:10.1016/S0076-6879(09)04603-5

Crespo, A. & Cubero, O. F. A molecular approach to the circumscription and evaluation of some genera segregated from Parmelia s. lat. Lichenologist (1998). doi:10.1006/lich.1998.0156

Crespo, A. & Lumbsch, H. T. Cryptic species in lichen-forming fungi. IMA Fungus (2011). doi:10.5598/imafungus.2010.01.02.09

Crespo, A. et al. Phylogenetic generic classification of parmelioid lichens (Parmeliaceae, Ascomycota) based on molecular, morphological and chemical evidence. Taxon (2010).

Crespo, A. et al. Testing morphology-based hypotheses of phylogenetic relationships in Parmeliaceae (Ascomycota) using three ribosomal markers and the nuclear RPB1 gene. Mol. Phylogenet. Evol. (2007). doi:10.1016/j.ympev.2006.11.029

Crespo, A., Divakar, P. K. & Lumbsch, H. T. Hyperdiversity closer to animals than to plants. Tree Life (2014).

Culberson, W. L., Brodo, I. M. & Hawksworth, D. L. Alectoria and Allied Genera in North America. Bryologist (2006). doi:10.2307/3242363

Dal Grande, F. et al. Adaptive differentiation coincides with local bioclimatic conditions along an elevational cline in populations of a lichen-forming fungus. BMC Evol. Biol. 17, (2017).

Dal Grande, F. et al. Molecular phylogeny and symbiotic selectivity of the green algal genus Dictyochloropsis s.l. (Trebouxiophyceae): A polyphyletic and widespread group forming photobiont-mediated guilds in the lichen family Lobariaceae. New Phytol. (2014). doi:10.1111/nph.12678

Dal Grande, F. et al. The draft genome of the lichen-forming fungus Lasallia hispanica (Frey) Sancho & A. Crespo. Lichenologist 50, 329–340 (2018).

de Paz, G. A., Cubas, P., Crespo, A., Elix, J. A. & Lumbsch, H. T. Transoceanic dispersal and subsequent diversification on separate continents shaped diversity of the Xanthoparmelia pulla group (Ascomycota). PLoS One (2012). doi:10.1371/journal.pone.0039683

de Vries, R. P. et al. Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus Aspergillus. Genome Biol. 18, (2017).

Debuchy, R. & Turgeon, B. Mating-type structure, evolution, and function in Euascomycetes. Growth, Differ. Sex. 46 (2006). doi:10.1007/3-540-28135-5_15

Del-Prado, R. et al. Molecular phylogeny and historical biogeography of the lichen-forming fungal genus Flavoparmelia (Ascomycota: Parmeliaceae). Taxon (2013). doi:10.12705/625.22

Del-Prado, R., Divakar, P. K. & Crespo, A. Using genetic distances in addition to ITS molecular phylogeny to identify potential species in the Parmotrema reticulatum complex: A case study. Lichenologist (2011). doi:10.1017/S0024282911000582

Denver, D. R., Brown, A. M. V., Howe, D. K., Peetz, A. B. & Zasada, I. A. Genome Skimming: A Rapid Approach to Gaining Diverse Biological Insights into Multicellular Pathogens. PLoS Pathog. (2016). doi:10.1371/journal.ppat.1005713

Desjardins, C. A. et al. Comparative genomic analysis of human fungal pathogens causing paracoccidioidomycosis. PLoS Genet. 7, (2011).

Dighton, J. & White, J. F.) 53–62 (CRC Press Taylor & Francis Group, 2017). doi:https://doi.org/10.1201/9781315119496

Dimijian, G. G. Evolving Together: The Biology of Symbiosis, Part 2. Baylor Univ. Med. Cent. Proc. (2000). doi:10.1080/08998280.2000.11927712

Divakar, P. K. et al. Evolution of complex symbiotic relationships in a morphologically derived family of lichen-forming fungi. New Phytol. (2015). doi:10.1111/nph.13553

Divakar, P. K. et al. Parallel Miocene dispersal events explain the cosmopolitan distribution of the Hypogymnioid lichens. J. Biogeogr. (2019). doi:10.1111/jbi.13554

Divakar, P. K., Kauff, F., Crespo, A., Leavitt, S. D. & Lumbsch, H. T. Understanding phenotypical character evolution in parmelioid lichenized fungi (Parmeliaceae, ascomycota). PLoS One (2013). doi:10.1371/journal.pone.0083115

Dyer, P. S. Evolutionary Biology: Genomic Clues to Original Sex in Fungi. Current Biology 18, (2008).

Dyer, P.S., Inderbitzin, P, Debuchy, R. et al. Growth, Differentiation and Sexuality. in The Mycota 1, 81–101 (2006).

Dyer, P.S., Inderbitzin, P, Debuchy, R. Mating-type structure, function, regulation and evolution in the Pezizomycotina. in Growth, Differentiation and Sexuality, 3rd Edition, The Mycota I. (ed. J. Wendland) (2016). doi:10.1007/978-3-319-25844-7

Ebersberger, I. et al. A consistent phylogenetic backbone for the fungi. Mol. Biol. Evol. (2012). doi:10.1093/molbev/msr285

Elgar, G. et al. Generation and analysis of 25 mb of genomic DNA from the pufferfish Fugu rubripes by sequence scanning. Genome Res. (1999). doi:10.1101/gr.9.10.960

Elix, J. & Elix, J. A. Progress in the Generic Delimitation of Parmelia Sensu Lato Lichens (Ascomycotina: Parmeliaceae) and a Synoptic Key to the Parmeliaceae. Bryologist (1993). doi:10.2307/3243867

Elix, J. A., Whitton, A. A. & Sargent, M. V. Recent Progress in the Chemistry of Lichen Substances. in Fortschritte der Chemie organischer Naturstoffe / Progress in the Chemistry of Organic Natural Products 103–234 (Springer, 1984). doi:https://doi.org/10.1007/978-3-7091-8717-3_2

Elix, J. Biochemistry and secondary metabolites. in Lichen Biology 155–180 (Cambridge University Press, 1996).

Emms, D. M. & Kelly, S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. (2015). doi:10.1186/s13059-015-0721-2

Fedorova, N. D. et al. Genomic islands in the pathogenic filamentous fungus Aspergillus fumigatus. PLoS Genet. (2008). doi:10.1371/journal.pgen.1000046

Feng, B. et al. Molecular cloning and characterization of WdPKS1, a gene involved in dihydroxynaphthalene melanin biosynthesis and virulence in Wangiella (Exophiala) dermatitidis. Infect. Immun. (2001). doi:10.1128/IAI.69.3.1781-1794.2001

Ferencova, Z. Estudio morfológico comparado de los caracteres generativos en relación con los linages monofiléticos de la familia Parmeliaceae. (Universidad Complutense de Madrid, 2012).

Feretzaki, M. & Heitman, J. Unisexual Reproduction Drives Evolution of Eukaryotic Microbial Pathogens. PLoS Pathog. 9, (2013).

Fernández-Moriano, C., Gómez-Serranillos, M. P. & Crespo, A. Antioxidant potential of lichen species and their secondary metabolites. A systematic review. Pharmaceutical Biology (2016). doi:10.3109/13880209.2014.1003354

Feuerer, T. & Hawksworth, D. L. Biodiversity of lichens, including a world-wide analysis of checklist data based on Takhtajan's floristic regions. Biodivers. Conserv. (2007). doi:10.1007/s10531-006-9142-6

Fischer, M. J. C., Rustenhloz, C., Leh-Louis, V. & Perrière, G. Molecular and functional evolution of the fungal diterpene synthase genes. BMC Microbiol. (2015). doi:10.1186/s12866-015-0564-8

Fu, C., Sun, S., Billmyre, R. B., Roach, K. C. & Heitman, J. Unisexual versus bisexual mating in Cryptococcus neoformans: Consequences and biological impacts. Fungal Genet. Biol. 78, 65–75 (2014).

Gaffoor, I. & Trail, F. Characterization of two polyketide synthase genes involved in zearalenone biosynthesis in Gibberella zeae. Appl. Environ. Microbiol. (2006). doi:10.1128/AEM.72.3.1793-1799.2006

Galagan, J. E. et al. Sequencing of Aspergillus nidulans and comparative analysis with A. fumigatus and A. oryzae. Nature (2005). doi:10.1038/nature04341

Gallo, A., Ferrara, M. & Perrone, G. Phylogenetic study of polyketide synthases and nonribosomal peptide synthetases involved in the biosynthesis of mycotoxins. Toxins (Basel). (2013). doi:10.3390/toxins5040717

Gardiner, D. M. & Howlett, B. J. Bioinformatic and expression analysis of the putative gliotoxin biosynthetic gene cluster of Aspergillus fumigatus. FEMS Microbiol. Lett. (2005). doi:10.1016/j.femsle.2005.05.046

Gargas, A., DePriest, P. T., Grube, M. & Tehler, A. Multiple origins of lichen symbioses in fungi suggested by SSU rDNA phylogeny. Science (80-.). (1995). doi:10.1126/science.7770775

Gaya, E. et al. The adaptive radiation of lichen-forming Teloschistaceae is associated with sunscreening pigments and a bark-to-rock substrate shift. Proc. Natl. Acad. Sci. 112, 11600–11605 (2015).

German National Academy of Sciences Leopoldina - Challenges and Opportunities of Integrative Taxonomy for Researchand Society. 0–56 (2014).

Gibbons, J. G. & Rokas, A. The function and evolution of the Aspergillus genome. Trends in Microbiology (2013). doi:10.1016/j.tim.2012.09.005

Gibbons, J. G. et al. Global transcriptome changes underlying colony growth in the opportunistic human pathogen Aspergillus fumigatus. Eukaryot. Cell (2012). doi:10.1128/EC.05102-11

Gillian Turgeon, B. et al. Cloning and analysis of the mating type genes from Cochliobolus heterostrophus. MGG Mol. Gen. Genet. 238, 270–284 (1993).

Grandjean, F. et al. Rapid recovery of nuclear and mitochondrial genes by genome skimming from Northern Hemisphere freshwater crayfish. Zool. Scr. (2017). doi:10.1111/zsc.12247

Green, T. G. A., Sancho, L. G., Türk, R., Seppelt, R. D. & Hogg, I. D. High diversity of lichens at 84°S, Queen Maud Mountains, suggests preglacial survival of species in the Ross Sea region, Antarctica. Polar Biol. 34, 1211–1220 (2011).

Greshake, B. et al. Potential and pitfalls of eukaryotic metagenome skimming: A test case for lichens. Mol. Ecol. Resour. (2016). doi:10.1111/1755-0998.12463

Grewe, F., Lagostina, E., Wu, H., Printzen, C. & Thorsten Lumbsch, H. Population genomic analyses of RAD sequences resolves the phylogenetic relationship of the lichen-forming fungal species Usnea antarctica and Usnea aurantiacoatra. MycoKeys (2018). doi:10.3897/mycokeys.43.29093

Grube M, W. M. Lichenized Fungi and the Evolution of Symbiotic Organization. Microbiol. Spectr. 4, (2017).

Grube, M. et al. Exploring functional contexts of symbiotic sustain within lichen-associated bacteria by comparative omics. ISME J. (2015). doi:10.1038/ismej.2014.138

Grube, M. et al. The next generation fungal diversity researcher. Fungal Biology Reviews (2017). doi:10.1016/j.fbr.2017.02.001

Grube, M., Cardinale, M., De Castro, J. V., Müller, H. & Berg, G. Species-specific structural and functional diversity of bacterial communities in lichen symbioses. ISME J. (2009). doi:10.1038/ismej.2009.63

Hagen, D. C., Bruhn, L., Westby, C. a & Sprague, G. F. Transcription of alpha-specific genes in Saccharomyces cerevisiae: DNA sequence requirements for activity of the coregulator alpha 1. Mol. Cell. Biol. 13, 6866–6875 (1993).

Hamed, A. et al. Terretonin O: a new meroterpenoid from Aspergillus terreus. Natural Product Research (2018). doi:10.1080/14786419.2018.1544977

Harris, R. S. Improved pariwise alignment of genomic DNA. The Pennsylvania State University (2007).

Hawksworth, D. L. & Honegger, R. The lichen thallus: a symbiotic phenotype of nutritionally specialized fungi and its response to gall producers. in Plant galls organisms, interactions, populations (1994).

Hawksworth, D. L. Lichen Chemotaxonomy. in Lichenology: Progress and Problems. (eds. DH, B., D.L, H. & R.H, B.) 139–184 (Academic Press, 1976).

Hawksworth, D. L. Lichenization: The origins of a fungal life-style. in Recent Advances in Lichenology: Modern Methods and Approaches in Lichen Systematics and Culture Techniques, Volume 2 1–10 (2015). doi:10.1007/978-81-322-2235-4_1

Hawksworth, D. L. The lichenicolous fungi of Great Britain and Ireland: An overview and annotated checklist. Lichenologist (2003). doi:10.1016/S0024-2829(03)00027-6

Hawksworth, D. L. The magnitude of fungal diversity: The 1.5 million species estimate revisited. in Mycological Research (2001). doi:10.1017/S0953756201004725

Heiss, A. A. et al. Combined morphological and phylogenomic re-examination of malawimonads, a critical taxon for inferring the evolutionary history of eukaryotes. R. Soc. Open Sci. (2018). doi:10.1098/rsos.171707

Hestmark, G. Sex, size, competition and escape-strategies of reproduction and dispersal in Lasallia pustulata (Umbilicariaceae, Ascomycetes). Oecologia 92, 305–312 (1992).

Holt, C. & Yandell, M. MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics 12, (2011).

Honegger, R. & Zippler, U. Mating systems in representatives of Parmeliaceae, Ramalinaceae and Physciaceae (Lecanoromycetes, lichen-forming ascomycetes). Mycol. Res. 111, 424–432 (2007).

Honegger, R., Zippler, U., Gansner, H. & Scherrer, S. Mating systems in the genus Xanthoria (lichen-forming ascomycetes). Mycol. Res. 108, 480–488 (2004).

Huang J.P., Kraichak, E., Leavitt, S.D., Nelsen, M.P. & Lumsbch, H.T. Accelerated diversifications in three diverse families of morphologically complex lichen forming fungi link to major historical events. Sci. Rep. (2019). doi:10/1038/s41598-019-44881-1

Huang, J.P., Leavitt, S.D. & Lumbsch, H.T. Testing the impact of effective population size on speciation rates - A negative correlation or lack thereof in lichenized fungi. Sci. Rep. (2018). doi: 10.1038/s41598-018-24120-971/journal.pone.0146537

Huffman, J., Gerber, R. & Du, L. Review Recent Advancements in the Biosynthetic Mechanisms for Polyketide-Derived Mycotoxins. Biopolymers (2010). doi:10.1002/bip.21483

Huneck, S. & Yoshimura, I. Identification of Lichen Substances. (Springer, Berlin, Heidelberg, 1996). doi:https://doi.org/10.1007/978-3-642-85243-5_2

Huson, D. H. et al. MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. PLoS Comput. Biol. 12, (2016).

Idnurm, A., Walton, F. J., Floyd, A. & Heitman, J. Identification of the sex genes in an early diverged fungus. Nature 451, 193–196 (2008).

Jacobson, E. S. Pathogenic roles for fungal melanins. Clinical Microbiology Reviews (2000). doi:10.1128/CMR.13.4.708-717.2000

Jahns, H. M., Herold, K. & Beltman, H. A. Chronological Sequence, Synchronization and Induction of the Development of Fruit Bodies in Cladonia furcata var. furcata (Huds.) Schrad. Nov. Hedwigia 30, 469–526 (1979).

Jaklitsch, W., Baral, H. O., Lücking, R., Lumbsch, H. T. Syllabus of Plant Families. Part 1/2 Ascomycota. (Borntraeger Science Publishers, 2016).

Jeon, J. R., Baldrian, P., Murugesan, K. & Chang, Y. S. Laccase-catalysed oxidations of naturally occurring phenols: From in vivo biosynthetic pathways to green synthetic applications. Microbial Biotechnology (2012). doi:10.1111/j.1751-7915.2011.00273.x

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A. & Jermiin, L. S. ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat. Methods (2017). doi:10.1038/nmeth.4285

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 44, D457–D462 (2016).

Katz, L. A. & Grant, J. R. Taxon-rich phylogenomic analyses resolve the eukaryotic tree of life and reveal the power of subsampling by sites. Syst. Biol. (2015). doi:10.1093/sysbio/syu126

Keller, N. P. Fungal secondary metabolism: regulation, function and drug discovery. Nature Reviews Microbiology (2019). doi:10.1038/s41579-018-0121-1

Keller, N. P. Translating biosynthetic gene clusters into fungal armor and weaponry. Nature Chemical Biology (2015). doi:10.1038/nchembio.1897

Keller, N. P., Turner, G. & Bennett, J. W. Fungal secondary metabolism - From biochemistry to genomics. Nature Reviews Microbiology (2005). doi:10.1038/nrmicro1286

Khaldi, N. & Wolfe, K. H. Evolutionary Origins of the Fumonisin Secondary Metabolite Gene Cluster in Fusarium verticillioides and Aspergillus niger . Int. J. Evol. Biol. (2011). doi:10.4061/2011/423821

Khaldi, N. et al. SMURF: Genomic mapping of fungal secondary metabolite clusters. Fungal Genet. Biol. (2010). doi:10.1016/j.fgb.2010.06.003

Kidron, G. J. & Temina, M. Lichen colonization on cobbles in the Negev Desert following 15 years in the field. Geomicrobiol. J. (2010). doi:10.1080/01490450903490805

Kimura, M. The neutral theory of molecular evolution: a review of recent evidence. Jpn. J. Genet. (1991).

Kraichak, E., Huang, J. P., Nelsen, M., Leavitt, S. D. & Thorsten Lumbsch, H. A revised classification of orders and families in the two major subclasses of Lecanoromycetes (Ascomycota) based on a temporal approach. Bot. J. Linn. Soc. (2018). doi:10.1093/botlinnean/boy060

Kroken, S., Glass, N. L., Taylor, J. W., Yoder, O. C. & Turgeon, B. G. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. Proc. Natl. Acad. Sci. (2003). doi:10.1073/pnas.2532165100

Kronstad, J. W. & Staben, C. MATING TYPE IN FILAMENTOUS FUNGI. Annu. Rev. Genet. 31, 245–276 (1997).

Kück, P. & Longo, G. C. FASconCAT-G: Extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. Front. Zool. 11, (2014).

Leavitt, S. D., Divakar, P. K., Crespo, A. & Lumbsch, H. T. A Matter of Time – Understanding the Limits of the Power of Molecular Data for Delimiting Species Boundaries. Herzogia (2016). doi:10.13158/heia.29.2.2016.479

Lee, S. C. et al. Microsporidia Evolved from Ancestral Sexual Fungi. Curr. Biol. 18, 1675–1679 (2008).

Lee, S. C., Ni, M., Li, W., Shertz, C. & Heitman, J. The Evolution of Sex: a Perspective from the Fungal Kingdom. Microbiol. Mol. Biol. Rev. 74, 298–340 (2010).

Legouin, B. et al. Specialized Metabolites of the Lichen Vulpicida pinastri Act as Photoprotective Agents. Molecules (2017). doi:10.3390/molecules22071162

Lemmon, E. M. & Lemmon, A. R. High-Throughput Genomic Data in Systematics and Phylogenetics. Annu. Rev. Ecol. Evol. Syst. (2013). doi:10.1146/annurev-ecolsys-110512-135822

Levin, R. A., Whelan, A. & Miller, J. S. The utility of nuclear conserved ortholog set II (COSII) genomic regions for species-level phylogenetic inference in Lycium (Solanaceae). Mol. Phylogenet. Evol. (2009). doi:10.1016/j.ympev.2009.08.016

Lewin, R. A. Symbiosis and Parasitism: Definitions and Evaluations. Bioscience 32, 254–260 (1982).

Lim, F. Y. et al. Genome-Based Cluster Deletion Reveals an Endocrocin Biosynthetic Pathway in Aspergillus fumigatus. Appl. Environ. Microbiol. (2012). doi:10.1128/aem.07710-11

Lin, X. & Heitman, J. Mechanisms of homothallism in fungi and transitions between heterothallism and homothallism. Sex fungi Mol. Determ. Evol. Implic. 35–57 (2007). doi:10.1128/9781555815837.ch3

Lind, A. L. et al. Drivers of genetic diversity in secondary metabolic gene clusters within a fungal species. PLoS Biol. (2017). doi:10.1371/journal.pbio.2003583

Lind, A. L. et al. Examining the Evolution of the Regulatory Circuit Controlling Secondary Metabolism and Development in the Fungal Genus Aspergillus. PLoS Genet. (2015). doi:10.1371/journal.pgen.1005096

Liu, L. et al. Bioinformatical analysis of the sequences, structures and functions of fungal polyketide synthase product template domains. Sci. Rep. (2015). doi:10.1038/srep10463

Liu, Y. et al. Phylogenomic analyses predict sistergroup relationship of nucleariids and Fungi and paraphyly of zygomycetes with significant support. BMC Evol. Biol. (2009). doi:10.1186/1471-2148-9-272

Louca, S. et al. High taxonomic variability despite stable functional structure across microbial communities. Nat. Ecol. Evol. (2017). doi:10.1038/s41559-016-0015

Ludwig, L. R., Summerfield, T. C., Lord, J. M. & Singh, G. Characterization of the mating-type locus (MAT) reveals a heterothallic mating system in Knightiella splachnirima. Lichenologist 49, 373–385 (2017).

Lumbsch, H. T. & Huhndorf, S. M. Myconet Volume 14. Part One. Outline of Ascomycota—2009. Part Two. Notes on Ascomycete Systematics. Nos. 4751–5113. Fieldiana Life Earth Sci. (2010). doi:10.3158/1557.1

Lumbsch, H. T. T. & Rikkinen, J. Chapter 4 Evolution of Lichens. in The Fungal Community: Its Organization and Role in the Ecosystem, Fourth Edition eds. Lumbsch, H. T. The taxonomic use of metabolic data in lichen-forming fungi. in Chemical Fungal Taxonomy. (eds. Frisvad, J. C., Bridge, P. D. & Arora, D. K.) 345–387 (CRC Press, 1998).

Lumbsch, H. T., Hipp, A. L., Divakar, P. K., Blanco, O. & Crespo, A. Accelerated evolutionary rates in tropical and oceanic parmelioid lichens (Ascomycota). BMC Evol. Biol. (2008). doi:10.1186/1471-2148-8-257

Lutzoni, F., Pagel, M. & Reeb, V. Major fungal lineages are derived from lichen symbiotic ancestors. Nature (2001). doi:10.1038/35082053

Malé, P. J. G. et al. Genome skimming by shotgun sequencing helps resolve the phylogeny of a pantropical tree family. Mol. Ecol. Resour. (2014). doi:10.1111/1755-0998.12246

Malinsky, M. et al. Whole genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. bioRxiv (2018). doi:10.1101/143859

Matsuda, Y., Wakimoto, T., Mori, T., Awakawa, T. & Abe, I. Complete biosynthetic pathway of anditomin: Nature's sophisticated synthetic route to a complex fungal meroterpenoid. J. Am. Chem. Soc. (2014). doi:10.1021/ja508127q

Mattsson, J. E. & Wedin, M. Phylogeny of the Parmeliaceae - DNA data versus morphological data. Lichenologist (1998). doi:10.1006/lich.1998.0143

Medina, E. M., Jones, G. W. & Fitzpatrick, D. A. Reconstructing the fungal tree of life using phylogenomics and a preliminary investigation of the distribution of yeast prion-like proteins in the fungal kingdom. J. Mol. Evol. (2011). doi:10.1007/s00239-011-9461-4

Meiser, A., Otte, J., Schmitt, I. & Grande, F. D. Sequencing genomes from mixed DNA samples - Evaluating the metagenome skimming approach in lichenized fungi. Sci. Rep. 7, (2017).

Menardo, F., Praz, C. R., Wicker, T. & Keller, B. Rapid turnover of effectors in grass powdery mildew (Blumeria graminis). BMC Evol. Biol. (2017). doi:10.1186/s12862-017-1064-2

Menardo, F., Wicker, T. & Keller, B. Reconstructing the Evolutionary History of Powdery Mildew Lineages (Blumeria graminis) at Different Evolutionary Time Scales with NGS Data. Genome Biology and Evolution (2017). doi:10.1093/gbe/evx008

Metzenberg, R. L. & Glass, N. L. Mating type and mating strategies in Neurospora. BioEssays 12, 53–59 (1990).

Meusemann, K. et al. A phylogenomic approach to resolve the arthropod tree of life. Mol. Biol. Evol. (2010). doi:10.1093/molbev/msq130

Miao, V. P. W., LeGal-Coeëffet, M. F., Brown, D. W. & Davies, J. E. Polyketide synthase genes in lichens. in Sixth International Mycological Congress (IMC6). Abstracts 84 (1998).

Misof, B. et al. Phylogenomics resolves the timing and pattern of insect evolution. Science (80-.). (2014). doi:10.1126/science.1257570

Misof, B. et al. Selecting informative subsets of sparse supermatrices increases the chance to find correct trees. BMC Bioinformatics 14, (2013).

Mistry, J., Finn, R. D., Eddy, S. R., Bateman, A. & Punta, M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. (2013). doi:10.1093/nar/gkt263

Mitchell, A. L. et al. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. (2015). doi:10.1093/nar/gkv1344

Molina, M. D. C. & Crespo, A. Comparison of development of axenic cultures of five species of lichen-forming fungi. Mycol. Res. 104, 595–602 (2000).

Müller, E. Die schweizerischen Arten der Gattung Ophiobolus Riess. Ber Schweiz Bot Ges 62, 307–339 (1952).

Murtagh, G. J., Crittenden, P. D. & Dyer, P. S. Reproductive systems: Sex and the single lichen. Nature 404, 564–564 (2000).

Nash, T. . Lichen Biology. Igarss 2014 (Cambridge University Press, 2008).

Nash, T. Lichen Biology. Cambridge Univ. Press (2008). doi:10.1007/s13398-014-0173-7.2

Nelson, M. A. Mating systems in ascomycetes: A romp in the sac. Trends in Genetics 12, 69–74 (1996).

Nguyen, L. T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274 (2015).

Ni, M., Feretzaki, M., Sun, S., Wang, X. & Heitman, J. Sex in Fungi. Annu. Rev. Genet. 45, 405–430 (2011).

Nicholson, T. P. et al. Design and utility of oligonucleotide gene probes for fungal polyketide synthases. Chem. Biol. (2001). doi:10.1016/S1074-5521(00)90064-4

Nimis, P. L. A critical appraisal of modern generic concepts in lichenology. Lichenologist (1998). doi:10.1006/lich.1998.0152

Normark, B. B., Judson, O. P. & Moran, N. A. Genomic signatures of ancient asexual lineages. in Biological Journal of the Linnean Society 79, 69–84 (2003).

Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. MetaSPAdes: A new versatile metagenomic assembler. Genome Res. 27, 824–834 (2017).

Olive, L. S. On the evolution of heterothallism in fungi. Am. Nat. 92, 233–251 (1958).

Osbourn, A. Secondary metabolic gene clusters: Evolutionary toolkits for chemical innovation. Trends in Genetics (2010). doi:10.1016/j.tig.2010.07.001

Otto, S. Sexual Reproduction and the Evolution of Sex. Nat. Educ. 1, 1(1):182 (2008).

Ovcharenko, I., Loots, G. G., Hardison, R. C., Miller, W. & Stubbs, L. zPicture: Dynamic alignment and visualization tool for analyzing conservation profiles. Genome Res. 14, 472–477 (2004).

Park, S.-Y. et al. Draft Genome Sequence of Cladonia macilenta KoLRI003786, a Lichen-Forming Fungus Producing Biruloquinone. Genome Announc. (2013). doi:10.1128/genomea.00695-13

Park, S.-Y. et al. Draft Genome Sequence of Endocarpon pusillum Strain KoLRILF000583. Genome Announc. (2014). doi:10.1128/genomea.00452-14

Park, S.-Y. et al. Draft Genome Sequence of Lichen-Forming Fungus Caloplaca flavorubescens Strain KoLRI002931. Genome Announc. (2013). doi:10.1128/genomea.00678-13

Park, S.-Y. et al. Draft Genome Sequence of Umbilicaria muehlenbergii KoLRILF000956, a Lichen-Forming Fungus Amenable to Genetic Manipulation. Genome Announc. (2014). doi:10.1128/genomea.00357-14

Park, S.-Y. S.-Y. et al. Draft Genome Sequence of Lichen-Forming Fungus Cladonia metacorallifera Strain KoLRI002260. Genome Announc. (2014). doi:10.1128/genomea.01065-13

Parra, G., Bradnam, K. & Korf, I. CEGMA: A pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23, 1061–1067 (2007).

Patron, N. J. et al. Origin and distribution of epipolythiodioxopiperazine (ETP) gene clusters in filamentous ascomycetes. BMC Evol. Biol. (2007). doi:10.1186/1471-2148-7-174

Philippe, H., Delsuc, F., Brinkmann, H. & Lartillot, N. Phylogenomics. Ann Rev Ecol Syst 36, 541–562 (2005).

Pizarro, D. et al. Phylogenomic analysis of 2556 single-copy protein-coding genes resolves most evolutionary relationships for the major clades in the most diverse group of lichen-forming fungi. Fungal Divers. 92, (2018).

Pizarro, D. et al. Whole-Genome Sequence Data Uncover Widespread Heterothallism in the Largest Group of Lichen-Forming Fungi. Genome Biol Evol. 11, (2019).

Pringle, A., Chen, D. & Taylor, J. W. Sexual Fecundity is Correlated to Size in the Lichenized Fungus Xanthoparmelia cumberlandia. Bryol. 106, 221–225 (2003).

Printzen, C. Lichen Systematics: The Role of Morphological and Molecular Data to Reconstruct Phylogenetic Relationships. in (2010). doi:10.1007/978-3-642-02167-1_10

Proctor, R. H. et al. Birth, death and horizontal transfer of the fumonisin biosynthetic gene cluster during the evolutionary diversification of fusarium. Mol. Microbiol. (2013). doi:10.1111/mmi.12362

Pyatt, F. . Ascospore germination in Pertusaria pertusa (L.) tuck. Rev. Bryol. Lichenol. 36, 316–320 (1968).

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J. & Segata, N. Corrigendum: Shotgun metagenomics, from sampling to analysis. Nature biotechnology (2017). doi:10.1038/nbt1217-1211b

Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842 (2010).

Rambaut, A. FigTree v1. 3.1: Tree figure drawing tool. Website http//tree. bio. ed. ac. uk/software/figtree (2009).

Ranković, B. Lichen secondary metabolites: Bioactive properties and pharmaceutical potential. Lichen Secondary Metabolites: Bioactive Properties and Pharmaceutical Potential (2015). doi:10.1007/978-3-319-13374-4

Ren, R. et al. Phylogenetic resolution of deep eukaryotic and fungal relationships using highly conserved low-copy nuclear genes. Genome Biol. Evol. (2016). doi:10.1093/gbe/evw196

Roach, K. C. & Heitman, J. Unisexual reproduction reverses Muller's ratchet. Genetics 198, 1059–1069 (2014).

Roach, K. C., Feretzaki, M., Sun, S. & Heitman, J. Unisexual reproduction. Adv. Genet. 85, 255–305 (2014).

Rokas, A., Williams, B. I., King, N. & Carroll, S. B. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature (2003). doi:10.1038/nature02053

Rolshausen, G., Dal Grande, F., Sadowska-Deś, A. D., Otte, J. & Schmitt, I. Quantifying the climatic niche of symbiont partners in a lichen symbiosis indicates mutualist-mediated niche expansions. Ecography (2017). doi:10.1111/ecog.03457

Sagan, L. On the origin of mitosing cells. J. Theor. Biol. (1967).

Salas-Leiva, D. E. et al. Conserved genetic regions across angiosperms as tools to develop single-copy nuclear markers in gymnosperms: An example using cycads. Mol. Ecol. Resour. (2014). doi:10.1111/1755-0998.12228

Scherrer, S., Zippler, U. & Honegger, R. Characterisation of the mating-type locus in the genus Xanthoria (lichen-forming ascomycetes, Lecanoromycetes). Fungal Genet. Biol. 42, 976–988 (2005).

Schmitt, I. & Lumbsch, H. T. Ancient horizontal gene transfer from bacteria enhances biosynthetic capabilities of fungi. PLoS One (2009). doi:10.1371/journal.pone.0004437

Schuelke, T. A. et al. Comparative Genomics of Pathogenic and Nonpathogenic Beetle-Vectored Fungi in the Genus Geosmithia. Genome Biol. Evol. (2017). doi:10.1093/gbe/evx242

Schumacher, J. et al. A Functional Bikaverin Biosynthesis Gene Cluster in Rare Strains of Botrytis cinerea Is Positively Controlled by VELVET. PLoS One (2013). doi:10.1371/journal.pone.0053729

Seymour, F. A. et al. Breeding systems in the lichen-forming fungal genus Cladonia. Fungal Genet. Biol. 42, 554–563 (2005).

Seymour, F. A., Crittenden, P. D. & Dyer, P. S. Sex in the extremes: Lichen-forming fungi. Mycologist 19, 51–58 (2005).

Shelest, E. Transcription factors in fungi: TFome dynamics, three major families, and dual-specificity TFs. Front. Genet. (2017). doi:10.3389/fgene.2017.00053

Shen, X. X., Liang, D., Feng, Y. J., Chen, M. Y. & Zhang, P. A versatile and highly efficient toolkit including 102 nuclear markers for vertebrate phylogenomics, tested by resolving the higher level relationships of the caudata. Mol. Biol. Evol. (2013). doi:10.1093/molbev/mst122

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212 (2015).

Singh, G., Dal Grande, F., Cornejo, C., Schmitt, I. & Scheidegger, C. Genetic Basis of Self-Incompatibility in the Lichen-Forming Fungus Lobaria pulmonaria and Skewed Frequency Distribution of Mating-Type Idiomorphs: Implications for Conservation. PLoS One 7, (2012).

Singh, G., Dal Grande, F., Werth, S. & Scheidegger, C. Long-term consequences of disturbances on reproductive strategies of the rare epiphytic lichen Lobaria pulmonaria: Clonality a gift and a curse. FEMS Microbiol. Ecol. 91, (2015).

Slot, J. C. & Rokas, A. Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi. Curr. Biol. (2011). doi:10.1016/j.cub.2010.12.020

Slot, J. C. & Rokas, A. Multiple GAL pathway gene clusters evolved independently and by different mechanisms in fungi. Proc. Natl. Acad. Sci. U. S. A. (2010). doi:10.1073/pnas.0914418107

Slot, J. C. Fungal Gene Cluster Diversity and Evolution. in Advances in Genetics (2017). doi:10.1016/bs.adgen.2017.09.005

Solhaug, K. A. & Gauslaa, Y. Secondary Lichen Compounds as Protection Against Excess Solar Radiation and Herbivores. in (2011). doi:10.1007/978-3-642-22746-2_11

Spatafora, J. W. et al. The Fungal Tree of Life: from Molecular Systematics to Genome-Scale Phylogenies. Microbiol. Spectr. (2017). doi:10.1128/microbiolspec.funk-0053-2016

Spribille, T. et al. Basidiomycete yeasts in the cortex of ascomycete macrolichens. Science (80-.). 353, 488–492 (2016).

Stajich, J. E. Fungal Genomes and Insights into the Evolution of the Kingdom. Microbiol. Spectr. 5, 1–15 (2017).

Standley, K., Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780 (2013).

Stanke, M., Steinkamp, R., Waack, S. & Morgenstern, B. AUGUSTUS: A web server for gene finding in eukaryotes. Nucleic Acids Res. (2004). doi:10.1093/nar/gkh379

Stayton, C. T. The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence. Evolution (N. Y). (2015). doi:10.1111/evo.12729

Stelzer, C.-P. Does the avoidance of sexual costs increase fitness in asexual invaders? Proc. Natl. Acad. Sci. 112, 8851–8858 (2015).

Stocker-Wörgötter, E. Metabolic diversity of lichen-forming ascomycetous fungi: Culturing, polyketide and shikimate metabolite production, and PKS genes. Natural Product Reports (2008). doi:10.1039/b606983p

Sutton, B. C. Bau und Entwicklung der Flechten–Pycnidien und ihrer Conidien. By Vobis Gernot. [Bibliotheca Lich. no. 14.] Vaduz: J. Cramer. 08 1980. Pp. 141, 215 photomicrographs. Price DM 50.00. Lichenol. 13, 100 (2007).

Tatusov, R. L. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 28, 33–36 (2000).

Teixeira, M. M. et al. Exploring the genomic diversity of black yeasts and relatives (Chaetothyriales, Ascomycota). Stud. Mycol. (2017). doi:10.1016/j.simyco.2017.01.001

Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y. O. & Borodovsky, M. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res. (2008). doi:10.1101/gr.081612.108

Thell, A. & Morberg, R. Nordic Lichen Flora, Volume 4. (Museum of Evoolution, 2011).

Thell, A. et al. A review of the lichen family Parmeliaceae - history, phylogeny and current taxonomy. Nordic Journal of Botany (2012). doi:10.1111/j.1756-1051.2012.00008.x

Thorsten Lumbsch, H. & Leavitt, S. D. Goodbye morphology? A paradigm shift in the delimitation of species in lichenized fungi. Fungal Diversity (2011). doi:10.1007/s13225-011-0123-z

Traperocasas, A. & Kaiser, W. J. Development of Didymella-Rabiei, the Teleomorph of Ascochyta-Rabiei, on Chickpea Straw. Phytopathology 82, 1261–1266 (1992).

Trapero-Casas, A. & Kaiser, W. J. Differences Between Ascospores and Conidia of Didymella rabiei in Spore Germination and Infection of Chickpea. Phytopathology 97, 1600–1607 (2007).

Tripp, E. A. & Lendemer, J. C. Twenty-seven modes of reproduction in the obligate lichen symbiosis. Brittonia 70, (2018).

Tripp, E. A., Zhuang, Y. & Lendemer, J. C. A review of existing whole genome data suggests lichen mycelia may be haploid or diploid. Bryologist 120, 302–310 (2017).

Turgeon, B. G. & Yoder, O. C. Proposed nomenclature for mating type genes of filamentous ascomycetes. Fungal Genetics and Biology 31, 1–5 (2000).

Türk, A. Ö., Yilmaz, M., Kivanç, M. & Türk, H. The Antimicrobial Activity of Extracts of the Lichen Cetraria aculeata and Its Protolichesterinic Acid Constituent. Zeitschrift fur Naturforsch. - Sect. C J. Biosci. (2003).

Wang, B. et al. Presence of three mycorrhizal genes in the common ancestor of land plants suggests a key role of mycorrhizas in the colonization of land by plants. New Phytol. (2010). doi:10.1111/j.1469-8137.2009.03137.x

Wang, Y. et al. Identification of a putative polyketide synthase gene involved in usnic acid biosynthesis in the lichen Nephromopsis pallescens. PLoS One (2018). doi:10.1371/journal.pone.0199110

Wang, Y. Y. et al. Genome characteristics reveal the impact of lichenization on lichen-forming fungus Endocarpon pusillum Hedwig (Verrucariales, Ascomycota). BMC Genomics 15, (2014).

Wang, Y., Yuan, X., Chen, L., Wang, X. & Li, C. Draft Genome Sequence of the Lichen-Forming Fungus Ramalina intermedia Strain YAF0013 . Genome Announc. (2018). doi:10.1128/genomea.00478-18

Whitehouse, H. L. K. HETEROTHALLISM and SEX IN THE FUNGI. Biol. Rev. 24, 411–447 (1949).

Whittle, C. A., Nygren, K. & Johannesson, H. Consequences of reproductive mode on genome evolution in fungi. Fungal Genetics and Biology 48, 661–667 (2011).

Wiemann, P. et al. Deciphering the Cryptic Genome: Genome-wide Analyses of the Rice Pathogen Fusarium fujikuroi Reveal Complex Regulation of Secondary Metabolism and Novel Metabolites. PLoS Pathog. (2013). doi:10.1371/journal.ppat.1003475

Wilson, A. M. et al. Homothallism: an umbrella term for describing diverse sexual behaviours. IMA Fungus 6, 207–214 (2015).

Wisecaver, J. H. & Rokas, A. Fungal metabolic gene clusters-caravans traveling across genomes and environments. Frontiers in Microbiology (2015). doi:10.3389/fmicb.2015.00161

Wisecaver, J. H., Slot, J. C. & Rokas, A. The Evolution of Fungal Metabolic Pathways. PLoS Genet. (2014). doi:10.1371/journal.pgen.1004816

Yamamoto, Y. Lichen Substances and Their Biological Activities. in Recent Advances in Lichenology: Modern Methods and Approaches in Lichen Systematics and Culture Techniques, Volume 2 (ed. Upreti D., Divakar P., Shukla V., B. R.) 181–199 (Springer, 2015). doi:https://doi.org/10.1007/978-81-322-2235-4_10

Yu, J. et al. Clustered Pathway Genes in Aflatoxin Biosynthesis. Applied and Environmental Microbiology (2004). doi:10.1128/AEM.70.3.1253-1262.2004

Yu, J. H. & Leonard, T. J. Sterigmatocystin biosynthesis in Aspergillus nidulans requires a novel type I polyketide synthase. J. Bacteriol. (1995). doi:10.1128/jb.177.16.4792-4800.1995

Yun, S.-H., Berbee, M. L., Yoder, O. C. & Turgeon, B. G. Evolution of the fungal self-fertile reproductive life style from self-sterile ancestors. Evolution (N. Y). 96, 5592–5597 (1999).

Zeng, L. et al. Resolution of deep angiosperm phylogeny using conserved nuclear genes and estimates of early divergence times. Nat. Commun. (2014). doi:10.1038/ncomms5956

Zhang, F. et al. Phylogenomics from low-coverage whole-genome sequencing. Methods Ecol. Evol. (2019). doi:10.1111/2041-210X.13145

ANEXO 1

A

ANEXO 1

Comunicaciones orales y Posters en congresos

1) 8th International Association of Lichenology Symposium (Helsinki, Finland 2016)
Genome-wide search, characterization and comparison of MAT gene families in lichen- forming fungi.

2) XXI Simposio de Botánica Criptogámica (Aranjuez, España 2017)
Genome-wide search of phylogenetic markers to resolve Deep evolutionary relations in fungi

3) Lichen Genomic Workshop II (Graz, Austria 2018)
Poster: Phylogenomic analysis of the most diverse family of lichen-forming fungi.

Publicaciones científicas

1) Pizarro, D., Divakar, P.K., Grewe, F. et al. Fungal Diversity (2018) 92: 31. https://doi.org/10.1007/s13225-018-0407-7

2) Pizarro, D., Grande, F.D., Leavitt, S.D et al. Genome Biology and Evolution (2019) 11: 721–730. https://doi.org/10.1093/gbe/evz027

3) Pizarro, D., Crespo, A., Divakar, P.K. IMA Fungus (2019). In revision.

4) Pizarro, D., Divakar, P.K., Lumbsch, H.T. Genome Biology (2019). Submmited.

Estancias Breves de Investigación

1) 2016. Institute for Cell Biology and Neuroscience, Göethe University, (Francfort, Alemania). 01/04/2016 - 29/04/2016

2) 2017. Integrative Research Center, The Field Museum. (Chicago, EEUU). 28/07/2017 - 29/09/2017

Actividades formativas

1) Curso Estadística Computacional (CNB, Madrid).
2) Curso MADPHYLO, Statistical Phylogenetic (RJB, Madrid).
3) Lichen Genomic Workshop II (University of Graz, Austria).
4) Workshop "The Complexity of Lichen Symbiosis: Novel Interdisciplinary approaches from Genomic to Functional perspectives", (Universidad de Valencia, España).

Actividad Docente

180 horas de Colaboraciones docentes en el Departamento de Farmacología, Farmacognosia y Botánica durante el desarrollo de la tesis.

Seminario de "Secuenciación NGS y Genómica en Líquenes" en el Departamento de Farmacología, Farmacognosia y Botánica, Facultad de Farmacia, UCM.

Seminario de "Genética Molecular y Genómica" impartido durante dos años en la asignatura de primer año "Biología e Introducción al Laboratorio biológico".