

Genomic Insights Into the Lichen Symbiosis: *Cladonia grayi* as a Model Lichen

by

Tami Renae McDonald

Department of Biology
Duke University

Date: _____

Approved:

François Lutzoni, Supervisor

Daniele Armaleo

Fred Dietrich

Laura Rusche

Tai-Ping Sun

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy in the Department of
Biology in the Graduate School
of Duke University

2011

ABSTRACT

Genomic Insights Into the Lichen Symbiosis: *Cladonia grayi* as a Model Lichen

by

Tami Renae McDonald

Department of Biology
Duke University

Date: _____

Approved:

François Lutzoni, Supervisor

Daniele Armaleo

Fred Dietrich

Laura Rusche

Tai-Ping Sun

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Biology in the Graduate School
of Duke University

2011

Copyright by
Tami Renae McDonald
2011

Abstract

Lichens are symbioses between a fungus and a photosynthesizing partner such as a green alga or a cyanobacterium. Unlike mycorrhizal or rhizobial symbioses, the lichen symbiosis is not well understood either morphologically or molecularly. The lichen symbiosis has been somewhat neglected for several reasons. Lichens grow very slowly in nature (≤ 1 cm a year), it is difficult to grow the fungus and the alga separately and, moreover, it remains difficult to resynthesize the mature symbiosis in the laboratory. It is not yet possible to delete genes, nor has any transformation method been established to introduce genes into the genomes of either the fungus or the alga. However, the lack of genetic tools for these organisms has been partially compensated for by the sequencing of the genomes of the lichenizing fungus *Cladonia grayi* and its green algal partner *Asterochloris* sp. This work uses the model lichen system *Cladonia grayi* and the associated genomes to explore one evolutionary and one developmental question concerning the lichen symbiosis.

Chapter One uses data from the genomes to assess whether there was evidence of horizontal gene transfer between the lichen symbionts in the evolution of this very intimate association; that is, whether genes of algal origin could be found in the fungal genome or vice versa. An initial homology search of the two genomes demonstrated that the fungus had, in addition to ammonium transporter/ammonia permease genes

that were clearly fungal in origin, ammonium transporter/ammonia permease genes which appeared to be of plant origin. Using cultures of various lichenizing fungi, plant-like ammonium transporter/ammonia permease genes were identified by degenerate PCR in ten additional species of lichen in three classes of lichenizing fungi including the Lecanoromycetes, the Eurotiomycetes, and the Dothidiomycetes. Using the sequences of these transporter genes as well as data from publically available genome sequences of diverse organisms, I constructed a phylogeny of 513 ammonium transporter/ammonia permease sequences from 191 genomes representing all main lineages of life to infer the evolutionary history of this family of proteins. In this phylogeny I detected several horizontal gene transfer events, including the aforementioned one which was demonstrated to be not a transfer from plants to fungi or vice versa, but a gene gain from a group of phylogenetically unrelated hyperthermophilic chemoautolithotrophic prokaryotes during the early evolution of land plants (Embryophyta), and an independent gain of this same gene in the filamentous ascomycetes (Pezizomycotina), which was subsequently lost in most lineages but retained in even distantly related lichenized fungi. Also demonstrated was the loss of the native fungal ammonium transporter and the subsequent replacement of this gene with a bacterial ammonium transporter during the early evolution of the fungi. Several additional recent horizontal gene transfers into lineages of eukaryotes were demonstrated as well. The phylogenetic analysis suggests that what has heretofore been conceived of as a protein family with

two clades (AMT/MEP and Rh) is instead a protein family with three clades (AMT, MEP, and Rh). I show that the AMT/MEP/Rh family illustrates two contrasting modes of gene transmission: AMT family as defined here exhibits standard parent-to-offspring inheritance, whereas the MEP family as defined here is characterized by several ancient independent horizontal gene transfers (HGTs) into eukaryotes. The clades as depicted in this phylogenetic study appear to correspond to functionally different groups, with ammonium transporters and ammonia permeases forming two distinct and possibly monophyletic groups.

In Chapter Two I address a follow-up question: in key lichenizing lineages for which ammonium transporter/ammonia permease (AMTP) genes were not found in Chapter One, were the genes lost? The only definitive information which can demonstrate absence of a gene from a genome is a full genome sequence. To this end, the genomes of eight additional lichenizing fungi in the key clades including the Caliciales (*sensu* Gaya 2011), the Peltigerales, the Ostropomycetidae, the Acarosporomycetidae, the Verrucariales, the Arthoniomycetidae and the Lichinales were sequenced using the Illumina HiSeq technology and assembled with the short reads assembly software Velvet. These genomes were searched for ammonium transporter/ammonia permease sequences as well as 20 test genes to assess the completeness of each assembly. The genes recovered were included in a refined phylogenetic analysis. The hypothesis that lichens symbiotic with a nitrogen-fixing

cyanobacteria as a primary photobiont or living in high nitrogen environments lose the plant-like ammonium transporters was upheld, but did not account for additional losses of ammonium transporters/ammonia permeases in the Acarosporomyetidae and Arthoniomycetes. In addition, the four AMTP genes from *Cladonia grayi* were shown to be functional by expression of the lichen genes in a strain of *Saccharomyces cerevisiae* in which all three native ammonium transporters were deleted, and assaying for growth on limiting ammonia as a sole nitrogen source.

In Chapter Three I use genome data to address a developmental aspect of the lichen symbiosis. The finding that DNA in three genera of lichenizing fungi is methylated in symbiotic tissues and not methylated in aposymbiotic tissues or in the free-living fungus (Armaleo & Miao 1999a) suggested that epigenetic silencing may play a key role in the development of the symbiosis. Epigenetic silencing involves several steps that are conserved in many eukaryotes, including methylation of histone H3 at lysine 9 (H3K9) in nucleosomes within the silenced region, subsequent binding of heterochromatin-binding protein (HP1) over the region, and the recruitment of DNA methyltransferases to methylate the DNA, all of which causes the underlying chromatin to adopt a closed conformation, inhibiting the transcriptional machinery from binding. In this chapter I both identify the genes encoding the silencing machinery and determine the targets of the silencing machinery. I use degenerate PCR and genome sequencing to identify the genes encoding the H3K9 histone methyltransferase, the heterochromatin

binding protein, and the DNA methyltransferases. I use whole genome bisulfite sequencing of DNA from the symbiotic structures of *Cladonia grayi* including podetia, squamules and soredia as well as DNA from cultures of the free-living fungus and free-living alga to determine which regions of the genome are methylated in the symbiotic and aposymbiotic states. In particular I examine regions of the genomes which appear to be differentially methylated in the symbiotic versus the aposymbiotic state. I show that DNA methylation is uncommon in the genome of the fungus in the symbiotic and aposymbiotic states, and that the genome of the alga is methylated in the symbiotic and aposymbiotic states.

Dedication

To Erik and Vivian, the other two partners in this tripartite symbiosis.

Contents

| | |
|---|------|
| Abstract | iv |
| List of Tables | xv |
| List of Figures | xvi |
| Acknowledgements | xvii |
| 1. Multiple Horizontal Gene Transfers of Ammonium Transporters/Ammonia Permeases from Prokaryotes to Eukaryotes: Toward a New Functional and Evolutionary Classification..... | 1 |
| 1.1 Introduction..... | 1 |
| 1.2 Materials and Methods | 5 |
| 1.2.1 Cultures and DNA extraction..... | 5 |
| 1.2.2 Degenerate PCR and Sequencing..... | 6 |
| 1.2.3 Genome Sequencing..... | 9 |
| 1.2.4 Database Mining..... | 9 |
| 1.2.5 Phylogenetic Methods | 11 |
| 1.3 Results and Discussion | 12 |
| 1.3.1 Overview of Phylogenetic Tree | 12 |
| 1.3.2 One Horizontal Gene Transfer During the Early Evolution of the Fungi Creates the MEP γ clade..... | 15 |
| 1.3.3 Horizontal transfers from hyperthermophilic chemoautolithotrophic prokaryotes into the Pezizomycotina (leotiomyceta) and land plants (Embryophyta) characterize the MEP α clade..... | 16 |
| 1.3.4 Recent Transfers into Eukaryotes occur during the evolution of the MEP β gene family | 18 |

| | |
|--|----|
| 1.3.5 Multiple independent horizontal gene transfers into restricted lineages of eukaryotes in the MEP grade..... | 20 |
| 1.3.6 Fungal MEP α gene preferentially retained in lichens..... | 22 |
| 1.3.7 Dating the transfers..... | 24 |
| 1.3.8 Functional differences of AMT-Euk, MEP and Rh transporters?..... | 26 |
| 1.3.9 Nomenclatural notes..... | 27 |
| 2. An in-depth phylogenetic and physiological exploration of the MEP α clade ammonium transporters/ammonia permeases..... | 29 |
| 2.1 Introduction..... | 29 |
| 2.2 Materials and Methods..... | 31 |
| 2.2.1 Media and culture conditions..... | 31 |
| 2.2.2 Construction of plasmids..... | 32 |
| 2.2.3 Construction of yeast strains..... | 35 |
| 2.2.4 DNA amplification and Sanger Sequencing..... | 36 |
| 2.2.5 Illumina sequencing..... | 38 |
| 2.2.6 Genome assembly..... | 39 |
| 2.2.7 Phylogenetic analysis..... | 40 |
| 2.3 Results..... | 41 |
| 2.3.1 <i>Cladonia grayi</i> Mep1a and Mep1b are functional ammonium transporters..... | 41 |
| 2.3.2 Genome sequencing and assembly..... | 44 |
| 2.3.3 A panel of 20 test genes is used to assess completeness of genome assembly | 46 |
| 2.3.4 Plant-like MEP α AMTPs are missing from the genomes of lichens growing in high-nitrogen habitats..... | 47 |

| | |
|--|----|
| 2.3.5 Plant-like MEP α AMTPs are missing from the genomes of lichens that are symbiotic with nitrogen-fixing cyanobacteria..... | 51 |
| 2.3.6 Plant-like AMTPs are found in the Ostropomycetidae but not the Acarosporomycetidae | 53 |
| 2.3.7 Distribution of MEP α in the Chaetothyriomycetidae is patchy | 54 |
| 2.3.8 The MEP α gene replaces the high-affinity MEP γ gene in <i>Graphis scripta</i> | 55 |
| 2.3.9 Phylogenetic tree | 56 |
| 2.4 Discussion..... | 58 |
| 3. DNA methylation of the lichen <i>Cladonia grayi</i> in the symbiotic and free-living states . | 66 |
| 3.1 Introduction..... | 66 |
| 3.1.1 Reproducing the Armaleo and Miao Effect..... | 75 |
| 3.1.2 Finding the silencing machinery | 75 |
| 3.1.3 Finding the silencing targets..... | 76 |
| 3.2 Methods | 76 |
| 3.2.1 Cultures..... | 76 |
| 3.2.2 Genomic DNA extractions | 78 |
| 3.2.2.1 Alkali lysis method of DNA extraction | 78 |
| 3.2.2.2 CTAB method of DNA extraction | 78 |
| 3.2.2.3 DTAB/CTAB method of DNA extraction | 79 |
| 3.2.3 DNA Amplification, Cloning and Sanger Sequencing | 81 |
| 3.2.4 Southern blotting..... | 83 |
| 3.2.5 DNA immunoblotting | 84 |

| | |
|---|-----|
| 3.2.6 Methylated DNA library construction | 85 |
| 3.2.7 Methylated DNA Immunoprecipitation (MeDIP)..... | 85 |
| 3.2.8 Chromatin immunoprecipitation (ChIP) | 88 |
| 3.2.9 Bisulfite sequencing of candidate methylated regions | 89 |
| 3.2.10 Bisulfite genome sequencing | 90 |
| 3.2.11 Genome mining | 91 |
| 3.2.12 Genome sequencing..... | 91 |
| 3.3 Results | 92 |
| 3.3.1 Evidence for increased fungal DNA methylation in the symbiotic state of <i>Cladonia grayi</i> | 92 |
| 3.3.2 Identification of DNA methyltransferases from lichenizing fungi from diverse taxonomical groups..... | 98 |
| 3.3.3 Identification of targets of DNA methylation in <i>Cladonia grayi</i> | 99 |
| 3.3.4 Whole genome bisulfite sequencing..... | 101 |
| 3.4 Discussion..... | 108 |
| Appendix A: Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences generated for this study and included in the phylogenetic tree | 115 |
| Appendix B: Evolution of ammonium transporters/ammonia permeases..... | 148 |
| Appendix C. Summary of assembly results for eight lichen genomes..... | 159 |
| Appendix D. Phylogenetic placement of ammonium transporters/ammonia permeases .. from eight lichen genomes..... | 162 |
| Appendix E. List of primers used in Chapter 3..... | 169 |

| | |
|--|-----|
| Appendix F. List of all contexts in which 5-methylcytosine is found in the fungus <i>Cladonia grayi</i> | 174 |
| Appendix G. Amino acid alignment of DNA methyltransferase proteins identified in lichen genomes to <i>Neurospora crassa</i> DIM-2 DNA methyltransferase..... | 177 |
| Appendix H. DNA methyltransferase proteins identified from lichen genomes aligned to <i>Neurospora crassa</i> RID-1 DNA methyltransferase..... | 183 |
| Appendix I. Distribution of methylated sites in the fungal genome from squamules and podetia of <i>Cladonia grayi</i> | 188 |
| Appendix J. Distribution of methylated sites in the fungal genome from squamules and podetia of <i>Cladonia grayi</i> | 190 |
| References | 218 |
| Biography | 231 |

List of Tables

| | |
|--|-----|
| Table 1: Summary of characteristics uniting prokaryotic members of the MEP α clade of ammonia permeases | 18 |
| Table 2: List of primers used for construction of AMTP expression plasmids | 34 |
| Table 3: Strains of <i>Saccharomyces cerevisiae</i> carrying ATMP genes from <i>Cladonia grayi</i> | 43 |
| Table 4: Statistics on the raw data for eight lichen genomes | 45 |
| Table 5: Presence/absence in the assembled genomes of eight lichen fungi of ammonium transporter/ammonia permease genes and of 20 conserved test genes | 48 |
| Table 6: Summary of systematic placement of some lichen genera..... | 62 |
| Table 7: List of candidate methylated regions from the first methylated library | 100 |
| Table 8: Selection of candidate methylated regions from MeDIP library of squamules and podetia | 101 |
| Table 9: Accession numbers of ammonium transporter/ammonia permease genes of sequences generated for this study and included in the phylogenetic tree..... | 115 |
| Table 10: Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree | 117 |
| Table 11: Summary of assembly results for eight lichen genomes | 159 |
| Table 12: Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree | 169 |
| Table 13: Distribution of methylated sites in the fungal genome from squamules and podetia of <i>Cladonia grayi</i> | 189 |

List of Figures

| | |
|--|-----|
| Figure 1: Degenerate primers and priming sites..... | 8 |
| Figure 2: Evolution of ammonium transporters/ammonia permeases | 13 |
| Figure 3: Schematic representation of the tree of life (organismal phylogeny), for taxa of interest for this study, showing the phylogenetic placement of the proposed horizontal gene transfer events during the evolution of eukaryotes, derived from the inferred AMTP gene tree..... | 22 |
| Figure 4. Heterologous expression of <i>Cladonia grayi</i> AMTP genes in a strain of <i>Saccharomyces cerevisiae</i> lacking all native AMTPs | 42 |
| Figure 5: Simplified phylogenetic tree of the Ascomycota showing the placement of the eight lichenizing fungi used for genome sequencing | 49 |
| Figure 6: Immunoblot showing HpaII (H) and MspI (M) digest of DNA from <i>Cladonia grayi</i> podetia probed with an antibody to 5-methylcytosine | 93 |
| Figure 7: Southern blot probed with labeled total fungal DNA..... | 94 |
| Figure 8: Genomic DNA immunoblotted with antibody to methylated cytosine..... | 96 |
| Figure 9. Consensus sequence of methylation sites in the genomes of the lichen <i>Cladonia grayi</i> | 104 |
| Figure 10: Plot of bisulfite-transformed cytosine and methylcytosine on a representative algal gene..... | 106 |

Acknowledgements

I would like to thank Duke Institute for Genome Sciences and Policy (IGSP), especially Huntington Willard and Gregory Wray, for funding support for the *Cladonia grayi* and *Asterochloris* sp. genome projects and Daniele Armaleo for his contribution to the original discovery of MEP α in *Cladonia grayi*. The ammonium transporter project was also funded in part by a subcontract (112442) to Daniele Armaleo, Fred Dietrich and François Lutzoni as part of the Pacific Northwest National Laboratory (PNNL) foundational scientific focus area (FSFA) under DOE-BER's genomic sciences program in collaboration with Scott Baker and Jon Magnuson. This work would not have been possible without the Duke Shared Cluster Resource (DSCR) and outstanding services provided by John Pormann and Tom Milledge.

I am very grateful to Tai-Ping Sun and Laura Rusche for guidance, thoughtful commentary and insight. I would like to thank Daniele Armaleo for his guidance throughout, for laboratory space, assistance with sample preparation and helpful discussion. I would like to thank Fred Dietrich for writing the script for genome data quality assessment and the scripts which processes BSMAP output, for access to computational resources, and patient instruction in handling genome data. I would especially like to thank François Lutzoni for years of support and guidance.

I would like to thank Geir Hestmark for specimens of *Umbilicaria*, Suzanne Joneson for sharing the culture of *Endocarpon pallidulum*, Brendan Hodkinson for sharing

transcriptome data from *Peltigera phyllidiosa*, Bernard Goffinet for sharing genomic data from *Peltigera dolichorhiza*, Mark Chee for the pGal and p416 plasmids, and Joseph Heitman for the MLY131 strains of *Saccharomyces cerevisiae*.

I would like to thank Charles Hall and Erik McDonald for additional critical comments on the manuscript of the ammonium transporter paper, Jolanta Miadlikowska for assistance with graphics software, and the Duke University sequencing facility, particularly Fangfei Ye for sample preparation and sequencing.

I am deeply grateful for funding support including a Grant-In-Aid of Research from the Department of Biology at Duke University, a Grant-In-Aid of Research from Sigma Xi, a Graduate Research Award from the Mycological Society of America, Luttrell Mentor Travel Award and Alexopoulos Mentor Travel Award from the Mycological Society of America, a FESIN/Nordforsk travel grant, and travel grants from the Duke University Graduate School, and an Anne T. and Robert M. Bass Fellowship for Undergraduate Instruction from Duke University Graduate School, and a James B Duke Fellowship from the Duke University Graduate School.

1. Multiple Horizontal Gene Transfers of Ammonium Transporters/Ammonia Permeases from Prokaryotes to Eukaryotes: Toward a New Functional and Evolutionary Classification

1.1 Introduction

Ammonium transporters/ammonia permeases (AMTPs) are membrane-spanning proteins composed of 11 highly conserved transmembrane domains that fold into a pore through which ammonia or ammonium translocates. These proteins were described in 1994 simultaneously from plants (Ninnemann, Jauniaux, and Frommer 1994) and from fungi (Marini et al. 1994). The plant protein was named AMT, for AMmonium Transporter, and the fungal protein was named MEP, for MEthylammonium Permease. Collectively, the homologues of these proteins have been called the AMT/MEP family, although most subsequently described proteins have been called AMT while the name MEP has been adopted for only a few of the fungal AMTP homologues. Shortly afterwards, Rhesus factors (Rh) were shown to be distantly related to the AMT/MEP family (Marini et al. 1997b). The Rh proteins have 12 transmembrane domains (Gruswitz et al. 2010) and have been shown to conduct ammonia and in some cases CO₂ (Kustu and Inwood 2006; Li et al. 2007). Together, the AMT/MEP and the Rh proteins form the ammonium transporter/methylammonium permease/Rhesus factor (AMT/MEP/Rh) family.

AMT/MEP/Rh genes have a complicated evolutionary history marked by duplication and larger gene family expansions (e.g. Couturier et al. 2007), loss, and horizontal gene transfer (McDonald, Plant, and Worden 2010). For example, land plants have, in addition to the AMTPs first defined in 1994 (plant family AMT1), a second family of AMTPs that has been shown to be more closely related to bacteria than to other plant AMTPs (plant family AMT2, Sohlenkamp et al. 2002). Fungal AMTPs are also more closely related to prokaryotic AMTPs than to most other eukaryotic AMTPs (Monahan et al. 2002). Many lineages of eukaryotes including apicomplexans, microsporidia, and diplomonads such as *Giardia* have no AMTPs at all in their genomes. Reconstructing the evolutionary history of the AMT/MEP/Rh family is thus not a trivial task. A large sampling of diverse groups of organisms is essential for untangling the evolutionary intricacies of such a problem. With over 1500 prokaryotic genomes and nearly 300 eukaryotic genomes publicly available, it is possible to assemble a large dataset with the requisite phylogenetic density, that is, a sampling of all of life.

Genomically, the most extensively sampled of all groups of eukaryotes are the fungi, because they have relatively small genomes compared to most plants and animals. However, key lineages of fungi are missing in public databases, notably the lecanoromycetes, a large class of lichenizing fungi. The genomes of the model lichen *Cladonia grayi*, a symbiotic system composed of the lecanoromycete fungus *Cladonia grayi* and the green alga *Asterochloris* sp. (Trebouxiophyceae) were recently sequenced

(Armaleo, Dietrich, and Lutzoni, unpublished). Initial similarity-based searches using these fungal and algal genomes revealed sequences encoding three AMTPs from the algal partner, all of which fell into the plant AMT1 family, as expected. It also revealed sequences encoding four AMTPs from the fungal partner. Two of these fungal AMTPs showed highest similarity to other fungal AMTPs. However, the remaining two fungal AMTPs showed highest similarity to land plant (Embryophyta) AMTPs of the plant AMT2 family (absent from all green algae with sequenced genomes) and subsequently to a small assemblage of unrelated prokaryotes. No putative homologues of the plant-like AMTPs from the fungal partner were found in the genome sequence of the green algal partner, leaving open the question of how this fungus came to acquire plant-like genes.

One way in which the fungus may have acquired a plant-like gene is through horizontal gene transfer. Horizontal gene transfer (HGT), sometimes called lateral gene transfer (LGT), is the process by which DNA from a donor organism is incorporated into the genome of a recipient organism of a different species (for example, the passing of antibiotic resistance genes between unrelated bacteria). This is opposed to the transmission of genetic material from parent to offspring via reproduction, which is known as vertical gene transfer. Although horizontal gene transfer is rampant among prokaryotes (Gogarten, Doolittle, and Lawrence 2002), the extent of horizontal gene transfer in the evolutionary history of eukaryotes is unclear. Recent works taking

advantage of the diversity and sheer number of fungal genomes available have demonstrated horizontal gene transfer between plants and fungi (Richards et al. 2009) and from bacteria to fungi (Marcet-Houben and Gabaldon 2010), suggesting that the fungi have been the recipients of horizontally transferred genes. Moreover, recent works have demonstrated horizontal transfer of genes encoding nitrate or ammonia/ammonium transporters (McDonald, Plant, and Worden 2010), including horizontal transfers in which fungi have been the recipients and donors of horizontally transferred genes (Slot and Hibbett 2007).

Recent horizontal gene transfer events are relatively easy to detect. They are characterized by signatures such as the gene being present in donor lineages and in an isolated, evolutionarily distant recipient lineage or lineages; features such as insertions, deletions or intron positions that are more similar to those of distantly related organisms than to those of close relatives; or lack of introns in a eukaryotic recipient if the donor organism was prokaryotic. Ancient events are more difficult to detect because these signatures are usually lost. However, phylogenetic analyses with broad sampling can uncover ancient horizontal transfers. The signatures of these events are gene tree topologies in which well-supported monophyletic groups violate well-established common ancestry among major groups of taxa.

To determine if the plant-like AMTPs in the genome of the fungal partner of the lichen *C. grayi* were the result of a horizontal gene transfer event and to suggest a

possible donor of these genes, we performed a PCR survey for plant-like AMTPs in cultures of fungi from nearly all main lichen-forming lineages. We used the resulting 16 sequences in a phylogenetic analysis of the AMT/MEP/Rh gene family from lichenized and non-lichenized fungi (81 sequences) as well as 416 published AMTP sequences from diverse prokaryotes and eukaryotes. Here we present the results of our phylogenetic analysis of the AMT/MEP/Rh family to detect horizontal gene transfers in the evolutionary history of AMTPs, to estimate their relative time of occurrence within the tree of life, and propose a new classification for AMTPs that reflects their evolutionary history and molecular function.

1.2 Materials and Methods

1.2.1 Cultures and DNA extraction

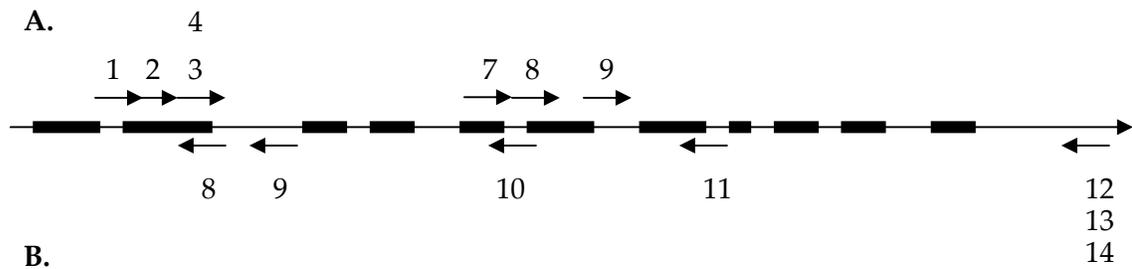
Apothecia from lichen-forming fungi were washed and affixed to the lid of an inverted petri dish, allowing the spores to be shot up onto various media. Germinated ascospores were subcultured onto solid media for long-term storage or liquid media for tissue production for DNA extraction. Media used were: Malt extract-yeast extract (Ahmadjian 1993), nutrient medium based on Bold's basal medium (Trembley et al. 2002), oatmeal (10 g homogenized oatmeal flakes), potato-carrot (as in Simmons 1992 except made as a 10x stock, and centrifuged rather than sieved, Dyer and Crittenden, personal communication); and yeast extract plus supplements (MP catalog #4101-532). A total of 20 g/l agar was added to all solid media. DNA was extracted from cultured

fungus partners of the lichens *Arthonia cf rubrocincta* (Arthoniomycetes), *Cladonia cristatella*, *Cladonia pezizaformis*, *Ramalina* sp., *Usnea strigosa* (Lecanoromycetidae, Lecanoromycetes), *Dibaeis baeomyces* (Ostropomycetidae, Lecanoromycetes), *Endocarpon cf. pallidulum*, *Pyrenula cruenta* (Chaetothyriomycetidae, Eurotiomycetes), *Laurera megasperma*, and *Trypethelium virens* (Dothideomycetes), and from washed whole dissected lichen thalli for *Lasallia papulosa*, *Parmotrema michauxiana*, *Stereocaulon tennesseense*, and *Usnea mutabilis* (Lecanoromycetes). DNA was extracted following the alkali lysis method outlined in Zolan and Pukkila (1986) as modified by Gueidan, Roux, and Lutzoni (2007). Briefly, each sample was ground to a powder, resuspended in 500 µl of a 2% SDS extraction buffer (2% SDS, 50mM Tris pH 8, 10mM EDTA, 150mM NaCl), extracted once with phenol:chloroform:IAA (25:24:1), then centrifuged at maximum speed for 5 min. The aqueous phase was removed to a new tube and extracted once with 500 µl of chloroform:IAA (24:1) and centrifuged at maximum speed for 5 min. The aqueous phase was moved to a new tube and the DNA precipitated with 300 µl (0.6 volumes) of isopropanol and centrifuged as above. The resulting pellet was washed with 70% ethanol, then allowed to air dry before resuspension in 25 µl sterile water.

1.2.2 Degenerate PCR and Sequencing

A list of the degenerate primers used for this study and a depiction of the placement of these primers relative to transmembrane domains is found in Figure 1.

PCR was performed on a MJ Research PTC200 thermocycler or an Applied Biosystems (Foster City, CA) Veriti thermocycler using a series of two touchdown PCR programs. In the first program, an initial denaturation step for 3 min at 94°C was followed by 10 cycles of 30 s at 94°C, 30 s at 60° C with a -1°C step down at each cycle, and 60 s at 72°C, followed by 25 cycles of 30 s at 94°C, 30 s at between 50°C and 47°C, and 45 s at 72°C with a final elongation step at 72°C for 7 min. If no products were formed or if faint products were seen, a second program was implemented consisting of no initial denaturation step, 24 cycles of 30 s at 94°C, 30 s at 55°C with a -0.4° C step down at each cycle, and 60 s + 2 s per cycle at 72° C , followed by 12 cycles of 30 s at 94° C, 30 s at 45° C, and 120 s + 3 s per cycle at 72°C with a final elongation step at 72°C for 10 min. PCR products were visualized on a TAE 1% agarose gel stained with SYBR Safe (Invitrogen, Carlsbad, CA). If necessary, faint products or products with multiple bands were cloned with the TOPO TA cloning kit (Invitrogen, Carlsbad, CA) following the manufacturer's instructions. For each cloning reaction, we screened at least 8 clones by colony PCR using T7 and M13R primers and a PCR program consisting of a 10 min initial denaturation step, followed by 25 cycles of 30 s at 94°C, 30 s at 52°C and 60 s at 72°C followed by a final elongation step of 7 min. PCR products were cleaned with a Montage PCR filter column (Millipore, Bilerica, MA) or with an Exo-SAP clean-up using 1 µl SAP dilution buffer, 0.5 µl Exonuclease 1, 0.5 µl Shrimp Alkaline Phosphatase, added to 10 µl PCR reaction and incubating on one of the aforementioned thermocyclers



| # | Primer name | Sequence (5' – 3') |
|----|----------------|---------------------------|
| 1 | _PLAMT_A_F1 | TGACAGCAGGTACTTTGGTCGC |
| 2 | Rama_PLANTA_F | CACAAATGGTCYATYAAAYTCCG |
| 3 | FMAFYAF | TTYATGGCITYTAYGCITT |
| 4 | FMVIFYAF | TTYATGGTITTYTAYGCITT |
| 5 | Laurera_F | AGCTTGGTGTCATTGACTATTCTGG |
| 6 | Laurera_F2 | GTTTTCTGGTCTGGATGGCATTG |
| 7 | WNGFNG | TGGAAYGGNTTYAAYGGNGG |
| 8 | Rama_PLANTA_R | TTYATGCCNTTYTAYGCNTT |
| 9 | Cgr_PLAMT_A_R1 | GGTAACATCTGAGTCCCAGAAAGCC |
| 10 | GGYVIH_R | GTGDATNACRTANCCNCC |
| 11 | FFKKPSV2 | ACIGAIGGYTTYTTRAARAA |
| 12 | AVHGEE_R | YTCYTCICCRTGIACNGC |
| 13 | AVHGEE_pen | CTCCTCACCGTGNACNGC |
| 14 | AIHGEE_R | YTCYTCICCRTGDATNGC |

Figure 1: Degenerate primers and priming sites. A. Placement of degenerate primers relative to coding regions for transmembrane domains of *Cladonia grayi mep1a* (MEP α clade, Figure. 2). Transmembrane domains shown as thickened bars. Primers shown as numbered arrows. B. Sequences of degenerate primers used to amplify “plant-like” MEP α ammonium transporters/ammonia permeases from lichenized fungi.

for 30 min at 37°C, then 15 min at 80°C (USB, Cleveland, OH). Cleaned PCR products were sequenced in 10 μ l reactions using: 1 μ l primer, 3 μ l purified PCR product, 0.5 μ l Big Dye (Big Dye Terminator Cycle sequencing kit, ABI PRISM version 3.1; PE Applied Biosystems, Foster City, CA), 1.5 μ l Big Dye buffer, and 4 μ l double-distilled water.

Automated reaction clean-up and visualization was performed at the Duke IGSP Genome Sequencing & Analysis Core Facility using Big Dye chemistry with an ABI 3730xl automated sequencer (PE Applied Biosystems, Foster City, CA). Sequencher version 4.8 (Gene Codes Corporation, Ann Arbor, MI) was used to edit sequences and assemble contigs.

1.2.3 Genome Sequencing

Cladonia grayi strain *Cgr/DA2myc/ss* was isolated from fungal ascospores (Armaleo, Sun, and Culbertson 2011). The unicellular green alga *Asterochloris* sp. *Cgr/DA1pho* was isolated from soredia (Armaleo and May 2009). A total of 5 µg of DNA was submitted to the Duke IGSP Genome Sequencing & Analysis Core Facility and sequenced with pyrosequencing (454) technology. The genome data used in this work were obtained through the *Cladonia grayi* genome project in progress at Duke University and currently housed at <http://genome.jgi-psf.org/Clagr2/Clagr2.home.html> and <http://genome.jgi-psf.org/Astpho1/Astpho1.home.html>.

1.2.4 Database Mining

All publicly available eukaryotic genomes as of 11/10/09 in Genbank (<http://www.ncbi.nlm.nih.gov/nucleotide>), the Broad Institute (<http://www.broadinstitute.org>), and the Department of Energy Joint Genome Initiative (<http://www.jgi.doe.gov>) were searched for AMTPs. Additional AMTPs from the

Cyanidioschyzon merolae genome project (<http://merolae.biol.s.u-tokyo.ac.jp>) and *Galdieria sulphuraria* genome project (<http://genomics.msu.edu/galdieria>) were also included. As AMTPs are present in multiple copies in most genomes (for example, four or five in most fungi, five or more in most plants, e.g. eight in *Clamydomonas reinhardii* and up to 15 as in *Populus trichocarpa*) it was necessary to choose one or a few genomes from each major group of interest to maximize the phylogenetic coverage of the analysis while maintaining a manageable number of genes. We used all of the AMTPs from each of the selected genomes. Whether a gene was an AMTP was determined using annotation, if available, blast similarity of greater than ~70% similarity to one of the query AMTPs: *Cladonia grayi mep1a*, *C. grayi mep2* (a high-affinity fungal AMTP, Marini et al. 1997a; Lorenz and Heitman 1998; Montanini et al. 2002; Javelle et al. 2003; Lopez-Pedrosa et al. 2006; Rutherford et al. 2008), *C. grayi mep3* (a low-affinity fungal AMTP, Javelle et al. 2006), *Arabidopsis thaliana* AMT2, or *Oryza sativa* AMT1;1), and by the presence of highly conserved AMTP motifs (FMAFYAF and variants, transmembrane domain 2; FQFAAIT and variants, transmembrane domain 3; WxWGGG and variants, transmembrane domain 4; GGYVIH or FAGGxxxH, transmembrane domain; WNGFNG and variants, transmembrane domain, and AIHGEE and variants, cytosolic tail). If structural or experimental data had been published for one or more of the transporters, we chose that genome over other genomes in the major group of interest. Some groups of interest, like the conifers, do not have completed genome projects. To represent this group and

others without completed genome projects, we chose EST data and therefore do not have a full representation of the AMTPs in the genome. In addition, we used all sequences published in Huang and Peng (2005). All publicly available prokaryotic genomes in Genbank were also searched and selected as above, endeavoring to include at least one representative of each major phylum. Genomes in the analysis are listed in Appendix A.

To determine the phylogenetic extent of the “plant-like” AMTPs within the Fungi, we searched all fungal genomes available (86 as of 11/10/09) through NCBI (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?organism=fungi) as well as the whole genome shotgun reads (40 additional fungi) using tblastx and blastp. We searched the fungal genomes at the Saccharomyces Genome Database site (SGD) (<http://www.yeastgenome.org/cgi-bin/blast-fungal.pl>, 2 additional fungi) and JGI (<http://genome.jgi-psf.org>, 23 additional fungi), and the Broad institute (http://www.broadinstitute.org/annotation/genome/multicellularity_project/MultiHome.html, 3 additional fungi) using blastx and tblastx. All AMTPs from fungal genomes containing plant-like AMTPs (4 genomes) were included in the analysis.

1.2.5 Phylogenetic Methods

Manual alignments were performed using MacClade 4.08 (Maddison and Maddison 2005). Ambiguously aligned regions and introns were delimited manually and excluded from phylogenetic analyses. Alignments are available through TreeBASE

(<http://purl.org/phylo/treebase/phyloids/study/TB2:S11394>). Models of molecular evolution were selected using the Akaike Information Criterion (AIC) implemented in jModeltest (Posada 2008; Guindon and Gascuel 2003) or MrModeltest 2.3 (Nylander 2004). Phylogenetic relationships and confidence values were inferred using a maximum likelihood approach at the nucleotide level. Maximum likelihood analysis at the nucleotide level used GTR GAMMAI (with a gamma parameter and a proportion of invariable sites, = GTR+ Γ +I). The program RAxML-VI-HPC (was used for the maximum likelihood search for the most likely tree. The same program using the same settings was used for the bootstrap (BS) analysis with 1000 BS replicates. The consensus tree was calculated and visualized using the majority rule consensus tree command in PAUP 4.0d701 (Swofford 2002).

1.3 Results and Discussion

1.3.1 Overview of Phylogenetic Tree

A phylogenetic tree of the AMT/MEP/Rh family resulting from a maximum likelihood analysis at the nucleotide level of 513 sequences is presented in Figure 2 and Appendix B. In this tree, two main monophyletic groups (AMT-Euk and MEP clades) and one basal grade (MEP grade) are evident. The Rh group is very distinct from the MEP and AMT family and has been defined previously (Huang and Peng 2005). We use the Rh family sequences to root our AMTP phylogeny. Of interest is the placement of some of the eukaryotic sequences within the tree.

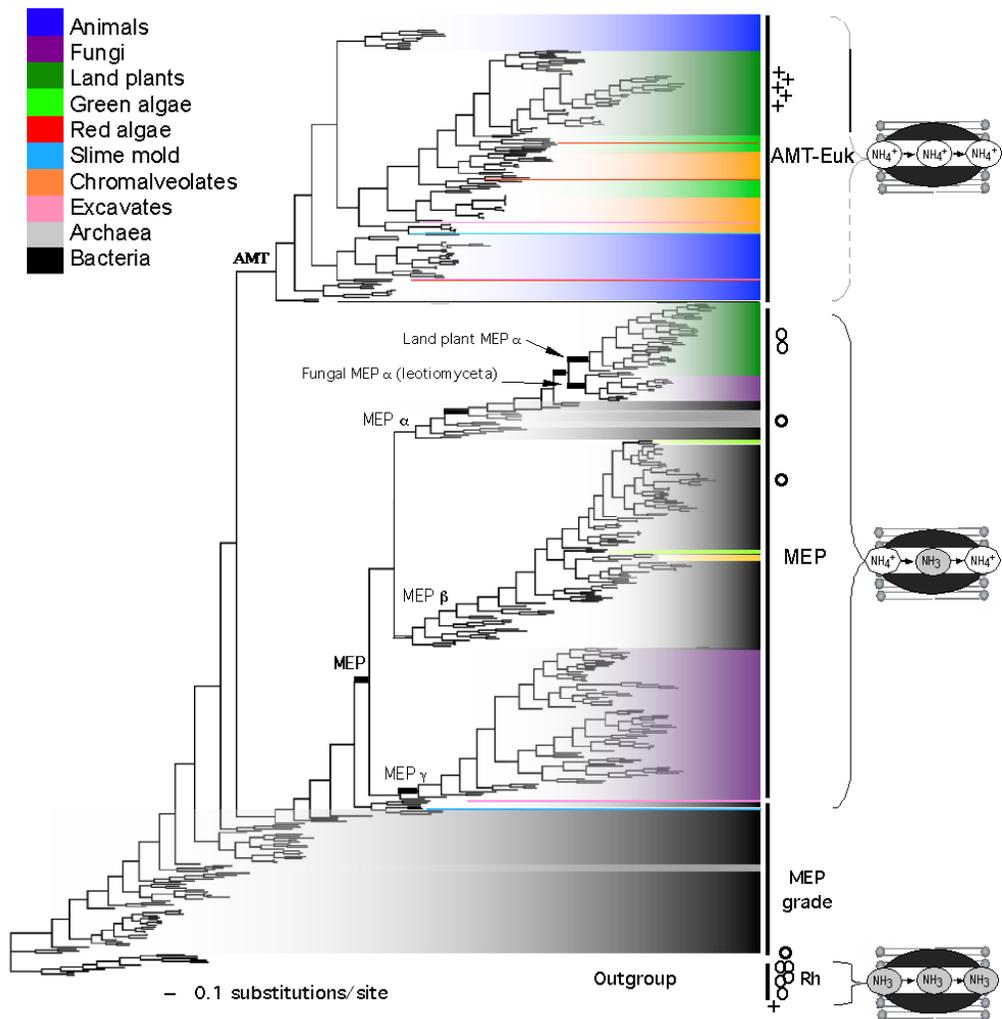


Figure 2: Evolution of ammonium transporters/ammonia permeases.

Maximum likelihood analysis of 513 ammonium transporter/ammonia permease genes shows the presence of a eukaryotic clade (AMT-Euk) representing vertical gene transfer of AMTPs into eukaryotes and a predominantly prokaryotic clade (MEP) in which eukaryotic lineages demonstrate horizontal gene transfer. AMT-Euk=eukaryotic ammonium transporters clade; MEP=methylammonium permease clade; MEP grade=prokaryotic ammonium transporter/ammonia permease grade; Rh=Rhesus family (outgroup). Thickened lines show $\geq 70\%$ bootstrap support for nodes discussed in the text. More detailed information about this phylogenetic tree, including all supported nodes, is shown in Appendix B. Proposed transport mechanisms are schematically represented to the right of each clade. + = electropositive transport demonstrated by

electrophysiology; O = electroneutral transport demonstrated by electrophysiology; O= electroneutral transport demonstrated by crystal structure.

Although the name “AMT” (AMmonium Transporter) is usually applied to all AMTP sequences not in the Rh family, here we restrict the term AMT to the sequences arising from node AMT (Figure 2). The AMT genes found in eukaryotes form a monophyletic group, referred to here as the AMT-Euk clade (corresponding to the AMT1 clade of McDonald, Plant, and Worden 2010). The AMT-Euk clade represents regular vertical transfer of AMTPs into eukaryotes. Notably, while AMTPs from animals, land plants, green algae, red algae, slime molds, chromalveolates and excavates are found in the AMT-Euk clade, fungal AMTPs are entirely missing from the AMT-Euk clade. Instead, they are found exclusively in the MEP clade (Figure 2 and Appendix B).

While the name “MEP” (MEthylammonium Permease) is currently applied to AMTPs from only some fungi, here we expand that usage to include all sequences arising from the well-supported MEP node supporting the MEP clade (corresponding to the AMT2 clade of McDonald, Plant, and Worden 2010; Figure 2). The MEP clade may be further subdivided into the plant-like MEP clade (MEP α) of putative archaeal origin, the prokaryotic MEP clade (MEP β) and the fungal MEP clade (MEP γ) of prokaryotic origin, discussed separately below. The MEP clade is a largely prokaryotic clade with unrelated groups of eukaryotes interspersed within it. Each eukaryotic lineage within the MEP clade therefore represents one or more horizontal gene transfer events, which

will be discussed in greater detail below, in chronological order from the earliest event to the latest event.

Broadly speaking, this gene tree follows the expectations for vertical gene transfer; a eukaryotic clade (AMT-Euk) arising from a grade of prokaryotic sequences (MEP grade) at the base of the AMTP tree. However, punctuating this vertical evolution are several recent and at least 3 ancient transfers from prokaryotes into diverse eukaryotic lineages, all found in the MEP clade. Thus, while all the eukaryotic sequences in the AMT-Euk clade represent vertical evolution, all the eukaryotic monophyletic groups in the MEP clade result from horizontal gene transfers.

1.3.2 One Horizontal Gene Transfer During the Early Evolution of the Fungi Creates the MEP γ clade

Fungi are the only major group of eukaryotes entirely missing AMTPs from the AMT-Euk clade. Fungal sequences are instead found in the MEP clade. Every fungus sampled so far has multiple AMTPs in the fungal MEP γ clade (Appendix A, B). Most have three to four copies, indicating that the MEP γ gene family has expanded through gene duplication. Because all fungi lack an AMT-Euk gene resulting from vertical inheritance and all fungi have at least one copy of the MEP γ gene, we interpret this phenomenon as a gene replacement event. Fungal ancestors, like other eukaryotes, must have originally had the AMT-Euk gene in their genomes. They then acquired the MEP γ gene near the time of the divergence of the fungi and lost the native AMT-Euk

gene. Therefore, the replacement of an AMT-Euk gene by a MEP γ gene could be a defining trait of the kingdom Fungi.

1.3.3 Horizontal transfers from hyperthermophilic chemoautolithotrophic prokaryotes into the Pezizomycotina (leotiomyceta) and land plants (Embryophyta) characterize the MEP α clade

Fungal AMTP genes are also found in the MEP α clade. This clade comprises a well-supported group of MEP genes from land plants (excluding green algae) sister to a well-supported group of genes from a subgroup of filamentous ascomycete fungi within the Pezizomycotina (referred to as the leotiomyceta; Schoch et al. 2009), where most of the Ascomycota species richness as well as all known lichen-forming ascomycetes are concentrated (Figure 2; Appendix B). This sister relationship between the MEP genes found in the leotiomyceta and those found in land plants explains the initial similarity search results in which these fungal MEP genes appeared to be most similar to AMTPs from land plants. These leotiomyceta-specific MEP α sequences are very distinct from the sequences in the fungal MEP γ clade and have a separate evolutionary origin. The same is true for the land plant MEP α sequences compared to their AMT-Euk plant counterparts. At the base of the MEP α clade is a grade of phylogenetically unrelated prokaryotes that are united by ecology: they are all hyperacidophilic, thermophilic, chemolithotrophic and chemoautolithotrophic Bacteria and Archaea (table 1) isolated from volcanic hot springs and similar extreme environments (Clark and Norris 1996; Schleper et al. 1996; Itoh et al. 1999; Suzuki et al. 2002; Goltsman et al. 2009). Most

closely related to the land plant and fungal MEP α are AMTPs from a small lineage of gamma proteobacteria (*Acidithiobacillus*) known from acid mine drainages (Kelly and Wood 2000; Valdes et al. 2009). The rest of the proteobacterial sequences, including the bulk of the gamma proteobacterial sequences, fall elsewhere in the tree (Appendix A, B). Likewise, *Leptospirillum rubrum* (Nitrospirae) and *Acidimicrobium ferrooxidans* (Actinobacteria) are unusual iron-oxidizing representatives of their phyla with AMTPs in this clade. More distantly related is a paraphyletic group of hyperacidophilic thermotolerant Archaea including *Caldivirga*, *Ferroplasma*, *Picrophilus*, and *Sulfolobus*. Except for the fungi, which were not extensively sampled in their work, this clade was also recovered by McDonald, Plant, and Worden (2010).

Because the early diverging diversity in this well-supported clade is archaeal, we suggest that land plants and filamentous ascomycetes in the leotiomyceta each inherited, independently, an archaeal-type AMTP. Although the original source of the gene is likely archaeal, the actual donor organism may have been a eubacterium. We suggest this because the gamma-proteobacterial sequences share a most recent common ancestor with the eukaryotic sequences, suggesting that they are more closely related to the sequences of the donor than are the archaeal sequences.

Table 1: Summary of characteristics uniting prokaryotic members of the MEP α clade of ammonia permeases. All are acidothermophilic chemolithotrophs or chemoautolithotrophs isolated from volcanic hot springs, acid mine seeps, and similar extreme environments. Carbon source: A=autotroph; H=heterotroph.

| Organism | Electron donor | pH optimum | Temperature optimum (°C) | Carbon source | GC content (%) |
|--|----------------|------------|--------------------------|---------------|---------------------|
| Bacteria | | | | | |
| <i>Acidimicrobium ferrooxidans</i> ¹ | Iron | 2 | 45-50 | A/H | 68.3 |
| <i>Acidithiobacillus caldus</i> ² | Sulfur | 2.0 – 2.5 | 45-50 | A | 61.4 ³ |
| <i>Acidithiobacillus ferrooxidans</i> ² | Iron, sulfur | 1.5 – 2 | 30-35 | A | 58-59 |
| <i>Leptospirillum rubrum</i> ⁴ | Iron | 1.1 | 41 | A | 55* |
| Archaea | | | | | |
| <i>Caldivirga maquilingensis</i> ⁵ | Sulfur | 3.7 – 4.2 | 85 | H | 43.1 |
| <i>Ferroplasma acidarmanus</i> ⁶ | Iron | 1.2 | 42 | H | 36.8 |
| <i>Picrophilus torridus</i> ^{7,8} | Iron | 0.7 | 60 | H | 36 |
| <i>Sulfolobus tokodaii</i> ⁹ | Sulfur | 2.5-3 | 75-80 | H/A | 32.8 |
| <i>Sulfolobus solfataricus</i> ^{9,10} | Sulfur | 2-4 | 80 | H/A | 35.79 ¹¹ |

* calculated from data available on 10/14/10 at

http://www.ebi.ac.uk/ena/data/view/Taxon:419542&portal=con_release&page=1

¹(Clark and Norris 1996)

² (Kelly and Wood 2000)

³(Valdes et al. 2009)

⁴(Goltsman et al. 2009)

⁵(Itoh et al. 1999)

⁶(Dopson et al. 2004)

⁷(Schleper et al. 1995)

⁸(Futterer et al. 2004)

⁹(Brock et al. 1972)

¹⁰(She et al. 2001)

¹¹ <http://microbes.ucsc.edu/cgi-bin/hgGateway?db=sulSol1> accessed on 10/14/10

1.3.4 Recent Transfers into Eukaryotes occur during the evolution of the MEP β gene family

Other eukaryotic sequences are found in the bacterial MEP β clade. MEP

sequences from the chromalveolate *Phytophthora infestans*, an oomycete which is the

causative agent of potato late blight, cluster together without support with one cyanobacterium (*Gloeobacter violaceus*) and *Deinococcus radiodurans* as a sister clade to the actinobacteria (Appendix B). All other chromalveolates, including diatoms and other stramenopiles, which are close relatives of oomycetes, are absent from the MEP β clade, indicating a relatively recent horizontal transfer of the MEP β gene from bacteria into the oomycetes. A horizontal gene transfer of AMTPs between fungi and oomycetes has been suggested elsewhere (McDonald, Plant, and Worden 2010). However, the more extensive prokaryotic sampling in this analysis allows us to exclude that possibility as the oomycetes clearly fall inside the bacterial MEP β clade and far from the well-supported MEP fungal clades within the MEP α and γ clades. Therefore, our results do not support a horizontal transfer scenario between fungi and oomycetes. Rather, the relationship previously detected likely reflects the separate bacterial origins of fungal and oomycete MEPs.

Shown here as a sister group without support to the aforementioned clade of oomycetes, actinobacteria, and intervening sequences, are the Mamielallean green algae *Ostreococcus* and *Micromonas*, suggesting a separate horizontal transfer of a MEP β AMTP from a bacterial donor to these green algae. One additional horizontal transfer event is indicated by the placement of another *Ostreococcus* sequence, which arises from a group of proteobacteria. These horizontal transfers of AMTPs into the Mamielalles

were also detected by McDonald, Plant and Worden (2010), who had the complete complement of AMTPs from Mamielalleean genomes in their analyses.

Because there is no support at the deeper nodes of the bacterial clade (MEP β), we cannot detect horizontal gene transfer events between groups of prokaryotes. We recover proteobacteria, actinobacteria, and firmicutes clades, but have no support for the relationships between these groups. Likewise, we recover a basal grade containing more proteobacterial and cyanobacterial sequences, as well as Bacteriodes, Themotoga, Green sulfur bacteria, and Purple bacteria, with no support for the relationships between these groups. Although eubacterial phyla are recovered, the relationships between the phyla cannot be resolved.

1.3.5 Multiple independent horizontal gene transfers into restricted lineages of eukaryotes in the MEP grade

Most unusual is a small but well-supported clade containing AMTP sequences from *Trypanosoma cruzi* (excavate), *Naegleria gruberi* (excavate), *Dictyostelium discoideum* (slime mold), and the phylogenetically unrelated prokaryotes *Ureaplasma urealyticum* (Tenericutes, Mycoplasmatales), *Enterococcus faecium* (Firmicutes), and *Methanosarcina* spp. (Archaea, Euryarchaeota). The lack of a copy of this AMTP in related lineages (including *Trypanosoma brucei*), suggests that each of the eukaryotic lineages gained this AMTP recently and independently. Among the prokaryotic members of this clade, *Methanosarcina* in particular has experienced horizontal transfer of AMTP genes from multiple sources as shown in this study. However, it is difficult to find a characteristic

that unites all the organisms in this group. Many are pathogens or commensals of animals. The association of eukaryotic and prokaryotic symbionts together in the GI tract, urinary tract, or blood stream of an animal host may have facilitated horizontal transfer. By extending the MEP clade to one node deeper into the AMTP phylogeny (Appendix B) the MEP clade would remain monophyletic and encompass four MEP subclades instead of three. However, because of our low sampling for this fourth MEP subclade and its uncertain phylogenetic placement within the MEP grade, we decided to keep it as part of the MEP grade.

In summation, the eukaryotic MEP genes are not most closely related to each other. Eukaryotic MEP clades have arisen through at least nine putative separate horizontal gene transfer events (Figure 3) from at least three lineages of Bacteria and from hyperthermoacidophilic Archaea (Appendix B). Our data show no evidence of horizontal gene transfer from eukaryotes to prokaryotes or from eukaryotes to eukaryotes.

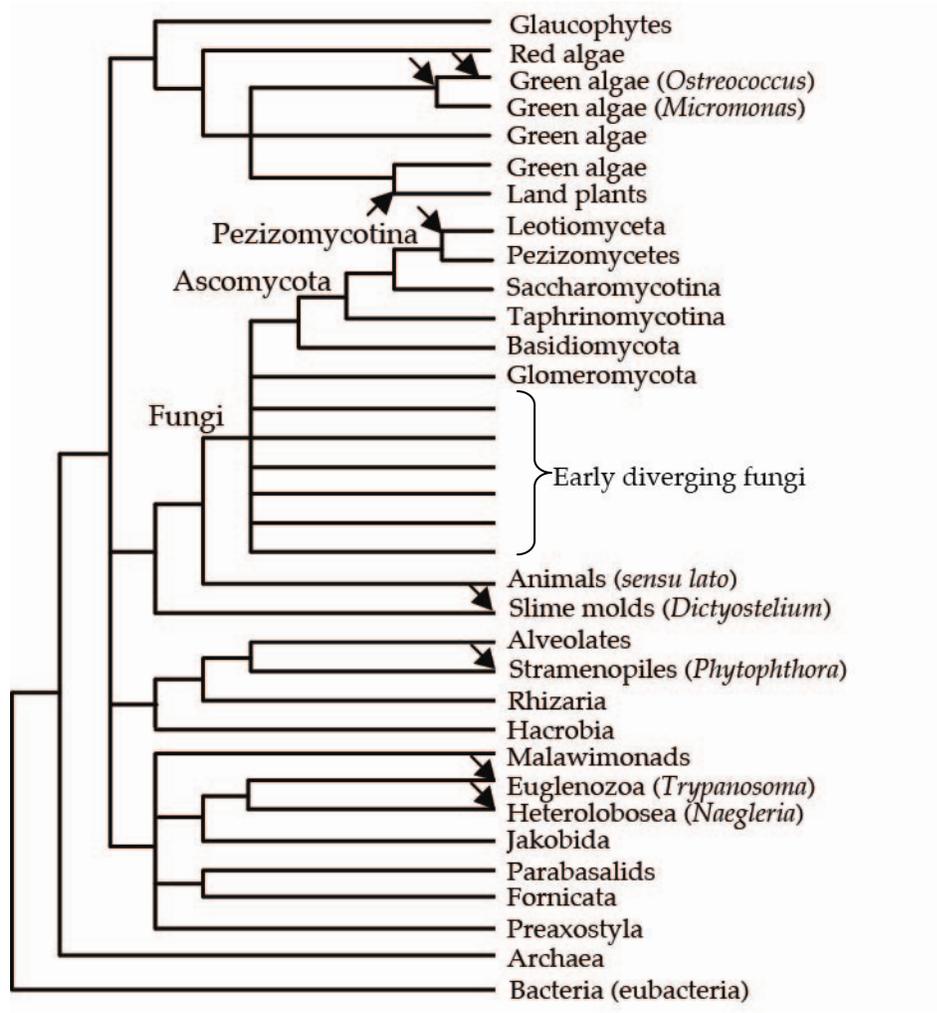


Figure 3: Schematic representation of the tree of life (organismal phylogeny), for taxa of interest for this study, showing the phylogenetic placement (arrows) of the proposed horizontal gene transfer events during the evolution of eukaryotes, derived from the inferred AMTP gene tree (Figure 2; Appendix B). **Classification and tree adapted from Hibbett et al. 2007; Schoch et al. 2009; Keeling, Leander, and Simpson 2009.)**

1.3.6 Fungal MEP α gene preferentially retained in lichens

This project was initiated by the discovery of two “plant-like” AMTP genes (MEP α) in the newly sequenced genome of the lichen-forming ascomycete *Cladonia grayi*. To determine the distribution of these “plant-like” MEP α genes among fungi, we queried all publicly available fungal genomes. In addition, we attempted to amplify by degenerate PCR the fungal MEP α gene from 15 cultures of lichenizing fungi from three classes of Pezizomycotina (Appendix A). Although most lichens are found in one large class of Ascomycota (Lecanoromycetes), lichens are also found in four other fungal classes that include non-lichenized fungi (Lutzoni, Pagel, and Reeb, 2001; Lutzoni et al. 2004; James et al. 2006; Schoch et al. 2009). We found the MEP α gene in nine out of 15 lichens sampled from Lecanoromycetes, Eurotiomycetes, and Dothidiomycetes (AppendixA). These three classes of Ascomycota belong to the leotiomyceta (Figure 3). As more lichen genomes accumulate, we should be able to refine our primers and identify this gene in additional lichenized fungal genomes.

By contrast, only four of the 151 non-lichen fungal genomes searched had the MEP α gene, two in a small subgroup of *Penicillium* (Eurotiomycetidae; only those with *Talaromyces* telomorphs) (also detected by McDonald, Plant and Worden 2010) and two in a subgroup of *Fusarium*. The *Penicillium* members are in the Eurotiomycetes, a class with lichenized members that retain the gene. The *Fusarium* members are in the class Sordariomycetes, which is entirely non-lichenized but share a common ancestor with

three lichen-rich groups (Lecanoromycetes, Lichinomycetes and Eurotiomycetes) as part of the leotiomyceta (Figure3).

We interpret this distribution as a horizontal gene transfer into filamentous ascomycetes (Pezizomycotina), after the divergence of the Pezizomycetes, but before the radiation of the leotiomyceta (James et al. 2006; Schoch et al. 2009). This was followed by a subsequent loss of the gene in almost all non-lichenized lineages, while the gene was retained and in fact duplicated in lichenized lineages of Lecanoromycetes, Eurotiomycetes and Dothidiomycetes (Appendix A and Appendix B).

Because only land plants and a relatively late-evolving group of fungi have the MEP α gene, we rule out the possibility of one transfer to the base of eukaryotes and subsequent losses by intervening eukaryotic clades, which would require at least 14 separate events. Instead we interpret this distribution as the result of two separate horizontal transfers, one to the ancestor of the land plants, and one to an ancestral lineage of the leotiomyceta, by the same or similar prokaryotic donors. This scenario explains the sister relationship between MEP α genes from land plants (Embryophyta) and a group of filamentous ascomycetes (leotiomyceta). Within leotiomyceta, the gene was then lost in most lineages, but retained by many lichenized lineages, no matter how distantly related.

1.3.7 Dating the transfers

The MEP α gene appears to have entered the fungal lineage after the divergence of the Saccharomycotina from the Pezizomycotina (≈ 773 MYA) but before the radiation of the leotiomyceta (≈ 673 MYA, Blair 2009), suggesting a window of about 100 MY for this horizontal transfer event into the Pezizomycotina. The MEP α gene appears to have entered the plant lineage after the divergence of the land plants from the various chlorophyte algae ≈ 936 MYA (Bhattacharya et al. 2009), and before the diversification of extant land plants, ≈ 440 -550 MYA.

Why this particular MEP gene should be retained in land plants and in lichenizing Ascomycota is unclear. But the fact that both groups rely on photosynthesis for carbon, colonized land at about the same time or in close succession when sources of nitrogen might have been limiting, and gained the ammonium transporter more or less contemporaneously suggests that this gene may be advantageous under these conditions. For example, it may allow a balanced efficiency in nitrogen transport between algal and fungal partners forming lichen symbioses and between land plants and their fungal partners forming mycorrhizal symbioses. It could be argued that if one of the partners was more efficient in transporting nitrogen, it might have prevented the establishment and maintenance of these symbioses and the subsequent diversification of the lichen-forming Ascomycota and land plants. Therefore, this gene might have been one in a suite of pre-adaptations to lichenization of ascomycetes, or a copy of the MEP

transporter might have eased the transition to land for fungi and plants. Functional studies are required to explore this and other hypotheses.

1.3.8 Functional differences of AMT-Euk, MEP and Rh transporters?

Although it was originally thought that all proteins in the AMT/MEP/Rh family were ion channels transporting ammonium, the publication of the crystal structure from *Escherichia coli* (Khademi et al. 2004; Zheng et al. 2004) and the archaean *Archaeoglobus fulgidus* (Andrade et al. 2005) challenged this view. These crystal structures suggested that at least some of the proteins are modified gas channels conducting ammonia (NH₃). Although there is some debate (Lamoureux, Klein, and Berneche 2007; Lamoureux et al. 2010), the current model for the mechanism of action of these modified gas channels is that the ammonium proton is stripped at the entrance to the pore, whereupon ammonia moves through the channel and is reprotonated at the cytoplasmic side (Winkler 2006; Li et al. 2006; Yang et al. 2007). These results have been extrapolated to the whole AMT/MEP/Rh family (Winkler 2006). However, because both of these proteins fall into the MEP family (the *A. fulgidus* protein belongs to the MEP α clade, and the *E. coli* protein belongs to the MEP β clade), it is unlikely that these structures represent the whole AMT/MEP/Rh family. In fact, the crystal structure of the *Nitrosomonas europaea* Rh protein (Li et al. 2007) suggests that Rh proteins are gas channels for ammonia (NH₃) and/or CO₂; no deprotonation and reprotonation are involved (Figure 2). Although no crystal structure is available for any protein in the AMT-Euk clade, some proteins in this

clade, notably from plants, have been identified as bona fide ammonium transporters (Figure 2) using electrophysiology experiments (Ludewig, von Wiren, and Frommer 2002; Mayer, Dynowski, and Ludewig 2006). Work on plant family AMT1 genes, which fall into the AMT-Euk clade, demonstrates that transport in this group is electropositive, meaning either that ammonium is transported or that ammonia and a proton are co-transported (Mayer and Ludewig 2006).

We raise the possibility that each of these evolutionarily ancient families of AMTPs functions in a unique manner: MEP proteins strip the proton and transport ammonia, Rh proteins transport ammonia and/or CO₂ without de- and re-protonation, and AMT-Euk proteins transport ammonium or co-transport ammonia and a proton (Figure 2). Another interpretation could be that, after land plants acquired a MEP AMTP, the native AMT-Euk copy was freed to neofunctionalize into a true ammonium transporter, such that ammonium transporters are restricted to a subset of AMT-Euk lineages related to plants. Electrophysiological experiments on members of the AMT-Euk clade that are distantly related to plants are needed to explore the mode of action of these proteins (Figure 2).

1.3.9 Nomenclatural notes

From the moment of publication in 1994, both “MEP” (Marini et al. 1994) and “AMT” (Ninnemann, Jauniaux, and Frommer 1994) have been used to name functionally characterized AMTPs. Because the name AMT was applied to a sequence

that falls into the eukaryotic clade (AMT-Euk, Figure 2), and because MEP was applied to a sequence falling into the other clade (MEP, Figure 2), we have named our clades for these “founding” sequences. This clade nomenclature emphasizes not only the phylogenetic distinctness of gene families, but also the functional differences found so far between characterized members of each clade: bona fide ammonium transporters showing electropositive transport all fall in the AMT clade, and ammonia permeases showing electroneutral transport fall into the MEP clade. Already in the molecular literature the designation AMT/MEP/Rh is in wide use. This designation corresponds well to the functional classes already recognized and also proposed here.

2. An in-depth phylogenetic and physiological exploration of the MEP α clade ammonium transporters/ammonia permeases

2.1 Introduction

Ammonium transporters/ammonia permeases (AMTPs) are highly conserved proteins found in most organisms, including prokaryotes and eukaryotes. These proteins are composed of 11 transmembrane helices that fold into a pore through which ammonia or ammonium moves. In their native conformation they trimerize, forming a tripartite pore. While some AMTPs have been shown to transport ammonium (NH_4^+), notably those proteins in the AMT2 family of land plants (Sohlenkamp *et al.* 2000; Simon-Rosin *et al.* 2003), most AMTPs have been shown to transport ammonia (NH_3).

All fungi have AMTPs. Fungal AMTPs are in a phylogenetic clade by themselves, separated from all other eukaryotic ammonium transporters. Within this fungal ammonium transporter clade are both low-affinity and high-affinity ammonium transporters (McDonald *et al.*, 2011). High-affinity ammonium transporters are found in all fungal genomes and have been characterized in several fungi (e.g. Marini *et al.* 1997, Rutherford *et al.* 2008). By contrast, low-affinity ammonium transporters are found only in Ascomycetes and Basidiomycetes (Dikarya), not in chytrids, zygomycetes, or glomeralean fungi, all of which are earlier diverging lineages within the Fungi. Low-affinity ammonium transporters are also absent from all other eukaryotes and prokaryotes. Together, high-affinity and low-affinity fungal ammonium transporters

form a well-supported subclade of the MEP family (MEP γ , McDonald et al, 2011).

Recently, a new clade of fungal ammonium transporters was described (McDonald et al, 2011). This new clade of fungal ammonium transporters is distinct from the fungal high- and low-affinity ammonium transporters of the MEP γ clade, and in fact is most closely related to ammonium transporters from land plants in the AMT2 family and to transporters from mostly hyperacidophilic chemoautolithotrophic prokaryotes inhabiting deep sea thermal vents, volcanic hot springs, acid mine drainages, and similar extreme environments (MEP α clade). Interestingly, only a subset of filamentous ascomycetes, most of which are symbiotic with green algae in lichen symbioses, have representatives of this new clade of ammonium transporters in their genomes. In fact, lichenizing fungi in three different taxonomic classes of fungi have actually duplicated this ammonium transporter. By contrast, only four non-lichenizing fungi in two genera (*Penicillium* with *Talaromyces* teleomorphs, *Fusarium* with *Gibberella* teleomorphs) out of more than 200 publicly available sequenced fungal genomes have representatives of this new clade of ammonium transporters in their genomes, and this transporter is not duplicated. This result suggests that lichenized fungi preferentially retained the MEP α gene after the initial horizontal gene transfer event during the early evolution of the leotiomyceta, while non-lichenized fungi have lost this gene.

Interestingly, the MEP α gene was not found in all lichens surveyed. In particular, the MEP α gene was never recovered from the two orders of lichens most

closely related to the order in which the original discovery was made. In one of these two orders, the Peltigerales, the lichens are symbiotic with nitrogen-fixing cyanobacteria. In the other order, the Teloschistales, many lichens inhabit high-nitrogen niches like bird perching sites. The availability of nitrogen sources from the environment or from a symbiont, coupled with the failure to identify the AMTP gene by PCR suggests that lichens in these two orders may no longer need the MEP α ammonium transporter and may have shed it from their genomes.

Here, I further characterize this new clade of fungal ammonium transporters. I correlate presence of the ammonium transporters of the new clade with nitrogen lifestyle by surveying lichen fungi closely related to the main lineages previously found but tolerating high nitrogen habitats, and some with nitrogen-fixing cyanobacteria rather than green algae as the primary symbionts. I also characterize the function of the ammonium transporters from one lichen, *Cladonia grayi*, by assaying for growth on ammonium as a sole nitrogen source. I present a phylogeny of fungal ammonium transporters to contextualize this clade.

2.2 Materials and Methods

2.2.1 Media and culture conditions

For yeast strains, standard yeast media were used including: Synthetic Limiting Ammonium and Dextrose (SLAD) composed of 0.17% yeast nitrogen base (without

amino acids and ammonium sulfate), 2% glucose and 2% agar (Gimeno *et al.* 1992; Lorenz & Heitman 1998) with modifications including 2 mM, 5 mM, 7 mM, 10 mM and 20 mM ammonium sulfate; and Synthetic Complete medium (SC) lacking uracil, which contains 6.7 g/l yeast nitrogen base without amino acids, supplemented with all amino acids except uracil, 2% glucose and 2% agar. Yeast strains were grown at 30°C or at room temperature. *Cladonia grayi* was maintained at room temperature in liquid shaking cultures of MEYE medium consisting of 20 g/l malt extract and 2 g/l yeast extract (Ahmadjian 1966).

2.2.2 Construction of plasmids

RNA was extracted from *Cladonia grayi* cultures growing in liquid medium. Tissue from liquid cultures was harvested, rinsed in distilled water, lyophilized, then ground to a fine powder under liquid nitrogen in a pre-chilled mortar and pestle. The ground tissue was resuspended in TRIzol Reagent (Invitrogen, Carlsbad, CA) and RNA extraction proceeded according to the manufacturer's instructions.

The first strand of the cDNA was generated using the following reverse transcription reaction mix: 1.0 µl 10 x PCR buffer, 2.0 µl 25mM MgCl₂ (both supplied with AmpliTaq DNA Polymerase, Applied Biosystems), 2.0 µl dNTPs (10 mM each), 0.5 µl RNase inhibitor (20 µg/µl, Applied Biosystems) 0.5 µl MuLV reverse transcriptase (Applied Biosystems), 1.75 µl of distilled water, and 1.25 µl reverse primer. The

thermocycling conditions for this first strand synthesis reaction were: 42°C for 35 minutes, then 99°C for 5 minutes. Primers are listed in Table 2.

PCR to regenerate the second strand of the cDNA was performed using the entire volume of first strand reaction plus 31.25 µl distilled water, 4 µl 20xPCR buffer, 3 µl 25 mM MgCl₂, 1.25 µl of a forward primer, 0.25 µl of AmpliTaq polymerase (Applied Biosystems) and 0.25 µl of an antibody to Taq polymerase (Clontech). The thermocycling program consisted of 24 cycles of 94°C for 30 seconds, 55°C for 30 seconds with a 0.4 degree decrease in temperature for each cycle, and 72° for 1 minute, followed by 12 cycles of 94°C for 30 seconds, 45°C for 30 seconds, and 72°C for 2 minutes with a 3 second increase per cycle, followed by a final extension at 72°C for 10 minutes.

cDNA of each AMTP from *Cladonia grayi* was cloned into the TOPO TA vector as described above. Plasmid DNA was restricted to release the full-length AMTP cDNA which was then subcloned into the pRS306-GAL1-TADH vector (unpublished, a gift from Mark Chee) carrying a uracil selectable marker, an inducible Gal promoter and a Cyc terminator or the p416-GPD vector (Mumberg *et al.* 1995) carrying a uracil selectable marker and a constitutive GPD promoter and Cyc terminator, for transformation into *Saccharomyces cerevisiae*. The DNA sequence of the constructed plasmids was confirmed by PCR and sequencing as described below.

Table 2: Primers used for construction of AMTP expression plasmids

| PCR and Sequencing primers (5' to 3') | |
|--|--|
| CgrayiA_F | CAATACTCACCATGGCTACCAAC |
| CgrayiA_R | GAAGCATTGGGTATCTCAACAAA |
| CgrayiAmpC_for_int | AGGTAAATTGCCCCGAGTCCT |
| CgrayiAMTpC_for_int2 | CGATTCTTGACTGATTGCACA |
| CgrayiAmpC_rev_int | GTTTATGAGGAAGGCGAGGA |
| CgrayiAmpC_rev_int2 | ATTGTGAGGACGGAAGTTGG |
| CgrayiAmpD_for_int | AGACCGAGGCTAACGATGTG |
| CgrayiAmpD_for_int2 | CTCATGCTGGGAAAGAGGAG |
| CgrayiAmpD_rev_int | GATGCCGAGAAGGAATTTGA |
| CgrayiAmpD_rev_int2 | TCCATGTCCAGCAAGCAATA |
| CgrayiB_F | CAATACTCACCATGGCTACCAAC |
| CgrayiB_R | GAAGCATTGGGTATCTCAACAAA |
| CgrayiC_F | TACTAAGCCAACCGGTACCTCTG |
| CgrayiC_R | CCCTCCCCCTACCCATGC |
| CgrayiD_F | CACAACAAGATGGCTTCAGGAC |
| CgrayiD_R | CGATGACAGTGTCGTAGCATTTC |
| MepA_F_nest2 | CTTCTTGGTCTTGGCTCTGG |
| MepA_F_nest3 | CTGTGCTGCTTTCGGTATCA |
| MepARev_nest1 | TGATACCGAAAGCAGCACAG |
| MepArev_nest2 | TGCTGCTACAAAACCAGCAG |
| Gapped plasmid construction plasmids (5' to 3') | |
| CgrayiA_p416F | AGTTTTAAAACACCAGAACTTAGTTTCGACGGATTCT AGAATGTCGTCCACCATCGCTGC |
| CgrayiA_p416R | CGATAAGCTTGATATCGAATTCCTGCAGCCCCGGGGGA TCCTTACTCTGTTTTCTTGGAGT |
| CgrayiA_PgalF | TCTATACTTTAACGTCAAGGAGAAAAAACCCTGGATC CACATGTCGTCCACCATCGCTGC |
| CgrayiA_PgalR | TCATAAATCATAAGAAATTCGCTTATTTAGAAGTTCTA GATTACTCTGTTTTCTTGGAGT |
| CgrayiC_p416F | AGTTTTAAAACACCAGAACTTAGTTTCGACGGATTCT AGAATGACGAGCTATCCCCGCTCC |
| CgrayiC_p416R | CGATAAGCTTGATATCGAATTCCTGCAGCCCCGGGGGA TCCCTACCCATGCGCCTGCCCCCT |
| CgrayiC_PgalF | TCTATACTTTAACGTCAAGGAGAAAAAACCCTGGATC CACATGACGAGCTATCCCCGCTCC |

Table 2 (continued): Primers used for construction of AMTP expression plasmids

| | |
|---------------|---|
| CgrayiC_PgalR | TCATAAATCATAAGAAATTCGCTTATTTAGAAGTTCTA GACTACCCATGCGCCTGCCCT |
| CgrayiD_PgalF | TCTATACTTTAACGTCAAGGAGAAAAACCCCGGATC CACATGGCTTCAGGACCGGTGCT |
| CgrayiD_PgalR | TCATAAATCATAAGAAATTCGCTTATTTAGAAGTTCTA GACTACCTCAAAGGCCCTCA |
| p416_D_F | AGTTTTAAAACACCAGAACTTAGTTTCGACGGATTCT AGAATGGCTTCAGGACCGGTGCT |
| p416_D_R | TGAGGGGGCCTTTGAGGTAGCCCCGGGCTGCAGGA ATTCGATATCAAGCTTATCG |
| p416_ScMep2_F | AGTTTTAAAACACCAGAACTTAGTTTCGACGGATTCT AGAATGTCTTACAATTTACAGG |
| P416_ScMEp2_R | CGATAAGCTTGATATCGAATTCCTGCAGCCCCGGGGGA TCCTTATACTATATGGTCAGTGT |
| P416B_F | AGTTTTAAAACACCAGAACTTAGTTTCGACGGATTCT AGAATGGCTACCAACCCACCATG |
| p416B_R | CGATAAGCTTGATATCGAATTCCTGCAGCCCCGGGGGA TCCTTAATCATGCTTTATCTCCC |
| p416F | CAGAACTTAGTTTCGACGGATTC |
| P416R | CGATAAGCTTGATATCGAATTCCT |

2.2.3 Construction of yeast strains

Yeast strains are listed in Table 3. Yeast transformations were performed using the TRAF0 lithium acetate method (Agatep *et al.* 1998) or by the gapped plasmid method (Raymond *et al.* 1999). Following the TRAF0 method, a yeast culture in log phase was centrifuged for 5 minutes at 6000 RPM, and the pellet washed with distilled water, and resuspended in an equal volume of distilled water, pelleted again, then resuspended in 1 ml of 100 mM lithium acetate. This solution was pelleted at maximum speed (>14,000 RPM) in a microcentrifuge and the pellet resuspended in approximately

400 μ l of 100 mM lithium acetate. A transformation mix of 50 μ l yeast cells, 240 μ l PEG, 36 μ l 1M lithium acetate, 50 μ l salmon sperm DNA and 34 μ l of diluted plasmid (approximately 100 ng plasmid) was incubated at 30 °C for 30 minutes and heat shocked at 42°C for 30 minutes. The mix was then centrifuged and the pellet resuspended in 300 μ l water. 100 μ l of the cell suspension was plated onto selective medium and transformants confirmed by colony PCR as described below.

For gapped plasmid construction, cDNA of ammonium transporters from *C. grayi* which had been cloned into the TopoTA vector (Invitrogen, Carlsbad, CA) was amplified with primers composed of 40 basepairs of homology on either side of the multiple cloning site of the centromeric plasmid p416GPD (Mumberg *et al.* 1995) carrying the uracil selectable marker and joined to the first or last 20 bp of the gene (Raymond *et al.* 1999). The PCR product was cleaned in a Microcon column (Millipore, Bilerica, MA) and approximately 1 μ g of amplicon was transformed into yeast along with 0.1 μ g of restricted plasmid DNA. Yeast recombine the two fragments into a functional plasmid. Transformants were screened by PCR and the inserts sequenced as described below.

2.2.4 DNA amplification and Sanger Sequencing

For PCR performed directly on colonies of yeast or *E. coli*, the PCR reaction mix consisted of 17.3 μ l distilled water, 2.5 μ l of 10x PCR buffer, 2.5 μ l dNTPs, 1.25 μ l each of a forward primer and a reverse primer, and 0.3 μ l Taq polymerase (Denville). Colonies

were touched lightly with the tip of a pipet which was then dipped briefly into the PCR tube and removed. PCR using was performed on a MJ Research PTC200 thermocycler or an Applied Biosystems (Foster City, CA) Veriti thermocycler using a cycling program consisting of 10 minutes at 94°C to lyse the cells and release the DNA, followed by 25 cycles of 30 seconds at 94°C, 30 seconds at 55° C and 2 minutes at 72°C, followed a final elongation step at 72°C for 7 minutes.

PCR products were visualized on a TAE 1% agarose gel stained with SYBR Safe (Invitrogen, Carlsbad, CA). If necessary, faint products or product with multiple bands were cloned with the TOPO TA cloning kit (Invitrogen, Carlsbad, CA) following the manufacturer's instructions. For each cloning reaction, at least 8 clones were screened by colony PCR using T7 and m13R primers and a PCR program consisting of a 10 minute initial denaturation step, followed by 25 cycles of 30 seconds at 94°C, 30 seconds at 52°C and 60 seconds at 72°C followed by a final elongation step of 7 minutes. PCR products were cleaned with a Montage PCR filter column (Millipore, Bilerica, MA) or with an Exo-SAP clean-up using 1 µl SAP dilution buffer, 0.5 µl Exonuclease 1, 0.5 µl Shrimp Alkaline Phosphatase, added to 10 µl PCR reaction and incubating on one of the aforementioned thermocyclers for 30 minutes at 37°C, then 15 min at 80°C . Cleaned PCR products were sequenced in 10 µl reactions using: 1 µl primer, 3 µl purified PCR product, 0.5 µl Big Dye (Big Dye Terminator Cycle sequencing kit, ABI PRISM version 3.1; PE Applied Biosystems, Foster City, CA), 1.5 µl Big Dye buffer, and 4 µl double-

distilled water. Automated reaction clean-up and visualization was performed at the Duke IGSP Genome Sequencing & Analysis Core Facility using Big Dye chemistry with an ABI 3730xl automated sequencer (PE Applied Biosystems, Foster City, CA). Sequencher version 4.8 (Gene Codes Corporation, Ann Arbor, MI) was used to edit sequences and assemble contigs.

2.2.5 Illumina sequencing

DNA from cultures of the lichenizing fungi *Acarospora strigata* (Acarosporomycetidae, Lecanoromycetes), *Arthonia cf rubrocincta* (Arthoniomycetidae), *Endocarpon pallidulum* (Eurotiomycetes) *Graphis scripta* (Ostropomycetidae, Lecanoromycetes) and from lichen thalli of *Dibaeis baeomyces* (Ostropomycetidae, Lecanoromycetes), *Leptogium* sp., *Physcia cf stellaris* (both Lecanoromycetidae, Lecanoromycetes), and *Peltula cylindrica* (Lichinomycetes) was prepared following the DTAB/CTAB method outlined in Armaleo and May, 2009. Briefly, tissue was ground to a fine powder under liquid nitrogen using a pre-chilled mortar and pestle. Twenty volumes of a DTAB solution (with 1% DNase-free RNase) was added to the powder and incubated for ~3 minutes at 65°C. Polysaccharides and other contaminants were precipitated by adding 1/3 volume 5 M NaCl and centrifuging at 6000 RPM for 5 minutes. The supernatant was removed and extracted once with one volume of phenol, after which was supernatant was removed and then extracted with one volume of chloroform. The supernatant was removed, and DNA precipitated with one volume of

isopropanol and left to incubate at room temperature for 5 minutes, then centrifuged at 6000 RPM for 5 minutes to pellet the DNA. The pellet was resuspended in DTAB+RNAse, with heating and gentle agitation. Remaining polysaccharides were precipitated by adding 1/3 volume of 5 M NaCl and centrifuging at 6000 RPM for 5 minutes. The supernatant was removed and undissolved material was pelleted by centrifugation at 6000 RPM. DNA was precipitated by adding 2 volumes of 100% ethanol, incubating at room temperature for five minutes, and centrifuging at maximum speed (14,000 RPM) for five minutes. The pellet was washed once with 70% ethanol, and allowed to dry. When dry, the pellet was resuspended in 100 μ l TE and quantified by Qubit. DNA was then further purified using the PowerClean kit (Mo Bio) following the manufacturer's instructions. DNA was again quantified after purification by Qubit.

2.2.6 Genome assembly

One to six μ g of purified DNA were submitted to the sequencing facility at Duke University for 72-basepair, barcoded Illumina sequencing using the HiSeq technology. Four lichen cultures were multiplexed onto a lane, and DNA from two lichen thalli (yielding both fungal and algal DNA) were multiplexed onto a lane. A total of three lanes were used to sequence the eight fungal genomes (~80 million to 600 million reads). Genomes were assembled using Velvet (Zerbino & Birney 2008) and various assembly parameters including k-mer (hash) sizes of 27 basepairs or 31 basepairs, and assembling total reads, the first 100 million reads or the first 50 million reads.

2.2.7 Phylogenetic analysis

Ammonium transporter sequences generated for this study were included in a modified alignment from McDonald et al 2011 and phylogenetic analysis was performed as in McDonald et al 2011. Genes in the MEP clade of AMTPs were retained in the analysis, while all others were discarded. All excluded regions were reexamined, and additional sites which became alignable were included in the analysis, including 90 basepairs of transmembrane region one, which in the previous analysis had been entirely excluded. The first well-supported clade outside of the MEP group in the MEP grade was used as an outgroup. In one analysis, the three ammonium transporters from the green alga *Asterochloris* sp., which fall into the AMT clade of ammonium transporters, were also included and used as an outgroup. Manual alignments were performed using MacClade 4.08 (Maddison and Maddison 2005). Ambiguously aligned regions and introns were delimited manually and excluded from phylogenetic analyses. Models of molecular evolution which had been selected using the Akaike Information Criterion (AIC) implemented in jModeltest (Posada 2008) or MrModeltest 2.3 (Nylander 2004) were implemented on this dataset. Phylogenetic relationships and confidence values were inferred using a maximum likelihood approach at the nucleotide level. Maximum likelihood analysis at the nucleotide level used GTR GAMMAI (with a gamma parameter and a proportion of invariable sites, = GTR+ Γ +I). The program RAxML-VI-HPC was used for the maximum likelihood search for the most likely tree.

The same program using the same settings was used for the bootstrap (BS) analysis with 1000 BS replicates run in batches of 100 replicates and pooled. Bootstrap values were calculated and visualized using the majority rule consensus tree command in PAUP 4.0d701 (Swofford 2002).

2.3 Results

2.3.1 *Cladonia grayi* Mep1a and Mep1b are functional ammonium transporters

To test if the proteins in the MEP α clade were functional transporters, each *Cladonia grayi* AMTP gene (i.e. *mep1a*, *mep1b* from the MEP α clade and *mep2* and *mep3* from the MEP γ clade) was introduced individually into a yeast strain lacking all three native yeast AMTPs. Transformed yeast strains were then grown on 5 mM ammonium as a sole nitrogen source, which allows growth only when a fungal strain has a functional AMTP. Growth on a plate containing 20 mM ammonium as a sole nitrogen source, an amount of ammonium sufficient to allow growth of the yeast strain lacking all ammonium transporters, was used as a positive control.

When incubated on medium containing 5 mM ammonium as the sole nitrogen source, the triple knockout strain completed a few cell divisions before arresting, while the triple knockout strain complemented with the *Saccharomyces cerevisiae* high-affinity ammonium transporter MEP2 as a positive control grew vigorously (Figure 4). The yeast triple knockout strain complemented with the *C. grayi mep1a* gene, the *C. grayi mep1b* gene (MEP α clade or the the *C. grayi mep2* gene (MEP γ clade) grew slightly less

well than the positive control, but appreciably better than the plasmid-only negative control. The diploid yeast triple knockout strain complemented with the fungal low-affinity *C. grayi mep3* gene grew vigorously, approaching the levels of the yeast triple knockout strain complemented with the high affinity AMTP from *S. cerevisiae*. Thus, all four *C. grayi* AMTPs seem to be functional if not efficient transporters of ammonia/ammonium.

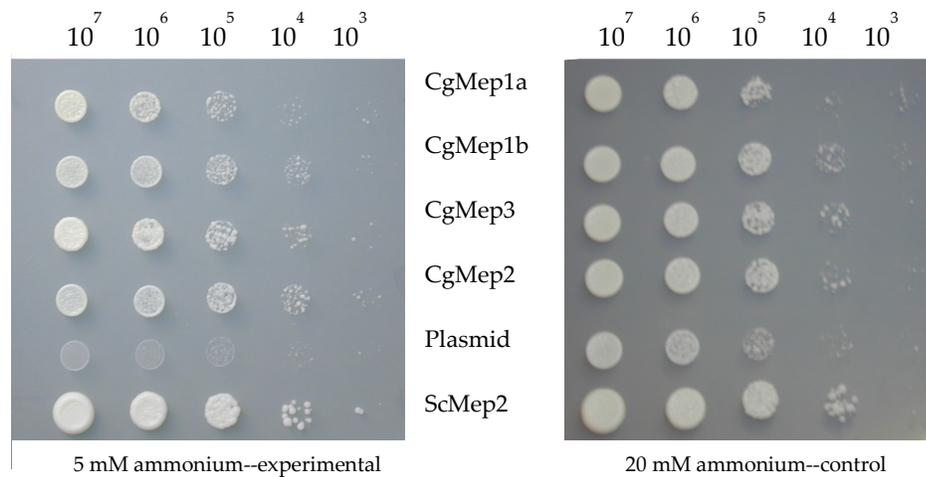


Figure 4. Heterologous expression of *Cladonia grayi* AMTP genes in a strain of *Saccharomyces cerevisiae* lacking all native AMTPs

Left: Ten-fold serial dilutions (10^7 to 10^3 cells/ml) of a *S. cerevisiae* (Sc) AMTP triple knockout complemented with AMTPs from the lichenizing fungus *Cladonia grayi* (Cg). Left, limiting ammonium (5 mM); right, control plate with excess ammonium (20 mM). Top to bottom, 1–6, transformants of MLY131 (a/ α) carrying a centromeric plasmid (p416-GPD) containing the following genes expressed the constitutive GPD promoter: 1. CgMep1a, 2. CgMep1b, 3. CgMep3, 4. CgMep2, 5. Plasmid only (negative control) 6. *Saccharomyces cerevisiae* Mep2 (positive control).

Table 3: Strains of *Saccharomyces cerevisiae* carrying ATMP genes from *Cladonia grayi*. All strains are congenic to Σ 1278b (Grenson *et al.* 1966) except S288C. All lichen genes are expressed from the p416-GPD plasmid unless otherwise noted.

| Strains (mating type) | Genotype | References and comments |
|--------------------------|---|---|
| MLY131 (a) | $\Delta mep1::LEU2, \Delta mep2::LEU2, \Delta mep3::G418, ura3-52, MAT a$ | Lorenz and Heitman, 1998 |
| MLY131 (α) | $\Delta mep1::LEU2, \Delta mep2::LEU2, \Delta mep3::G418, ura3-52, MAT \alpha$ | Lorenz and Heitman, 1998 |
| FD131 (a/ α) | $\Delta mep1::LEU2/\Delta mep1::LEU2, \Delta mep2::LEU2/\Delta mep2::LEU2, \Delta mep3::G418/\Delta mep3::G418, ura3-52/ura3-52, MATa$ | MLY131(a) x MLY131(α) |
| TRM Mep1a (a) | $\Delta mep1::LEU2/\Delta mep1::LEU2, \Delta mep2::LEU2/\Delta mep2::LEU2, \Delta mep3::G418/\Delta mep3::G418, ura3-52/ura3-52, MATa, ura4-CgMep1a$ | Transformant of MLY131(a) |
| TRM Mep1a (a/ α) | $\Delta mep1::LEU2/\Delta mep1::LEU2, \Delta mep2::LEU2/\Delta mep2::LEU2, \Delta mep3::G418/\Delta mep3::G418, ura3-52/ura3-52, MATa/\alpha, URA4-CgMep1a-CYC$ | Transformant of FD131(a/ α) |
| TRM Mep1b (a) | $\Delta mep1::LEU2/\Delta mep1::LEU2, \Delta mep2::LEU2/\Delta mep2::LEU2, \Delta mep3::G418/\Delta mep3::G418, ura3-52/ura3-52, MATa MATa, ura4::CgMep1b$ | Transformant of MLY131(a) (Expressed from pRS306-GAL1-TADH) |
| TRM Mep1b (a/ α) | $\Delta mep1::LEU2/\Delta mep1::LEU2, \Delta mep2::LEU2/\Delta mep2::LEU2, \Delta mep3::G418/\Delta mep3::G418, ura3-52/ura3-52, MATa MATa/\alpha, ura4::CgMep1b$ | TRMB x MLY 131(α) |
| TRM MepC (a) | $\Delta mep1::LEU2/\Delta mep1::LEU2, \Delta mep2::LEU2/\Delta mep2::LEU2, \Delta mep3::G418/\Delta mep3::G418, ura3-52/ura3-52, MATa, ura4::CgMepC$ | Transformant of MLY131 |
| TRM MepC (a/ α) | $\Delta mep1::LEU2/\Delta mep1::LEU2, \Delta mep2::LEU2/\Delta mep2::LEU2, \Delta mep3::G418/\Delta mep3::G418, ura3-52/ura3-52, MATa/\alpha, ura4::CgMepC$ | Transformant of FD131(a/ α) |

Table 3 (continued): Strains of *Saccharomyces cerevisiae* carrying ATMP genes from *Cladonia grayi*.

| | | |
|---------------------------|---|-------------------------------------|
| TRM MepD (a) | $\Delta mep1::LEU2/\Delta mep1::LEU2$, $\Delta mep2::LEU2/\Delta mep2::LEU2$, $\Delta mep3::G418/\Delta mep3::G418$, <i>ura3-52/ ura3-52</i> , <i>MATa ura4::CgMepD</i> | Transformant of MLY131(a) |
| TRM MepD (a/ α) | $\Delta mep1::LEU2/\Delta mep1::LEU2$, $\Delta mep2::LEU2/\Delta mep2::LEU2$, $\Delta mep3::G418/\Delta mep3::G418$, <i>ura3-52/ ura3-52</i> , <i>MATa MATa/α, ura4::CgMepD</i> | Transformant of FD131(a/ α) |
| TRM ScMep2 (a) | $\Delta mep1::LEU2/\Delta mep1::LEU2$, $\Delta mep2::LEU2/\Delta mep2::LEU2$, $\Delta mep3::G418/\Delta mep3::G418$, <i>ura3-52/ ura3-52</i> , <i>MATa MATa, ura4::ScMep2</i> | Transformant of MLY131(a) |
| TRM ScMep2 (a/ α) | $\Delta mep1::LEU2/\Delta mep1::LEU2$, $\Delta mep2::LEU2/\Delta mep2::LEU2$, $\Delta mep3::G418/\Delta mep3::G418$, <i>ura3-52/ ura3-52</i> , <i>MATa MATa/α, ura4::ScMep2</i> | Transformant of FD131(a/ α) |
| S288C | Wild type | |

2.3.2 Genome sequencing and assembly

Summary statistics for the raw genome data are shown in Table 4. Appendix C details the results of these explorations. Initial assemblies were performed with the full data and requiring 27 basepairs of perfect match before allowing two reads to contig. Twenty-seven basepairs was chosen since this value had yielded the best assemblies of *Ashbya gossypii* (Fred Dietrich, personal communication). Overall, the assemblies made using these parameters had few large contigs and many small contigs on the order of a few hundred basepairs. Assemblies were improved by reducing the amount of data included. Data were not sorted by quality for this exercise; rather, the first 50 million

Table 4: Statistics on the raw data for eight lichen genomes. Species names in bold type denote genome projects from culture (all reads represent the lichenizing fungus); species names in regular type denote genome projects from lichens (reads represent the lichenizing fungus, the lichenizing alga, and the associated microbial community). Estimated high quality coverage calculated by dividing the number of reads with at least the first 30 bases of high quality by the estimated genome size (of cultured fungi) or genome sizes (of lichenizing fungus and lichenizing alga but excluding the associated microbial community, and therefore overestimating coverage).

| | | |
|---------------------------------|-----------------------------------|---------------------------------------|
| | <i>Acarospora strigata</i> | <i>Arthonia cf rubrocincta</i> |
| Total sequences | 80464902 | 103562964 |
| first 30 bases high quality | 36326870 | 44129871 |
| first 50 bases high quality | 31585425 | 36870361 |
| first 70 bases high quality | 27701126 | 30691804 |
| all 75 bases high quality | 26289676 | 28531448 |
| Estimated high quality coverage | 68X | 82x |
| | <i>Dibaeis baeomyces</i> | <i>Endocarpon pallidulum</i> |
| Total sequences | 243144702 | 92018336 |
| first 30 bases high quality | 128915320 | 43595646 |
| first 50 bases high quality | 113229286 | 38937482 |
| first 70 bases high quality | 99892761 | 35132351 |
| all 75 bases high quality | 94764211 | 33697486 |
| Estimated high quality coverage | 80x | 81 x |
| | <i>Graphis scripta</i> | <i>Leptogium sp.</i> |
| Total sequences | 92331310 | 204893442 |
| first 30 bases high quality | 42665651 | 93698059 |
| first 50 bases high quality | 37267191 | 77111438 |
| first 70 bases high quality | 32664052 | 66199646 |
| all 75 bases high quality | 30956142 | 62669378 |
| Estimated high quality coverage | 173x | 148x |
| | <i>Peltula cylindrica</i> | <i>Physcia cf stellaris</i> |
| Total sequences | 152251382 | 201984964 |
| first 30 bases high quality | 79517806 | 90751190 |
| first 50 bases high quality | 67269367 | 78374611 |
| first 70 bases high quality | 57821683 | 68425794 |
| all 75 bases high quality | 28531448 | 64745297 |
| Estimated high quality coverage | 126x | 57x |

reads or the first 100 million reads were used regardless of quality. This reduction of data greatly improved the assemblies. Additionally, all the aforementioned assembly parameters were implemented using a 31-basepair k-mer size. In general, requiring more overlap of reads yielded a better assembly, although not in every case (Appendix C). None of the assemblies reached the median contig size reported for Velvet assemblies using paired end data or single read data (Zerbino and Birney, 2008), suggesting that more optimization is possible. Sorting the data by quality before assembly, especially when using partial data, could likely improve the assemblies. For samples containing both fungal and algal DNA, an initial segregation of fungal and algal data so that assembly of the two genomes could proceed more or less independently would likely improve the assembly.

2.3.3 A panel of 20 test genes is used to assess completeness of genome assembly

Since it is not possible to argue absence of a gene from the genome from the failure of PCR reactions, there is no way to definitively assess absence of a gene from a genome except the sequencing of the whole genome. Even then, however, it could be argued that a gene is not found simply because the assembly is incomplete, meaning that the gene is in fact in the genome of interest but is missing from the genome assembly. In order to test the completeness of the assembled genomes, 20 putatively single-copy nuclear genes were chosen as test queries. The 20 test genes selected are conserved throughout the fungi (Miadlikowska et al, unpublished data) and are of

roughly comparable size to the ammonium transporter genes. If part or all of each of the 20 test sequences could be recovered from a genome, then it was considered likely that homology searches for genes of interest, i.e. ATMPs and particularly the MEP α gene(s), would identify at least a portion of the gene(s) if such were present in the genome.

All 20 of the test genes were recovered from each of the eight genomes (Table 5). In addition, unexpected duplicate copies of some of the genes, including Fal1 and Mcm7, were discovered. It was therefore concluded that the genomes were complete enough to recover genes of interest, and further to lend confidence to calls of absence were the genes not recovered. It was deemed important to be able to comment on absence since initial attempts to amplify MEP α from several of the lineages included here consistently failed. Using the four *Cladonia grayi* MEP genes as queries, the MEP genes from all eight genomes were identified (Table 5) as described below.

2.3.4 Plant-like MEP α AMTPs are missing from the genomes of lichens growing in high-nitrogen habitats

In the work described in the previous chapter, the MEP α genes were found in the Lecanorales but not in any other order sampled within the subclass Lecanoromycetidae (Figure 5). However, the failure to amplify the MEP α genes by degenerate PCR is not definitive proof of loss; only a full genome sequence can be used to demonstrate loss of a gene. To this end, two additional orders within the Lecanoromycetidae, the Caliciales (Gaya *et al.* 2011) and the Peltigerales, were sampled

Table 5: Presence/absence in the assembled genomes of eight lichen fungi of ammonium transporter/ammonia permease genes and of 20 conserved test genes.

Because each of the 20 test genes could be identified in each genome, assemblies were of sufficient quality that a homology search should identify ammonium transporter genes were actually present in a genome.

| | <i>Acarospora strigata</i> | <i>Arthonia cf rubrocincta</i> | <i>Dibaeis baeomyces</i> | <i>Endocarpon pallidulum</i> | <i>Graphis scripta</i> | <i>Leptogium sp.</i> | <i>Peltula cylindrica</i> | <i>Physcia cf stellaris</i> |
|--|----------------------------|--------------------------------|--------------------------|------------------------------|------------------------|----------------------|---------------------------|-----------------------------|
| <i>mep1a</i> (MEP α) | 0 | 0 | X | 0 | X | 0 | 0 | 0 |
| <i>mep1b</i> (MEP α) | 0 | 0 | X | 0 | 0 | 0 | 0 | 0 |
| <i>mep2</i> (MEP γ , high-affinity) | X | X | X | X | 0 | X | X | X |
| <i>mep3</i> (MEP γ , low-affinity) | X | X | X | X | X | X | X | X |
| <i>amd1</i> | X | X | X | X | X | X | X | X |
| <i>bub2</i> | X | X | X | X | X | X | X | X |
| <i>cct2</i> | X | X | X | X | X | X | X | X |
| <i>cox15</i> | X | X | X | X | X | X | X | X |
| <i>ctk1</i> | X | X | X | X | X | X | X | X |
| <i>dpb3</i> | X | X | X | X | X | X | X | X |
| <i>eft2</i> | X | X | X | X | X | X | X | X |
| <i>fal1</i> | X | X | X | X | X | X | X | X |
| <i>frs2</i> | X | X | X | X | X | X | X | X |
| <i>gln4</i> | X | X | X | X | X | X | X | X |
| <i>krr1</i> | X | X | X | X | X | X | X | X |
| <i>mcm7</i> | X | X | X | X | X | X | X | X |
| <i>ost1</i> | X | X | X | X | X | X | X | X |
| <i>ret1</i> | X | X | X | X | X | X | X | X |
| <i>rio1</i> | X | X | X | X | X | X | X | X |
| <i>rpa135</i> | X | X | X | X | X | X | X | X |
| <i>rpb1</i> | X | X | X | X | X | X | X | X |
| <i>rpb2</i> | X | X | X | X | X | X | X | X |
| <i>rpo31</i> | X | X | X | X | X | X | X | X |
| <i>sec15</i> | X | X | X | X | X | X | X | X |

here. One more order, the Teloschistales, will shortly be represented by a genome of the lichenizing fungus *Xanthoria parietina* produced by another group (Paul Dyer, personal communication) so it was not sampled for this work. These orders are particularly interesting in terms of nitrogen tolerance and acquisition.

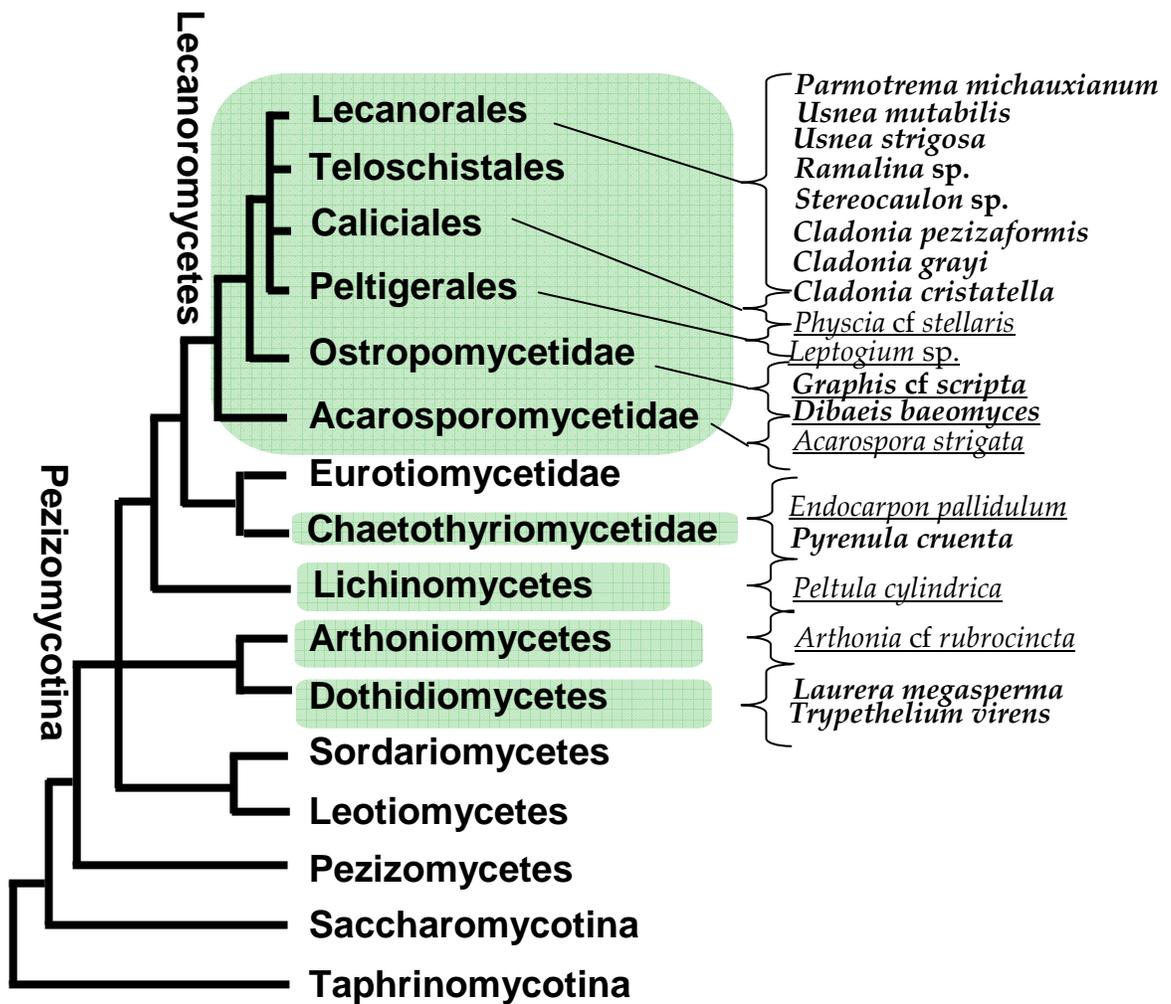


Figure 5: Simplified phylogenetic tree of the Ascomycota showing the placement of the eight lichenizing fungi used for genome sequencing. Lichenizing fungi used for genome sequencing are underlined. Species with plant-like transporters in the MEP α clade are shown in bold.

Lichens in the Teloschistales and Caliciales tend to grow in nitrogen-rich environments, such as rocks or tree limbs on which birds perch. Because nitrogen is not limiting in these environments, it was hypothesized that the failure to amplify by degenerate PCR any MEP α genes in the lichens in the Teloschistales and Caliciales (formerly a suborder of the Teloschistales) sampled for the work described in Chapter 1, including *Xanthomendoza*, *Caloplaca* or *Physcia*, was evidence that the plant-like AMTPs genes had been lost in the Teloschistales and Caliciales. No plant-like AMTP was found in *Physcia cf stellaris* as predicted (Table 4), while a high-affinity fungal AMTP and a low-affinity fungal ATMP were identified, as were all 20 test genes, suggesting that the assembly is complete enough to recover a MEP α gene were it present in the genome. If no plant-like ammonium transporter is found in *Xanthoria parietina*, the hypothesis that lichens in the Teloschistales and Caliciales have lost the plant-like AMTP will be strengthened. The loss of plant-like AMTPs in the lichens in these orders might suggest why many lichens in these orders are constrained to living in high-nitrogen environments; alternatively, the expansion of lichens in the Teloschistales and Caliciales into high nitrogen niches may have removed the selective pressure to retain the MEP α genes.

2.3.5 Plant-like MEP α AMTPs are missing from the genomes of lichens that are symbiotic with nitrogen-fixing cyanobacteria

Another order of lichens within the Lecanoromycetidae for which no plant-like ammonium transporters were amplified by degenerate PCR is the Peltigerales. This order is primarily composed of lichens in which the photobiont is a nitrogen-fixing cyanobacterium (*Nostoc*). Also in this order are lichens with both a green alga and a cyanobacterium as photobionts, but these lichens have secondarily regained the green algal photobiont (Miadlikowska & Lutzoni 2004). Data from two unpublished genome projects on various species of *Peltigera* showed that each species of *Peltigera* had one high affinity fungal ammonium transporter and one low-affinity fungal ammonium, but no plant-like ammonium transporters (Ólafur S. Andr sson, personal communication and unpublished data; Bernard Goffinet, unpublished data). The fact that one lichen genus which is symbiotic with a nitrogen fixer seemed to have lost the plant-like ammonium transporters led to the hypothesis that cyanolichens in general have lost the Mep α AMTPs. To test this hypothesis, the lichen *Leptogium* sp. in a separate suborder of the Peltigerales was chosen for genome sequencing. *Leptogium* is in the family Collemataceae in the suborder Collematineae within the Peltigerales (Miadlikowska & Lutzoni 2004). Both high-affinity fungal and low-affinity fungal AMTPs were identified in the genome of *Leptogium*, as were all 20 test genes. However, no MEP α AMTPs were found. Thus, because lichens in both suborders of the Peltigerales lack the plant-like ammonium transporters, it is probable that all the lichens in the Peltigerales lack the

plant-like ammonium transporters. Previous studies have reconstructed the ancestral state of the Peltigerales as a bimembered association with a cyanobacterium (Miadlikowska & Lutzoni 2004). It is possible that, having acquired a nitrogen fixer as a symbiont, the need for the nitrogen presumably provided by the horizontally transferred AMTPs was alleviated, allowing this gene to be lost during the early evolution of the Peltigerales. This suggests that the Mep α AMTPs will not be found in any of the green-algal associated members of the order either, as the symbiosis with a green alga is secondarily acquired. This may also suggest a reason why even most green algal members of the order continue to retain cyanobacteria in cephalodia—to supplement the nitrogen budget. Likewise, it is unusual for a fungus to have just two AMTP genes (Appendix A), particularly a filamentous ascomycete. Perhaps the presence of a nitrogen fixing symbiont provides sufficient nitrogen so that there is no pressure to duplicate the AMTPs.

In addition to the Peltigerales, cyanolichens are also found in the Lichinomycetes, a small class of lichenizing fungi evolutionarily distant from the bulk of lichens found in the Lecanoromycetes (Figure 5). Lichen-forming fungi in the Lichinales are mostly symbiotic with cyanobacteria other than *Nostoc*, such as *Anacystis* (Brodo et al. 2001). It was hypothesized that, like the lichens in the Peltigerales, lichens in the Lichinomycetes would also have lost the MEP α AMTPs upon acquisition of a cyanobacterium as a photobiont. To test this hypothesis, the genome of *Peltula cylindrica*

was sequenced. No MEP α AMTPs were found in this genome. However, a fungal high affinity AMTP and a fungal low-affinity AMT were identified, as were all 20 test genes. Thus, it would appear that *Peltula* also has lost the MEP α AMTP. It seems likely that the other members of this class which are also symbiotic with cyanobacteria will also have lost the MEP α AMTP.

2.3.6 Plant-like AMTPs are found in the Ostropomycetidae but not the Acarosporomycetidae

If loss of MEP α AMTPs is strictly associated with the photobiont of a lichen, then it would be expected that lichens with green algal photobionts should retain the MEP α gene(s). However, the failure to identify MEP α genes in other subclasses within the Lecanoromycetes by degenerate PCR argues against this hypothesis. To determine whether MEP α are present in other subclasses within the Lecanoromycetes, namely the Ostropomycetidae and the Acarosporomycetidae (Figure 5), the genomes of *Dibaeis baeomyces* (Ostropomycetidae) and *Graphis scripta* (Ostropomycetidae) and *Acarospora strigata* (Acarosporomycetidae) were sequenced.

MEP α genes were identified in *Graphis cf. scripta* and *Dibaeis baeomyces*, both members of the Ostropomycetidae (Table 5, Figure 5). The Ostropomycetidae is the sister subclass to the Lecanoromycetidae, the subclass in which are placed most of the lichens in which MEP α genes had previously been identified. Previous attempts to amplify MEP α genes from the Ostropomycetidae had failed, so this finding represents an extension of the known distribution of MEP α genes. Further, *Dibaeis* and *Graphis* are

in different orders within this subclass, suggesting that the MEP α gene may be widespread in the suborder.

The earliest diverged subclass within the Lecanoromycetidae is likely to be the Acarosporomycetidae (Figure 5), represented here by *Acarospora strigata*. No plant-like ammonium transporters were found in this genome, which was surprising as even lichens in distantly related classes, such as *Trypethelium virens* in the Dothidiomycetes and *Pyrenula cruenta* in the Chaetothyriomycetidae were shown to have the MEP α genes. Again, the fact that all 20 test genes were recovered and that other ammonium transporters were found in the genome suggests that the gene is actually absent from the genome, not just missing from the assembly. Possibly a wider sampling of lichens within the Acarosporomycetidae would reveal fungi in this subclass that retain the MEP α genes, or could reveal some other factor influencing the loss or retention of AMTP genes.

2.3.7 Distribution of MEP α in the Chaetothyriomycetidae is patchy

Previously, one lichen in the subclass Chaetothyriomycetidae (*Pyrenula cruenta*) was shown to have a MEP α gene. Thus, it was expected that a second lichen in that subclass but in a different order might also have a MEP α gene in its genome. But, genome sequencing of *Endocarpon pallidulum* revealed that no MEP α AMTP was present in the genome. Again, high affinity fungal AMTPs and low-affinity fungal AMTPs were

identified, as were all 20 test genes, suggesting that the failure to recover a MEP α gene from this genome is not due to incompleteness of the assembly.

2.3.8 The MEP α gene replaces the high-affinity MEP γ gene in *Graphis scripta*

Low-affinity fungal AMTPs falling into the MEP γ clade were found in all genomes, as expected given the presence of these genes in the genomes of all Dikarya fungi sequenced to date (Table 5). Likewise, high-affinity fungal AMTPs in the MEP γ clade have been found in all fungi to date. However, high affinity fungal ATMPs were recovered in only seven of the eight lichen genomes sequenced here. *Graphis scripta* appeared to be lacking a high-affinity MEP γ gene, which is unprecedented among the Dikarya fungi. The similarity of MEP genes is high throughout the transmembrane domains, such that any MEP gene used as a query will find all other MEP genes in the genome. Furthermore, the *Graphis scripta* genome was produced from the cultured fungus, and the genome is very high coverage (173x). Therefore, it is unlikely that no portion of the high-affinity MEP γ gene was sequenced if it were in the genome. Thus, it seems likely that this fungus actually lacks the high-affinity MEP γ gene.

Other lichens like *Peltigera* spp. or *Acarospora strigata* that have only two ammonium transporters have the fungal high affinity gene and the fungal low affinity gene. *Dibaeis baomyces*, which like *Graphis scripta* is in the Ostropomycetidae, has two copies of a MEP α gene as well as one fungal high affinity ammonium transporter and one fungal low-affinity ammonium transporter, suggesting that perhaps the ancestor to

the Ostropomycetidae or even the Lecanoromycetes had all four ammonium transporters. It would appear that in *Graphis* the MEP α gene has replaced the low-affinity fungal AMTP from the MEP γ clade. *Graphis* has a very thin crustose thallus, quite unlike the large thalli produced by many lichens in the Lecanorales. It is perhaps no surprise that less nitrogen is needed to support this comparatively reduced thallus. That the MEP α gene is retained in preference to the high affinity fungal ammonium transporter suggests that it was perhaps more efficient than the high affinity fungal ammonium transporter.

2.3.9 Phylogenetic tree

The sequences of the various MEP genes mined from the eight lichen genomes were included in a phylogenetic analysis based on the alignment used in Chapter 1, but restricted to the MEP clade (Appendix D). All MEP genes were used, including fungal MEP genes and MEP genes from the cyanobacterial symbionts of *Peltula cylindrica* and *Leptogium* sp. Green algal AMTPs all fall into the AMT family and thus were outside of the scope of this particular analysis. No algal AMT genes were recovered from *Dibaeis baeomyces*, since the DNA was prepared from podetia which, while being symbiotic tissue, are largely free of algal cells. No algal AMT genes from *Physcia cf stellaris* were included because the small average contig size (~150 basepairs) made it difficult to be certain how to assemble the various pieces of the alga AMTs recovered. The remaining

four genome assemblies were produced from cultured fungus and contained no algal sequences.

Restricting the analysis to the MEP clade allowed an additional 90 sites corresponding to transmembrane region 1 to be added to the analysis. It was hoped that these extra data would increase the support on critical branches. This was not the case. While the MEP family remained significantly supported, as did the MEP α clade and the MEP γ clade, the MEP β clade did not become significantly supported in this analysis (Appendix D). Within the MEP α clade, the land plant MEP α clade remained significantly supported, but the fungal MEP α clade lost support.

In general, the sequences from the lichenizing fungi were placed as expected in the high-affinity fungal clades and low-affinity fungal clade of the MEP γ clade. One exception was a small clade of one sequence from *Phycia cf stellaris* and one sequence from *Peltigera* sp. These sequences were placed closer to sequences from the Eurotiomycetes than to sequences from other Lecanoromycetes. Since the genomes of both of these lichens were prepared from DNA extracted from a thallus rather than from a culture, it is possible that these sequences represent AMTPs from endolichenic fungi rather than from the lichenizing fungus.

2.4 Discussion

Nitrogen is an important currency in the cell, arguably second in importance only to carbon. A recent work on fungal genomes has validated 323 horizontal gene transfer events into fungi (Richards *et al.* 2011), of which the two top categories are genes involved in the acquisition and metabolism of carbon and of nitrogen. If all the categories involving nitrogen are grouped, horizontal gene transfer of genes involved in nitrogen acquisition and metabolism are in fact the most populous category. It thus appears that fungi have been capturing new technology for nitrogen acquisition since the very beginning.

AMTPs in particular show an interesting pattern of expansion in the fungi. Preliminary data from the genome of *Rozella*, the earliest diverging lineage of the Fungi known (James *et al.*, 2006), suggests that *Rozella* entirely lacks AMTP genes (Tim James, personal communication). The microsporidia, also considered early diverging fungi (James *et al.*, 2006) also lack AMTPs. Interestingly, *Rozella* and the microsporidia are intracellular pathogens. Intracellular pathogens tend to experience genome contraction as they outsource more and more of their life functions to their hosts. If other early fungi also lack AMTP genes, it would lend support to the hypothesis raised in Chapter 1 that fungi first lost the eukaryotic ammonium transporters and then obtained by horizontal gene transfer a prokaryotic AMTP. In fact, the seemingly high levels of horizontal gene transfer seen in the fungi could be explained similarly, as fungi slowly

rebuilding their genomic toolkits after a period of gene loss as intracellular pathogens during their early evolution.

After a period of time with no AMTPs, fungi acquired by horizontal gene transfer a bacterial AMTP. Sampling of the Neocallimastigomycota and the Monoblepharidomycetes (chytrids *sensu lato*) is poor, so it isn't yet possible to pinpoint the entry of AMTPs into the fungi, but AMTPs had entered the fungi by the time of the divergence of the chytrid *Batrachomyces*. Then, as the Dikarya diverged there was a duplication event followed by a subfunctionalization into high-affinity (low-capacity) and low-affinity (high-capacity) ammonium transporters. In mycorrhizal basidiomycetes (but not rusts or smuts), a gene family expansion followed. In ascomycetes, there may also have been a slow gene family expansion perhaps starting as early as the Taphrinomycotina. Layered on top of this slow gene family expansion was a second horizontal transfer from hyperacidophilic chemoautolithotrophic prokaryotes into the leotiomyceta, followed by subsequent losses in almost all non-lichenized lineages and in lichenized lineages with a rich enough internal or external source of nitrogen. Interestingly, lichenized lineages appear not to have duplicated the fungal high-affinity or fungal low-affinity AMTPs, suggesting that the AMTP gene family expansion in non-lichenized leotiomyceta may post-date the acquisition and subsequent loss of the MEP α gene. Instead, the MEP α gene may have been duplicated either early in the evolution of the leotiomyceta or independently in the Lecanorales and in the

Dothideomycete lichens. The retention of the MEP α gene apparently obviates the need to multiply the fungal high-affinity and fungal low-affinity genes, since each lichenized fungus sequenced retains exactly one copy of high-affinity and low-affinity fungal AMTPs.

What, then, does the MEP α gene do for the lichenizing fungus, especially in light of the fact that the MEP α genes are evidently rapidly lost (as in both lineages of cyanolichens sequenced, among others) after selection pressure is removed? The MEP α genes must not be intimately involved in the lichen symbiosis, assuming that there is some ancestral core “symbiosis program” shared by ascomycete lichens. The MEP α genes are also evidently not absolutely mandatory to supplying nitrogen to the lichenizing fungus, since they are missing in lichens like *Acarospora* and *Endocarpon* which are not growing in high-nitrogen environments and are not symbiotic with cyanobacteria. It is possible that the MEP α genes are involved specifically with balancing the nitrogen budget between the fungal and green algal symbionts. If lichenization is considered a controlled parasitism of the photobiont by lichenizing fungi, it follows that the fungus would also control the nitrogen budget of the alga, perhaps by outcompeting the photobiont for nitrogen but then exporting, for example, amino acids or some other form of nitrogen on which the alga is dependent. Alternatively, the alga may also scavenge its own ammonium, and may in fact be fed by

the fungus releasing ammonium through ammonium exporters (e.g. those encoded by the *ATO* genes in yeast).

To further explore the reasons behind the loss of the MEP α AMTPs in lichens, it would be wise to explore tripartite symbioses. One lichen sampled in Chapter 1, *Stereocaulon tennessense*, is a tripartite lichen symbiotic with both a primary photobiont that is a green alga and a secondary photobiont that is a cyanobacterium which is confined to a specialized structure called a cephalodium. This genus has retained the MEP α gene. Were the presence of MEP α AMTPs strictly negatively correlated with the presence of a cyanobacterial nitrogen fixer (i.e, if presence of cyanobacteria in the lichen thallus means that the MEP α AMTPs are lost) then *Stereocaulon* should have lost its MEP α AMTP. So it cannot be said that the loss of the MEP α AMTP is a consequence of acquiring a nitrogen fixer. Taking the data from Peltigerales and the Lichinomycetes into account, it seems more specifically that being dependent exclusively, or in large part, on cyanobacteria for nitrogen precipitates the loss of the MEP α genes. To test this refined hypothesis, one could examine other tripartite lichens in the Lecanorales that are symbiotic with cyanobacteria in cephalodia, for example, *Pilophorus* (Cladoniaceae), *Stereocaulon* (Stereocaulaceae), and *Arctomia* (Arctomiaceae, Table 6) . All lichens in the Lecaromycetes sampled to date have the MEP α genes, so it seems likely that all the aforementioned lichens will too regardless of the presence of a cyanobacterial secondary photobiont. Outside of the Lecanorales, there are tripartite lichens in the

Ostropomycetidae such as *Placopsis* (Agyriales) and *Coccotrema* (Pertusariales) (Brodo 1973). Given that both members of the Ostropomycetidae sampled here retained the plant-like genes, it seems possible that these lichens with cephalodia will also retain the MEP α genes, although the loss of the MEP γ homologue in *Graphis scripta* suggests that there is selective pressure to reduce the copy number of AMTPs in the genome, and it may be easier to lose the more recently horizontally acquired MEP α genes rather than the less recently acquired MEP γ genes. There are no reports of lichens with cephalodia in the Acarosporomycetidae, but there are also lichens with cephalodia in the Eurotiomycetes (Chaetotheriomycetidae) such as *Sporodictyon* (Verrucariaceae, Eurotiomycetes) (Gueidan *et al.* 2007; Savic *et al.* 2008). Given the sporadic loss of the MEP α genes in the Chaetotheriomycetidae, it is unclear whether tripartite members of

Table 6: Summary of systematic placement of some lichen genera

| Genus | Family | Order | Subclass | Class |
|---------------------|-----------------|---------------|-----------------------|-----------------|
| <i>Pilophorus</i> | Cladoniaceae | Lecanorales | Lecanoromycetidae | Lecanoromycetes |
| <i>Stereocaulon</i> | Stereocaulaceae | Lecanorales | Lecanoromycetidae | Lecanoromycetes |
| <i>Arctomia</i> | Arctomiaceae | Lecanorales | Lecanoromycetidae | Lecanoromycetes |
| <i>Peltigera</i> | Peltigeraceae | Peltigerales | Lecanoromycetidae | Lecanoromycetes |
| <i>Placopsis</i> | Agyriales | Agyriales | Ostropomycetidae | Lecanoromycetes |
| <i>Coccotrema</i> | Coccotrema | Pertusariales | Ostropomycetidae | Lecanoromycetes |
| <i>Peltula</i> | Peltulaceae | Lichinales | N/A | Lichinomycetes |
| <i>Sporodictyon</i> | Verrucariaceae | Verrucariales | Chaetothyriomycetidae | Eurotiomycetes |

this subclass will retain or lose the MEP α genes. And of course, any clade arising from a bipartite cyanolichen is hypothesized to lose the MEP α genes.

Because *Acarospora strigata* and *Endocarpon pallidulum* have lost the MEP α AMTPs but are symbiotic with a green alga, there must be additional factors beyond type of photobiont or the nitrogen richness of the substrate that precipitate the loss of the MEP α AMTPs. What these factors might be is not immediately obvious, although from the data presented here there is some suggestion that thallus size (i.e. macrolichens versus microlichens) may play a role. More sampling from the Acarosporomycetidae and perhaps other clades of uncertain placement within the Lecanoromycetes could shed light on this question. For example, the Umbilicariaceae are large lichens. Based on their large size and their preferred substrate, which is rock, these lichens should have retained the MEP α AMTPs. An ecologically similar lichen genus is *Dermatocarpon*, which is in the Chaetothyriomycetidae. Based on habitat and thallus size, this lichen should also retain the MEP α AMTPs, although at least one lichen in this subclass has lost the MEP α AMTPs.

Additional sequences from the Acarosporomycetidae and sequences from members of the Candelariaceae could clarify if the early diverging lichens within the Lecanoromycetes have the spottily distributed pattern of MEP α genes as in the Chaetothyriomycetidae, or if the MEP α genes have been lost entirely, as is hypothesized for the Lichinomycetes and the Peltigerales.

Given, on one hand, the apparent ease of loss of the MEP α genes relative to the AMTPs in the MEP γ clade, and on the other hand, the apparent selective pressure to

retain these genes in lichenizing fungi relative to non-lichenizing fungi (McDonald *et al.* 2011), it would appear that the MEP α genes have some sort of function, at least a vestigial function, in lichens. In *Cladonia grayi*, these genes were shown to be functional ammonium transporters, suggesting that this function may be retained in the AMTP genes of other lichens. Aside from transporting ammonium/ammonia, AMTP proteins (specifically, only high-affinity fungal AMTPs) have also been shown to be involved in pseudohyphal filamentation in *Saccharomyces cerevisiae* and in *Cryptococcus neoformans*, where they serve as sensors of nitrogen starvation (Lorenz and Heitman 1998; Rutherford *et al.* 2008). High-affinity ammonium transporters from the basidiomycetes *Ustilago maydis* and *Hebeloma cylindrosporum* can complement this function in strains of *Saccharomyces cerevisiae* with the high-affinity ammonium transporter *MEP2* deleted (Javelle *et al.* 2003; Smith *et al.* 2003). Presumably, the high-affinity fungal ammonium transporter from *C. grayi* should also complement this function, while the low-affinity fungal AMTP should not, in keeping with the functional data from other fungi. It is not yet known whether the MEP α AMTPs also can stimulate pseudohyphal filamentation, since they are high-affinity ammonium transporters, or in a perhaps more likely scenario, will fail to do so due to amino acid sequence differences in regions shown to be critical for pseudohyphal filamentation (Lorenz and Heitman 1998).

Lastly, transcriptional profiling of the fungus *C. grayi* grown on nitrogen-rich or nitrogen-poor media would be helpful. Identifying genes co-regulated with the MEP α

AMTPs could reveal whether MEP α AMTPs are involved in mating, morphological change, growth, nutrient scavenging, or other pathways. Understanding this could shed light on why these genes have been retained in some lichenizing fungi but not others, and why these genes were horizontally transferred in the first place.

3. DNA methylation of the lichen *Cladonia grayi* in the symbiotic and free-living states

3.1 Introduction

Epigenetics is the study of heritable changes in gene expression involving modifications of chromatin that do not alter the underlying sequence of genomic DNA. Put in another way, epigenetics is the study of the various “flags” or combination of flags that mark a stretch of chromatin as actively transcribed euchromatin or as closed, tightly packed heterochromatin. Generally, these flags remain associated with a particular region throughout the cell cycle and allow for information about the region to be transmitted from one cell division to the next. Epigenetics is concerned not only with the nature of the flags themselves, but also the meaning carried by the flags or combination of flags and the machinery that places and removes the flags from chromatin.

The fundamental repeating units of chromatin-- the units onto which flags are placed--are nucleosomes. Nucleosomes are the “beads” on the chromatin string famously captured in the micrograph by Oudet and co-authors (Oudet *et al.* 1975). Nucleosomes are composed of DNA and of small, positively charged proteins called histones. The protein component of each nucleosome is an octamer made up of two of each of four core histones: H2A, H2B, H3 and H4. Histones H3 and H4 form a tetramer, which is flanked by two dimers of H2A/H2B to complete the nucleosome. Around this octamer is wrapped a 147-basepair length of DNA. Most properly, the term nucleosome

refers both to the core histones and to the DNA wrapped around it. Two additional histones that are not part of the core, H1 and H5, are considered linker histones and serve to bind adjacent nucleosomes together to compact chromatin into higher order chromatin fibers.

Each of the core histones is composed of a central domain, which contacts both DNA and the other histones in the octamer, and an N-terminal “tail” which extends out of the nucleosome into the nucleoplasm. The N-terminal tail is subject to post-translation modifications (flags) including acetylation, methylation, phosphorylation, ubiquitination, and sumoylation, of certain amino acids. Lysines may be acetylated, methylated, sumoylated or ubiquitinated, while serines and threonines may be phosphorylated. Not every lysine or serine is actually modified in this way—certain lysines are subject to methylation, for example, while others are not. Histone (lysine) acetylases (HATs or KATs), histone (lysine) deacetylases (HDACs or KDACs), and histone (lysine) methyltransferases (HMTs or KMTs), are examples of the chromatin modifying proteins which modify histones by adding these post-translational modifications, hereafter referred to as “marks”.

Chromatin modifying proteins often recognize only one, or at most a few, sites along the histone tail. For example, a particular histone methyltransferase may recognize only histone H3 at lysine 9, but not histone H3 at lysine 27, nor histone H4 at lysine 20, which are recognized by separate histone methyltransferases. The specificity

of each chromatin modifying enzyme is determined by the amino acid sequence surrounding the target amino acid, providing a binding platform for the chromatin modifying enzyme. The specificity of chromatin modifying enzymes, and early indications that post-translational modifications singly or in combination marked stretches of chromatin as certain varieties of euchromatin or heterochromatin, i.e centromeric heterochromatin versus active euchromatin for instance, led to the idea of a finite language of histone modification which could subdivide chromatin into different types, each recognizable by the unique signature of post-translational modifications to the histones in that neighborhood and by the proteins recruited and bound to the modified histones. This idea came to be known as the “histone code” (Strahl & Allis 2000) by loose analogy with the DNA code.

When first advanced, it was believed that these chromatin marks, or a subset of these marks, might be more or less universal, but recent work shows the picture to be much more variable, not just between kingdoms but even within kingdoms and phyla. Although not truly universal, there are nevertheless some generalizations that can be made about post-translational histone modifications that specify different types of chromatin. In general, open, accessible, actively transcribed euchromatin is characterized by acetylation of histones. Acetylation changes a positively charged lysine residue to a more negatively charged residue, decreasing the affinity for DNA which is negatively charged, and thus opening up the chromatin and allowing the transcriptional

machinery access to the DNA. Constitutive heterochromatin, on the other hand, is characterized by histones with deacetylated lysines. Heterochromatin may also be characterized by methylation of certain lysines, particularly di- or tri-methylation of H3K9, as well as the presence of scaffolding proteins like heterochromatin binding protein 1 (HP1 or Swi6). All of these tend to compact chromatin, making the DNA less accessible to the transcription machinery.

The composition of heterochromatin is not universal; the heterochromatin of different organisms may have some or all of the aforementioned marks. In addition, different kinds of heterochromatin within one organism may have different suites of marks. For example, pericentromeric heterochromatin in *Schizosaccharomyces pombe* is characterized by dimethylation of histone H3 at lysine 9 and by proteins which specifically bind this epigenetic mark, whereas the heterochromatin at the central kinetochore region of the centromere lacks dimethylation of H3 at lysine 9, and instead is characterized by nucleosomes containing a variant of histone H3 called CENP-A (Allshire and Selker, 2007).

Much that is known about the formation of heterochromatin comes from the study of fungi. Easy and fast to grow and amenable to genetic manipulation, they are superior genetic and epigenetic models. In particular, three model fungi have contributed greatly to the understanding of heterochromatin. These fungi are the ascomycete yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, and the

filamentous ascomycete *Neurospora crassa*. Interestingly, in all these fungi the composition of heterochromatin is slightly different.

Of the three fungi, heterochromatin formation in *N. crassa* is closest to that in plants and animals. It is characterized by trimethylation of histone H3 at lysine 9 (H3K9me3), heterochromatin binding protein 1 (HP1) and DNA methylation. Heterochromatin is found at the centromeres, telomeres and relics of transposons and other repeated elements.

In *S. pombe* as in *N. crassa*, heterochromatin is found at the telomeres and the centromeres. In addition, however, heterochromatin is also found at the silent mating type cassettes, which are a feature of yeasts in the Saccharomycotina and Taphinomycotina but are lacking in the filamentous ascomycetes. Although the actual mechanism of switching varies from yeast to yeast, there are nevertheless basic similarities in the switching process. Each cell carries an intact but silenced copy of information for each mating type (**a** or **α** in *Saccharomyces cerevisiae*, + or - for *Schizosaccharomyces pombe*). The mating type locus itself can carry information from either mating type. During meiosis, the mating type information at the mating type locus is “erased” and “rewritten” using information from the silenced copy of the opposite mating type, resulting in a switch of mating type from mother to daughter cell.

In *Saccharomyces cerevisiae*, only the telomeres and the silent mating cassettes flanking the mating type locus are heterochromatic; the centromere is a point

centromere and is not flanked by pericentric heterochromatin unlike the centromeres in *S. pombe* and *N. crassa* (Grunstein and Gasser, 2007). Heterochromatin in *S. cerevisiae* is characterized only by deacetylated histones; histone methylation, DNA methylation, and HP1 homologues are absent in the heterochromatin of this fungus (Hickman *et al.* 2011). Thus, within the fungi, and even within the Ascomycetes, there are varied ways of creating heterochromatin. Of note is that neither histone methylation nor DNA methylation is absolutely required.

DNA methylation is the best-studied epigenetic modification of the genomes of eukaryotes. Once thought to be strictly a mark of silent chromatin, DNA methylation is now also known to mark actively transcribed gene bodies in organisms as varied as monocots, green algae, and invertebrates, and to be present in genomes at levels ranging from negligible in *Drosophila melanogaster* and *Caenorhabditis elegans* to upwards of 90 percent in the green alga *Chlorella* (Zemach *et al.* 2010). It functions in diverse processes such as genomic imprinting, X-chromosome inactivation, silencing of transposable elements and other repetitive DNA in genome defense, expression or silencing of genes, and delineation of structural chromatin domains like telomeres and centromeres. Plants, animals and fungi all have DNA methylation, although the frequency of this mark in the genome and its function are not conserved (Zemach *et al.* 2010).

DNA methylation is generally less common among fungi than among plants and animals. An early survey of 20 diverse fungi found little to no DNA methylation

(Antequera *et al.* 1984). Most yeasts in the Saccharomycotina and Taphrinomycotina lack DNA methylation, as does *Aspergillus nidulans*. Among ascomycetes, DNA methylation is perhaps best characterized in the filamentous mold *Neurospora crassa*. In *Neurospora*, the phenomenon seems to be restricted to foreign DNA elements which are subject to a type of gene silencing called RIP (repeat-induced point mutation) (Miao *et al.* 2000). A genome-wide pull-down of methylated DNA sequences revealed that almost all were obvious transposable elements or remnants thereof (Selker *et al.* 2003). Of the ascomycetes studied so far, only *Ascobolus immersus*, *Candida albicans* and *Uncinocarpus reesii* seem to have DNA methylation in gene regions (Barry *et al.* 1993; Zemach *et al.* 2010; Mishra *et al.* 2011).

DNA methylation is also known from the basidiomycetes *Agaricus bisporus*, *Coprinopsis cinnerea*, *Laccaria bicolor*, *Postia placenta* and *Ustilago maydis* (Antequera *et al.* 1984; Zolan & Pukkila 1986; Zemach *et al.* 2010; Mishra *et al.* 2011). Genome-wide assessment of methylation levels in these diverse basidiomycetes demonstrated that most of these fungi have low levels of methylation localized at repetitive regions, comparable to most ascomycetes. The exception among fungi is the zygomycete *Phycomyces blakesleeanus*, in which nearly 10 percent of the genome was shown to be methylated (Zemach *et al.* 2010).

In contrast to the low levels of DNA methylation reported for most fungi, initial reports from the lichenizing fungus *Cladonia grayi* suggested that the levels of DNA

methylation and presumably of epigenetic silencing in this fungus were much greater, but only when the fungus was in symbiosis with the alga. When isolated in pure culture, the fungus also showed negligible levels of DNA methylation (Armaleo & Miao 1999b). Armaleo and Miao demonstrated this phenomenon by restricting DNA from various lichen tissues using restriction enzyme isoschizomers that are differentially sensitive to DNA methylation: both MspI and HpaII cut DNA at the sequence CCGG, but HpaII is blocked from cutting if the DNA is methylated, whereas MspI can cut regardless of DNA methylation at the site. Armaleo and Miao demonstrated differential cutting, indicating the presence of DNA methylation, in DNA isolated from the cleaned podetia (stalks supporting the apothecia) and in the soredia (asexual dispersal propagules composed of algal cells and fungal hyphae), and no differential cutting, indicating low or no DNA methylation, in the apothecia and in the fungal mycelia grown from spores. Taking advantage of the methylation sensitivity of the restriction enzyme EcoRI, they also performed a Southern blot of EcoRI/BamHI double digests of DNA isolated from the mycobiont and the cleaned podetia. They probed the blot with labeled portions of four polyketide synthase genes from the fungal partner of this lichen. Three of these probes hybridized to larger fragments in the digests from the lichenized tissue, while one showed the opposite pattern. The interference of digestion in the symbiotic state was interpreted as evidence of DNA methylation specifically in the symbiotic state of the fungus.

These results suggested that specifically when in the symbiosis, the fungus was silencing some large portion of its genome, possibly to facilitate the symbiosis by inactivating those genes or genome regions antagonistic to symbiotic life. Genome-wide DNA methylation thus seemed to correlate with the developmental switch from the free-living fungal state to a symbiotic state, suggesting that DNA methylation might be an important cause or consequence of lichenization.

Extrapolating from the work of Armaleo and Miao, which demonstrated high levels of DNA methylation in the symbiotic state and essentially no methylation in the aposymbiotic state of the lichen *Cladonia grayi*, I hypothesized that at each locus with DNA methylation, other components of the epigenetic silencing apparatus, i.e trimethylated H3K9 and heterochromatin binding protein HP1, would also be present. I further hypothesized that these silenced regions were perhaps critical in the initiation and/or maintenance of the symbiotic state, since the free-living fungus evidently did not require these regions to be silenced. Excitingly, this suggested a way to begin to untangle the signaling pathway controlling lichen development—by finding silenced regions, one could understand what the fungus shut down to enter into symbiotic life and could then track backward along the sequence of events by which a free-living fungus identifies a suitable photobiont and builds the symbiotic structure called a lichen thallus.

The three main goals of this work were 1) To reproduce the effect shown in the work of Armaleo and Miao using additional methods; 2) To identify the silencing *machinery* (that is, the genes encoding the proteins that silence portions of the genome and 3) To find the *targets* of silencing (that is, the genes or regions of the genome which are epigenetically silenced by the silencing machinery) in both the free-living state and in various structures of the symbiotic state to determine if epigenetic silencing and/or DNA methylation is correlated with symbiosis in the lichen *Cladonia grayi*.

3.1.1 Reproducing the Armaleo and Miao Effect

I attempted to replicate the effect shown in Armaleo and Miao using three additional techniques. I performed a southern blot using total labeled fungal DNA as a probe rather than individual genes as probes in order to assess the overall methylation of the fungal genome. I used an antibody to 5-methylcytosine to probe HpaII and MspI digests of DNA extracted from cleaned podetia in a southwestern blot to attempt to attribute the differential restriction to DNA methylation. I performed an immunoblot of DNA isolated from lichenizing fungi and aposymbiotically cultured mycobionts from diverse groups of lichens in order to determine how phylogenetically widespread DNA methylation in the symbiotic state might be.

3.1.2 Finding the silencing machinery

At the outset of this work, no genome project was in progress or in planning, so I used degenerate PCR to identify the genes encoding the H3K9-specific histone

methyltransferase and the heterochromatin-binding protein HP1. Armaleo was searching for the DNA methyltransferase. Because the *Cladonia grayi* system lacked important genetic tools and techniques such as gene deletion and transformation, these genes would have had to be expressed in another model fungus such as *Aspergillus nidulans* or *Neurospora crassa* in order to be studied. Later, once the genome was sequenced, I mined the genome for other genes encoding the proteins involved in epigenetic gene silencing.

3.1.3 Finding the silencing targets

In parallel, I attempted to identify regions that were silenced, using both methylation of H3K9 and methylation of DNA as proxies. I used ChIP to attempt to find regions with trimethylated H3K9. I used MeDIP to find regions with DNA methylation, and confirmed the methylation status of the candidate regions with bisulfite sequencing. I then used full genome bisulfite sequencing to identify the targets of DNA methylation in both the free-living fungus and the free-living alga and in the symbiotic structures of the lichen *Cladonia grayi* including the podetia, the squamules and the soredia.

3.2 Methods

3.2.1 Cultures

Lichens were collected from nature, washed briefly, and allowed to rehydrate in a very small volume of distilled water. After blotting to remove excess water, small

portions of thallus with apothecia were excised and affixed to the lid of an inverted Petri dish with tape. The spores were allowed to shoot up onto the agar (Ahmadjian 1966). Germlings were isolated and subcultured onto agar plates for growth or slants for storage. Media used were: Malt extract-yeast extract (MY) (Ahmadjian 1993), nutrient medium based on Bold's basal medium (NMBBM) (Trembley *et al.* 2002), oatmeal (10 g homogenized oatmeal flakes), potato-carrot (as in Simmons 1992 except made as a 10x stock, and centrifuged rather than sieved; Dyer and Crittenden, personal communication); and yeast extract plus supplements (YES) (MP catalog #4101-532). A total of 20 g/l agar was added to all solid media. For long-term storage, cultures were transferred to agar slants and maintained at room temperature or at 4°C. Vouchers of cultures were preserved as water stocks in which a slice of agar from a growing culture was fitted into a 2 ml tube completely filled with water and maintained at room temperature (Arnold *et al.* 2009). For algal cultures, a cleaned piece of thallus was crushed in distilled water in a 1.5 ml tube with a micropestle, and the slurry was diluted and plated onto water agar (16% agar), TMBBM or BBM 1NV (Bold's basal medium with nitrogen and vitamins, University of Texas algal culture facility). After 1 – 3 months, individual photobiont cells or colonies were subcultured. The identity of the fungal and algal cultures was confirmed by PCR amplification and Sanger sequencing (as described below) of the internal transcribed spacer (ITS) region of the ribosomal RNA tandem repeats.

3.2.2 Genomic DNA extractions

Genomic DNA from lichen thalli obtained from nature and from cultured fungal tissue obtained as described above was prepared in one of three ways. In all cases, purified DNA was quantified by Nanodrop or by Qubit.

3.2.2.1 Alkali lysis method of DNA extraction

DNA was extracted following the alkali lysis method outlined in Zolan and Pukkila (1986) as modified by Gueidan, Roux, and Lutzoni (2007). Each sample was ground to a powder, resuspended in 500 μ l of a 2% SDS extraction buffer (2% SDS, 0.15 M NaCl, 50 mM Tris pH 8, 10 mM EDTA), extracted once with phenol:chloroform:IAA (25:24:1), then centrifuged at maximum speed for 5 minutes. The aqueous phase was removed to a new tube and extracted once with 500 μ l of chloroform:IAA (24:1) and centrifuged at maximum speed for 5 minutes. The aqueous phase was moved to a new tube and the DNA precipitated with 300 μ l (0.6 volumes) of isopropanol and centrifuged as above. The resulting pellet was washed with 70% ethanol, then allowed to air dry before resuspension in 25 μ l sterile water.

3.2.2.2 CTAB method of DNA extraction

DNA was extracted using a CTAB method described in Cubero et al (1999). Tissue was homogenized under liquid nitrogen with a mortar and pestle, and the resulting powder was resuspended in 500 μ l pre-warmed extraction buffer (1% w/v CTAB; 1M NaCl; 100 mM Tris ; 20 mM EDTA; 1% w/v polyvinyl polypyrrolidone ,

PVPP). The mixture was heated at 70°C for 30 minutes, then extracted once with an equal volume of chloroform. The aqueous phase was added to two volumes of precipitation buffer (1% w/v CTAB; 50 mM Tris-HCl; 10 mM EDTA; 40 mM NaCl) and the mixture centrifuged for 15 minutes at 13,000 g. The pellet was resuspended in 350 µl of 1.2 M NaCl and incubated at 37°C for 30 minutes with 1 µl of 10 mM RNase. The mixture was extracted once with an equal volume of chloroform, and the aqueous phase recovered. Then 0.6 volumes of isopropyl alcohol were added and the mixture stored at 4°C for 15 minutes. The mixture was then centrifuged at 14,000 g for 20 minutes, the supernatant poured off and the pellet washed once with 70% EtOH. The resulting pellet was left to dry, then resuspended in TE (10 mM Tris pH 7.4, 1 mM EDTA) and quantified.

3.2.2.3 DTAB/CTAB method of DNA extraction

DNA was extracted using a DTAB/CTAB method as described in Armaleo and May (2009). Samples were ground to a powder under liquid nitrogen using a pre-chilled mortar and pestle. The resulting powder was weighed and resuspended in 20 volumes (w/v) of DTAB lysis buffer (40 mM Tris HCl pH 8, 20 mM NaOAc, 1 mM EDTa, 1%w/v SDS, with 1 µl RNase at 25 mg/ml added per 250 µl buffer). The sample was heated briefly at 65°C for 3 minutes, while mixing with a pipet to resuspend the powder. To this slurry was added 1/3 volume of 5M NaCl. After thorough mixing by inversion, the sample was centrifuged at 6,000 RPMs for 5 minutes. The supernatant was transferred

to a new tube and one volume of phenol/chloroform/IAA (25:24:1) added. After repeated inversion for 2 – 3 minutes, the mixture was centrifuged at 6,000 RPMs for three minutes. The upper phase was removed to a new tube and heated briefly at 65°C, and 0.11 volumes of pre-warmed CTAB buffer (10% CTAB, 0.7 M NaCl) was added and thoroughly mixed. The mixture was incubated at 65°C briefly. One volume of chloroform was added and the mixture was inverted for 2 minutes and then centrifuged for 3 minutes at 6,000 RPM. The aqueous phase was move to a new tube and 0.5 volumes of isopropanol was added. The sample was incubated at room temperature for five minutes, then centrifuged for five minutes at 6,000 RPMs. The supernatant was removed and the pellet was retained.

To further purify the DNA, the pellet was resuspended in lysis buffer at 5% of the initial volume and heated for 2 minutes at 65°C. At this stage, the pellet was broken up and resuspended as much as possible by gentle pipetting. The sample was transferred to an eppendorf tube and 1/3 volume of 5M NaCl was added. The sample was mixed and centrifuged at maximum speed for 5 minutes to pellet the undissolved portions of the pellet. The supernatant was transferred to a new tube and extracted once with an equal volume of chloroform. The mixture was .centrifuged for 5 minutes at maximum speed. The aqueous phase was transferred to a new tube and two volumes of ethanol were added to precipitate the DNA. The sample was centrifuged for one minute at maximum speed, and the supernatant removed. The pellet was dried in a speed vac

for ~10 minutes, then resuspended in 50 µl of distilled water by allowing the pellet to imbibe water at 65°C for 5 -10 minutes, then storing on ice in an ice bucket for at least one hour, or overnight.

After the pellet was fully hydrated, it was resuspended by gentle agitation with a pipet. The resuspension was then centrifuged for 10 minutes to remove trace insoluble material, and the supernatant containing the purified DNA moved to a new tube. The DNA was quantified by Qubit following the manufacturer's instructions.

3.2.3 DNA Amplification, Cloning and Sanger Sequencing

PCR using degenerate primers (Appendix E) was performed on a MJ Research PTC200 thermocycler or an Applied Biosystems (Foster City, CA) Veriti thermocycler using a series of two touchdown PCR programs. In the first program, an initial denaturation step for 3 minutes at 94°C was followed by 10 cycles of 30 seconds at 94°C, 30 seconds at 60° C with a -1°C step down at each cycle, and 60 seconds at 72°C, followed by 25 cycles of 30 seconds at 94°C, 30 seconds at between 50°C and 47°C, and 45 seconds at 72°C with a final elongation step at 72°C for 7 minutes. If no products were formed or if faint products were seen, a second program was implemented consisting of no initial denaturation step, 24 cycles of 30 seconds at 94°C, 30 seconds at 55°C with a -0.4° C step down at each cycle, and 60 seconds + 2 seconds per cycle at 72° C , followed by 12 cycles of 30 seconds at 94° C, 30 seconds at 45° C, and 120 seconds + 3 seconds per cycle at 72°C with a final elongation step at 72°C for 10 minutes.

Degenerate primers to the preSET and SET domains for the histone H3 lysine 9 methyltransferase (H3K9 HMT) and to the chromodomain and chromoshadow domain for heterochromatin binding protein 1 (HP1) were used to amplify fragments of the respective genes (Appendix E). Primers to the C5 methyltransferase domain and the bromo-adjacent homology domain (BAH domain) were used to attempt to amplify the Dim-2 DNA methyltransferase homologue (Appendix E). Successful PCR reactions were cloned and sequenced as above. Primers specific to the *Cladonia grayi* H3K9 histone methyltransferase and HP1 sequences were used in conjunction with the Genomewalker kit (Clontech) following the manufacturer's instructions to amplify and sequence successive portions of the genes and ultimately to identify the full-length genes.

PCR of genomic DNA using specific primers (Appendix E) was performed as above except using a program consisting of an initial denaturation step for 3 minutes at 94°C followed by 25 cycles of 30 seconds at 94°C, 30 seconds at 55° C and either 60 or 150 seconds at 72°C, followed a final elongation step at 72°C for 7 minutes.

PCR products were visualized on a TAE 1% agarose gel stained with SYBR Safe (Invitrogen, Carlsbad, CA). If necessary, faint products or product with multiple bands were cloned with the TOPO TA cloning kit (Invitrogen, Carlsbad, CA) following the manufacturer's instructions. For each cloning reaction, at least 8 clones were screened by colony PCR using T7 and M13R primers and a PCR program consisting of a 10

minute initial denaturation step, followed by 25 cycles of 30 seconds at 94°C, 30 seconds at 52°C and 60 seconds at 72°C followed by a final elongation step of 7 minutes. PCR products were cleaned with a Montage PCR filter column (Millipore, Bilerica, MA) or with an Exo-SAP clean-up using 1 µl SAP dilution buffer, 0.5 µl Exonuclease 1, 0.5 µl Shrimp Alkaline Phosphatase, added to 10 µl PCR reaction and incubating on one of the aforementioned thermocyclers for 30 minutes at 37°C, then 15 min at 80°C . Cleaned PCR products were sequenced in 10 µl reactions using: 1 µm primer, 3 µl purified PCR product, 0.5 µl Big Dye (Big Dye Terminator Cycle sequencing kit, ABI PRISM version 3.1; PE Applied Biosystems, Foster City, CA), 1.5 µl Big Dye buffer, and 4 µl double-distilled water. Automated reaction clean-up and visualization was performed at the Duke IGSP Genome Sequencing & Analysis Core Facility using Big Dye chemistry with an ABI 3730xl automated sequencer (PE Applied Biosystems, Foster City, CA). Sequencher version 4.8 (Gene Codes Corporation, Ann Arbor, MI) was used to edit sequences and assemble contigs.

3.2.4 Southern blotting

Genomic DNA was extracted from cultured *Cladonia grayi*, cultured alga *Asterochloris* sp. and from squamules and cleaned podetia of the *Cladonia grayi* lichen following the alkali lysis method. DNA was subjected to gel electrophoresis as described above. DNA was transferred overnight onto Amersham Hybond-N+ nitrocellulose membrane (GE Healthcare Life Sceinces) and Southern blotting was performed with the

Amersham Gene Images AlkPhos Direct Labeling and Detection System (GE Healthcare Life Sciences) following the manufacturer's instructions. Total DNA extracted from cultured *Cladonia grayi* was labeled using the CDP-Star Detection System according to the manufacturer's instructions and used to probe the blot. Kodak Biomax-MR film was used to capture the image.

3.2.5 DNA immunoblotting

Fresh lichens including *Parmotrema* sp., *Cladonia* sp., *Cladonia grayi*, *Usnea mutabilis*, *Stereocaulon tennesseense*, *Physcia* cf. *aipolia*, *Peltigera apthosa*, *Lasallia* sp., *Umbilicaria* sp., and *Dibaeis baeomyces* were dissected by removing the upper cortex and the algal layer in order to remove any signal from methylated algal DNA. Genomic DNA was then prepared following the alkali lysis method from the remaining fungal tissues, including the medulla and lower cortex. Genomic DNA from cultures of *Cladonia grayi*, *Xanthomendoza hasseana*, *Physcia* cf. *aipolia*, *Dibaeis baeomyces*, *Saccharomyces cerevisiae*, and the alga *Asterochloris* sp. was isolated directly following the alkali lysis method. Undigested genomic DNA was subjected to gel electrophoresis, then transferred overnight to a nitrocellulose membrane as for Southern blotting. The membrane was blotted with a mouse anti-5 methylcytosine antibody (Abcam) and a horseradish peroxidase conjugated anti-mouse IgG secondary antibody (Abcam) following the manufacturer's standard western blotting procedure.

In a related procedure, genomic DNA from cultures of the photobiont *Asterochloris* sp., from cultures of the fungus *Cladonia grayi*, and from squanules and cleaned podetia from the lichen *Cladonia grayi* was digested with HpaII, then subjected to gel electrophoresis, blotted and transferred as above.

3.2.6 Methylated DNA library construction

Purified genomic DNA was restricted with HpaII or with MspI (recognition site: CCGG) to completion as determined by parallel digests spiked with lambda DNA. The DNA was concentrated and desalted through a Microcon column (Millipore, Bilerica, MA). The 5' overhangs generated by the restriction were filled in with a reaction containing 1 µg digested DNA, 1 µl dNTPs (2mM each) 1 µl 10x PCR buffer and 0.5 µl Taq polymerase (Denville) in a 10 µl reaction which was incubated at 72°C for 10 minutes in one of the aforementioned thermocyclers. The fragments were cloned as described above and 20 clones were sequenced.

3.2.7 Methylated DNA Immunoprecipitation (MeDIP)

Methylated DNA immunoprecipitation (MeDIP) was performed as previously described (Weber *et al.* 2005). DNA was restricted with MseI (recognition site TTAA) overnight, then desalted and concentrated using the Qiaquick PCR Purification Kit according to the manufacturer's instructions. Adapters constructed from two single-stranded oligonucleotides (Vos *et al.* 1995) (Appendix E) were ligated to the purified fragments overnight at 8°C using T4 DNA ligase (New England BioLabs, Inc.) according

to the manufacturer's instructions. To check for successful ligation, PCR was performed using the touch-down protocol as described above except using a one-minute extension time and MeDIP primer 4 (Appendix E) as the only primer. The DNA was concentrated and desalted through a Microcon column (Millipore, Bilerica, MA) . The DNA was quantified by Nanodrop and examined for quality by gel electrophoresis as described above.

Four μg of DNA were diluted 1:100 in TE, denatured for 10 minutes at 100°C in one of the aforementioned thermocyclers. DNA was purified through a Microcon column as described above. To this DNA was added 100 μl 5x IP buffer (100 mM sodium phosphate pH 7.0, 1.4 M NaCl, 0.5 % Triton X-100), 1 μl of an antibody directed against 5-methylcytididine (Mab-5MECYT-100, Diagenode, Liège, Belgium). This primary antibody was conjugated to the DNA for two hours to overnight at 4°C with rotation.

While the conjugation reaction was proceeding, 40 μl of a secondary antibody conjugated to magnetic beads (Dynabeads M-280 Sheep anti-mouse IgG, Dynal Biotech #112.01) per sample were prewashed twice in 800 μl PBS + 0.1% bovine serum albumin for 5 minutes with shaking at room temperature. The washed beads were collected using a magnetic stand, and were resuspended in 40 μl 1x IP buffer. The washed Dynabeads were added to the sample, and the sample was incubated for 2 hours at 4°C with rotation.

Beads were collected in a magnetic tube rack and washed with 700 μ l 1x IP buffer for 10 minutes at room temperature with shaking. The beads were collected and the wash repeated twice. The beads were again collected, and resuspended in 250 μ l Proteinase K digestion buffer (50 mM Tris pH 8.0, 10 mM EDTA, 0.5 % SDS) to which 7 μ l 10 mg/ml proteinase K was added. The sample was incubated for 3 hours at 50°C with rotation. The sample was then extracted once with 250 μ l phenol. The aqueous phase was collected and then extracted with 250 μ l chloroform. The DNA was precipitated overnight in a freezer by adding 20 μ l of 5M NaCl, 1 μ l of glycogen, and 500 μ l of 100% EtOH. The resulting pellet containing single-stranded DNA putatively enriched for methylated regions was resuspended in 30 μ l TE.

To regenerate the second strand from this purified immunoprecipitated DNA, PCR was performed as described in Section 3.2.3, using the MeDIP4 PCR primer (Appendix E). The resulting fragments were cloned and sequenced as described in Section 3.2.3.

After this MeDIP procedure, the purified single stranded DNA was amplified by PCR using primers complementary to the adapters to create double-stranded DNA fragments for cloning and sequencing as described in Section 3.2.3. One hundred clones from a library made of squamule DNA and 67 clones of a library made of podetial DNA were sequenced by Sanger sequencing to determine the targets of methylation. Selected candidate genes were confirmed by bisulfite sequencing as described below. Universal

Methylated DNA Standard (Zymo Research) and the included control primers were used as a positive control for the bisulfite treatment and downstream steps.

3.2.8 Chromatin immunoprecipitation (ChIP)

To determine if methylation of a region of DNA may signify that the region is epigenetically silenced, chromatin immunoprecipitation for proteins known to co-localize with methylated DNA in silenced heterochromatin was performed as follows.

Fresh lichens are collected and fixed in formaldehyde to cross-link DNA-binding proteins to DNA. DNA was extracted and purified using the alkali lysis method described above. Adapters were ligated as for MeDIP as described above.

Immunoprecipitation with a commercial antibody to the histone H3 tri-methylated lysine 9 residue (H3K9me3) (Abcam) was performed following standard procedures (Rusche & Rine 2001). Enrichment for histone H3K9me3 among candidate regions was determined by quantitative PCR of the immunoprecipitated fraction compared to the input DNA. Quantitative PCR was performed as described in Armaleo and May (2009) on an MJ Research PTC-200 Peltier Thermal Cycler outfitted with a Chromo 4 real-time PCR detector. A qPCR master mix was prepared including all reagents except primers, which were added separately. Three 15- μ l replicates for each sample were placed into a 96-well PCR plate (Biorad) and the plate was sealed with microseal 'B' film (Biorad). The final composition of each reaction was: 5 μ l of genomic DNA (3 ng/ μ l); 1.5 μ l of Invitrogen 10 \times PCR buffer (200 mM Tris HCl, pH 8.4, 500 mM KCl); 5.62 μ l of water; 0.3

µl of Invitrogen 50× Rox Dye; 0.03 µl of a 100× dilution of Invitrogen Sybr Green I (50,000× stock); 0.9 µl of MgCl₂ (50 mM stock); 1.2 µl of a dNTP mix (each nucleotide at 1.25 mM); 0.3 µl of Taq DNA polymerase (Apex, 1 unit/µl); 0.075 µl of each primer (10 µM stock). Quantitative PCR was performed using the following cycling program: 10 min at 95°C followed by 40 cycles of 30 seconds at 95°C, 30 seconds at 60°C, 30 seconds at 72°C. The fluorescence signal was measured after each extension at 72°C. Ct values were obtained with the MJ OpticonMonitor Analysis software version 3.1.

3.2.9 Bisulfite sequencing of candidate methylated regions

To confirm the methylation status of the DNA of the candidate regions identified by MeDIP or by CHIP, bisulfite sequencing of the candidate regions was performed. Genomic DNA was extracted from podetial and squamule DNA from *Cladonia grayi* using the alkali lysis method. DNA was subjected to bisulfite treatment using the EZ DNA Methylation kit (Zymo Research) according to the manufacturer's instructions. To amplify the candidate regions from the bisulfite treated DNA, primers were designed using the MethPrimer program (Li & Dahiya 2002). The fragments amplified from the PCR reactions using the primers specific for the bisulfite-treated DNA (Appendix E) were cloned and sequenced as above to determine the methylation state of each cytosine in the sequence.

3.2.10 Bisulfite genome sequencing

Genomic DNA was extracted from the *Cladonia grayi* mycobiont *Cgr/DA1myc* (Armaleo & May 2009) growing in pure culture, from the photobiont *Asterochloris* sp. *Cgr/DA1ph* (Armaleo & May 2009) growing in pure culture, and from cleaned podetia, squamules and soredia of the lichen *Cladonia grayi* following the DTAB/CTAB method and further purified using the PowerClean kit (Mo Bio laboratories). DNA was submitted to the Duke IGSP Genome Sequencing & Analysis Core Facility for the initial steps of Illumina library preparation, and retrieved after methylated adapters designed by Illumina were attached. DNA was subjected to bisulfite treatment twice using the Epiect Whole Bisulfite Kit (Qiagen) and the resulting library size-separated on an agarose gel. A band corresponding to 200 – 500 basepairs was excised. Gel purification of the library was performed using a commercial gel purification kit (Qiagen), the cleaned library was amplified using a PCR reaction mix containing 2 µl bisulfite PCR primers (Illumina), 14.6 µl distilled water, 27 µl bisulfite-treated DNA, 5 µl 10x Pfu Turbo Cx Hotstart buffer, 0.4 µl dNTPs (25mM), and 1 µl PfuTurbo Cx Hotstart DNA Polymerase (Agilent Technologies). The reaction was amplified with an initial hold for 2 minutes at 95°C, then 30 seconds at 98°C, followed by 18 cycles of 10 seconds at 98°C, 30 seconds at 65°C, and 30 seconds at 72°C; a final extension of 5 min at 72°C; and an indefinite hold at 4°C (Ólafur S. Andrésón, personal communication). The amplified library was purified using a Qiagen PCR Purification kit (Qiagen) and eluted twice in 50

μL of EB buffer. The amplified library was quantified by Qubit, then returned to the sequencing facility for the final steps of Illumina sequencing. Data was assembled to the reference genomes of the fungus *Cladonia grayi* and the alga *Asterochloris* sp. at <http://genome.jgi-psf.org/Clagr2> and <http://genome.jgi-psf.org/Astpho1> using the short reads aligning software BSMAP (Xi and Li, 2009).

3.2.11 Genome mining

The genome sequences of *Cladonia grayi* and of *Asterochloris* sp. are hosted by JGI. Although not yet publicly available, the data are housed at <http://genome.jgi-psf.org/Clagr2> and <http://genome.jgi-psf.org/Astpho1>. To identify genes encoding the proteins involved in forming silent chromatin, the genomes were searched for the *C. grayi* homologues of a histone H3K9 methyltransferase and heterochromatin binding protein 1 (HP1) using portions of the genes identified by degenerate PCR to confirm the sequence of these genes obtained by use of the Genomewalker kit (Clontech). The sequences of additional genes involved in histone modification were extracted from the lichen genomes using the blastX algorithm employing, as query sequences, genes from *Neurospora crassa*, *Aspergillus nidulans*, and *Saccharomyces cerevisiae*.

3.2.12 Genome sequencing

DNA was extracted from the cultured mycobiont of *Acarospora strigosa* (Acarosporomycetidae), *Graphis scripta*, (Ostropomycetidae), *Arthonia cf rubrocincta* (Arthoniomycetes), and *Endocarpon pallidum* (Eurotiomycetes) and from the intact lichen

thallus of *Physcia cf stellaris*, *Leptogium* sp. (Lecanoromycetidae), *Dibaeis baeomyces* (Ostropomycetidae) and *Peltula cylindrica* (Lichinomycetes) using the CTAB method described above. Alternatively, DNA was extracted following the CTAB/DTAB method described above, and the resulting DNA was purified one additional time using the PowerClean kit (Mo Bio laboratories). One to six micrograms of DNA was submitted to Duke IGSP Genome Sequencing & Analysis Core Facility for sequencing using the Illumina HiSeq technology. Each genome was barcoded by the facility. One quarter lane was sequenced for each cultured fungus and one half lane sequenced for each lichen (~80 million to 600 million reads). Genomes were assembled using Velvet (Zerbino & Birney 2008) and various assembly parameters as discussed in Chapter 2.

3.3 Results

3.3.1 Evidence for increased fungal DNA methylation in the symbiotic state of *Cladonia grayi*

To explore the extent of DNA methylation in the genome of the fungus in the symbiotic state, DNA was extracted from podetia with soredia and apothecia removed, leaving largely fungal tissue representative of the fungus in the symbiotic state. Aliquots of the DNA were restricted to completion with either MspI or HpaII, separated by gel electrophoresis, then transferred to a membrane and probed with an antibody to 5-methylcytosine. The high background notwithstanding, the pattern of differential cutting seen by Armaleo and Miao is emphasized by the antibody (Figure 4). This DNA

immunoblot suggested that the differential cutting demonstrated by Armaleo and Miao could be attributable to DNA methylation

To dissect out what portion of the signal could be attributable specifically to methylation in the fungus rather than the alga given that it is not possible to remove all traces of algal DNA from the podetia even by thorough scraping, a southern blot was

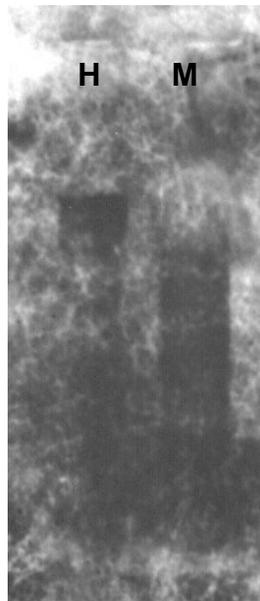


Figure 6: Immunoblot showing HpaII (H) and MspI (M) digest of DNA from *Cladonia grayi* podetia probed with an antibody to 5-methylcytosine.

performed using total fungal DNA from cultured *C. grayi* as a probe (Figure 5). DNA from the alga is not recognized by the probe, while fungal DNA is recognized.

Thus, the signal on the southern blot can be attributed to the fungal DNA. In the podetia sample, DNA restricted by HpaII yields more fragments of larger size than does the DNA restricted by MspI. Because MspI can restrict DNA regardless of methylation, and because restriction by HpaII is inhibited by methylation, this result suggests that at

least in the podetia, fungal DNA is methylated. In the squamule sample by contrast,

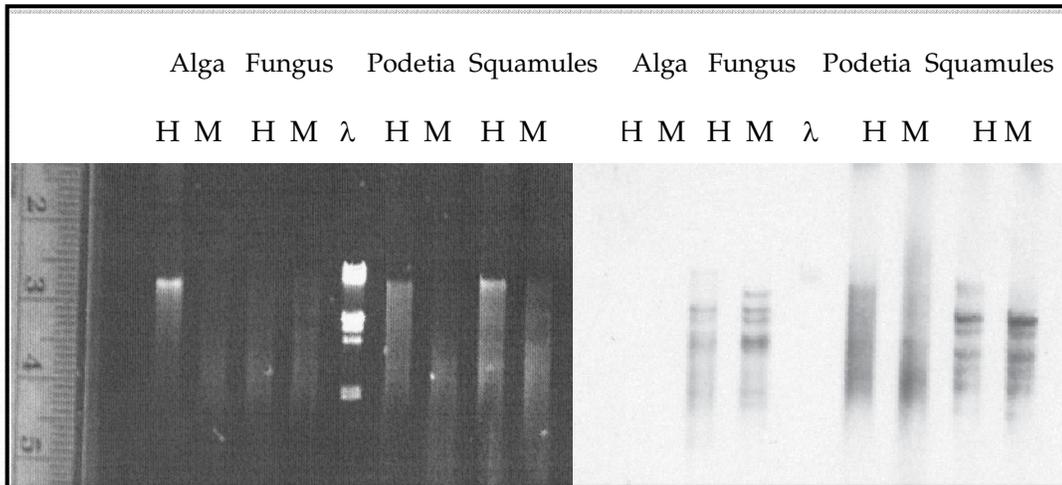


Figure 7: Southern blot probed with labeled total fungal DNA. Left panel: Agarose gel before transferring to nitrocellulose membrane for Southern blotting. Right panel: Southern blot probed with labeled total fungal DNA from cultured *Cladonia grayi*. H = HpaII, M=MspI, λ = lambda DNA (size standard).

there was only a small difference in the distribution of fragment sizes between the HpaII digest and the MspI digest, suggesting that fungal DNA methylation, although present, is less significant in this lichenized tissue. DNA from the cultured mycobiont showed no evidence of DNA methylation. The brighter signal toward the top of the MspI lane is probably due to a greater concentration of DNA in this lane. Finally, ethidium staining of the restriction digest (Left panel) shows that algal DNA appears strongly methylated both when the alga is free living and in symbiosis in the squamules. The lack of any signal in the free-living alga lanes suggests that all signal seen on the blot can be attributed to fungal DNA.

In a follow-up experiment to determine whether the putative DNA methylation in fungal DNA from lichenized structures was unique to *Cladonia* or was more widespread among lichens, several lichen thalli of lichens both closely and more distantly related to *Cladonia grayi* were assayed by DNA immunoblotting with an antibody to 5-methylcytosine. Lichens assayed included one other *Cladonia* species as well as lichens in closely related families (*Parmotrema michauxiana* and *Usnea mutabilis*, Parmeliaceae; *Stereocaulon tennesseense*, Stereocaulaceae); in closely related orders (*Physcia cf. aipola*, *Xanthomendoza hasseana*, Teloschistales) and more distantly related orders (*Peltigera aphthosa*, Peltigerales; *Umbilicaria* sp. and *Lasallia* sp., Umbilicariaceae), also in different subclass (*Dibaeis baeomyces*, Ostropomycetidae).

In lieu of performing a dot blot, in which different DNA extracts are spotted on a membrane that is then hybridized with the antibody, unrestricted genomic DNA was run out on a gel to allow any residual RNA to be separated out to ensure that the methylation signal was only from high molecular weight genomic DNA. The DNA from the gel was transferred to a membrane and assayed with the antibody to 5-methylcytosine (Figure 6). The results show a signal in most of the lanes in which DNA from dissected lichen thalli was run, indicating that DNA methylation in the symbiotic state may be present in *Parmotrema* sp., *Cladonia* sp., *Stereocaulon tennesseense*, *Physcia cf. aipolia*, and *Peltigera aphthosa*. The results are inconclusive for *Lasallia* sp. and *Umbilicaria* sp., as the DNA extraction was not very efficient and little DNA is evident. Surprisingly,

while there is clearly DNA in the lane for *Usnea mutabilis*, no DNA methylation is evident. This result could be indicative of no DNA methylation in the symbiotic state in this fungus, which would be surprising given its relatively close relationship to *Cladonia*. It could also be an experimental artifact of the preparation

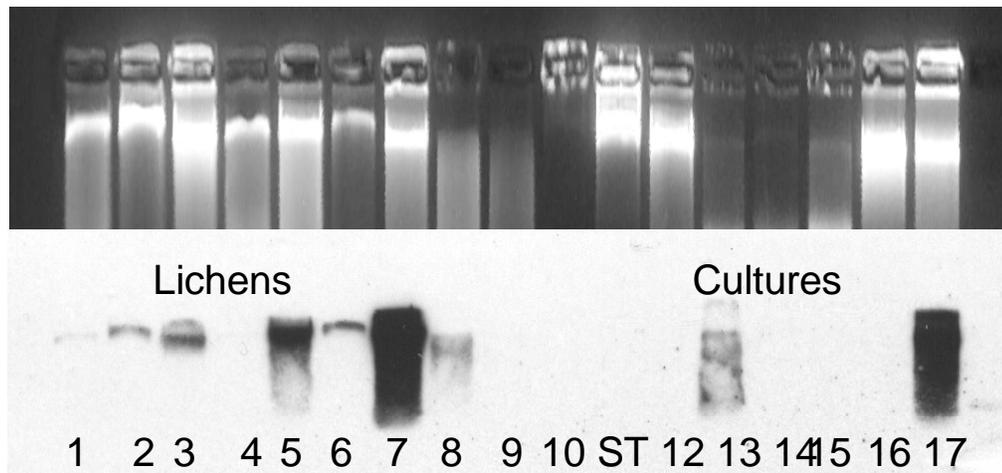


Figure 8: Genomic DNA immunoblotted with antibody to methylated cytosine. Top, gel image. Bottom, immunoblot. Lanes 1 - 10, lichens cleaned of algal cells. Lanes 12 - 17, cultures. 1. *Parmotrema* sp. 2. *Cladonia* sp.; 3. *Cladonia grayi*; 4. *Usnea mutabilis*; 5. *Stereocaulon tennesseense*; 6. *Phyiscia* cf. *aipolia* 7. *Peltigera apthosa*; 8. *Lasallia* sp.; 9. *Umbilicaria* sp.; 10. *Dibaeis baeomyces*; 11. Lambda DNA; 12. *Cladonia grayi* 13. *Xanthomendoxa hasseana*; 14. *Phyiscia* cf. *aipolia*; 15. *Dibaeis baeomyces*; 16. *Saccharomyces cerevisiae*; 17. Lichen alga *Asterochloris* sp.

technique, in which the outer cortex of this fruticose lichen is removed by a razor blade.

In practice, much of the medulla is also removed, leaving mostly the dense central core, a very atypical lichen tissue. Perhaps there is little or no DNA methylation in this portion of the thallus but there is DNA methylation in the medulla, which is in close contact with the alga. Or perhaps this sample is the only sample free of all algal cells.

An important caveat to these results is that these thalli are from nature and while they have been washed, it is impossible to remove all bacteria from the thallus. Bacteria have DNA methylation, although in a different sequence context than eukaryotes (i.e. 5' CC^{me} A/T GG' instead of 5' C^{me}G 3'). Part of the signal on the blot showing unrestricted DNA could be attributable to bacterial DNA methylation by the Dcm methyltransferase, as is almost certainly the case for the bright signal evident in the *Peltigera* lane, as *Peltigera* is symbiotic not with a green alga, as are all the other lichens shown, but with a cyanobacterium (*Nostoc*) and was not dissected. Additionally, as the thalli are not perfectly clean of algal cells or of DNA from destroyed algal cells, part or all of the signal on the blot could be attributable to methylation of the DNA from the alga.

Attempts were made to aposymbiotically culture the mycobiont from all the lichens shown in the blot, but not all isolations were successful, so matched pairs of lichenized and free-living fungi are not shown in all cases. For instance, *Lasallia* spores germinated and began to grow but failed to thrive on any of the five media used including both nutrient-rich and nutrient-poor media. Ascospores from *Umbilicaria* failed to germinate on any of the media tried, but conidia from one specimen did germinate, then failed to thrive in liquid media. *Peltigera* and other cyanolichens in the Peltigerales have only very infrequently been induced to germinate and grow (Stocker-Wörgötter, personal communication) and cannot routinely be germinated. Nevertheless, the trend shown from the available mycobionts and cultures appeared to be that DNA

methylation is present in the symbiotic state of the lichen but largely absent in the aposymbiotic state of lichenized fungi.

3.3.2 Identification of DNA methyltransferases from lichenizing fungi from diverse taxonomical groups

The *Neurospora crassa dim-2* gene and the *Cladonia grayi dmt1* genes (both encoding DNA methyltransferases), as well as the *masc1* gene from *Ascobolus*, the *rid-1* gene from *Neurospora crassa*, the *dmtA* gene from *Aspergillus nidulans* (Lee et al. 2008), and the homologue of these genes in the *Cladonia grayi* genome were used as query sequences to search the assemblies of lichen genomes for a DNA methyltransferase gene and for genes encoding other components of the heterochromatin silencing machinery. The genomes searched included those of the lichenizing fungi *Acarospora strigosa* (Acarosporomycetidae, Lecanoromycetes); *Arthonia cf rubrocincta* (Arthoniomycetes); *Graphis scripta* (Ostropomycetidae, Lecanoromycetes); *Endocarpon pusillum* (Chaetothyriomycetidae, Eurotiomycetes); *Dibaeis baeomyces* (Ostropomycetidae, Lecanoromycetes); *Leptogium* sp. (Lecanoromycetidae, Lecanoromycetes); *Physcia cf stellaris* (Lecanoromycetidae, Lecanoromycetes); and *Peltula cylindrica* (Lichinomycetes). Together with *Cladonia grayi*, these full and partial genomes span all major groups of lichens excluding the basidiomycete and dothideomycete lichens. Portions of the DNA methyltransferase gene were identified in seven lichens (Appendix H) and portions of the *masc1/rid-1/dmtA* gene were identified in all eight lichens (Appendix I). These

findings suggest some role for DNA methylation in the lichenizing fungi, because lichens from very different classes and subclasses have retained the DNA methyltransferase gene(s) even though other ascomycetes such as *Aspergillus nidulans* and most yeasts have lost these genes.

3.3.3 Identification of targets of DNA methylation in *Cladonia grayi*

Given that DNA methylation seemed to be specific to the symbiotic state in *Cladonia* and also in other lichens, it was natural to ask what sequences are subject to DNA methylation, and presumably to silencing, in the symbiotic state. To identify the DNA sequences which are methylated in the symbiotic state, various cloning strategies were attempted.

Initial attempts to create a plasmid library enriched for sequences subject to DNA methylation by cutting out the portion of a gel containing large HpaII fragments and then restricting and subcloning these fragments were unsuccessful. No DNA methylation could be demonstrated by bisulfite sequencing for any of the twenty test mycobiont sequences selected (Table 4). Therefore, a different strategy was adopted in order to minimize the cloning of unmethylated sequences.

A DNA methylation library was created using MeDIP (Methylated DNA Immunoprecipitation). MeDIP was performed on DNA extracted from cleaned podetia and on squamules in order to screen the various tissues for methylated sequences. Ninety-six clones from the squamule library were sequenced, yielding 66 unique

Table 7: List of candidate methylated regions from the first methylated library. Plasmid library was constructed from the high molecular weight fragments of a HpaII digest of *Cladonia grayi* podetial DNA. Regions selected for bisulfite sequencing are highlighted in bold type.

| 5' Gene or region | 3' Gene or region | Insert size (kb) |
|------------------------------|--|------------------|
| No homology | Integral membrane protein | 3 |
| No homology | No homology | 4 |
| Bacterial sugar transporter | Bacterial sugar transporter | 1 |
| MFS sugar transporter | No homology | 5 |
| No homology | P450 | 5 |
| Pre-mRNA splicing complex | Pre-mRNA splicing complex | 4 |
| MRD1(RNA processing) | Urea/AA transporter | 5 |
| No homology | P450 | 5 |
| Hypothetical protein | AreA | 2.5 |
| Hypothetical protein | Hypothetical protein | 0.5 |
| Dim1 | Dim1 | 2.5 |
| ERM | ERM | 0.5 |
| Endoribonuclease ATP binding | Sigma factor 54 | 6 |
| Disulfide isomerase | Disulfide isomerase | 2 |
| Bacterial sugar transporter | Bacterial sugar transporter | 1 |
| RPB2 | RNA binding (Loss of silencing) | 4.5 |
| Hypothetical protein | Ankryn repeat | 3.5 |
| No homology | Amino acid transporter | 5.5 |
| Hypothetical protein | NAD biosynthesis | 2.5 |
| Snf2 helicase/Sth1 | Snf2 helicase/Sth1 | 1.2 |

sequences of which 16 were fungal, 3 were algal and 47 were undetermined. From the podetial library, 67 clones were sequenced, yielding 30 unique sequences of which all were fungal. A list of candidate regions chosen for bisulfite sequencing is shown in Table 7.

Table 8: Selection of candidate methylated regions from MeDIP library of squamules and podetia. Algal genes in italics; fungal genes in regular type.

| Squamule library | Podetial library |
|-----------------------------------|-----------------------------------|
| Riboflavin biosynthesis | Conidiation-specific laccase |
| <i>Hypothetical protein</i> | GTPase |
| Sulfur transferase | Late embryogenesis domain protein |
| 2-nitropropane dioxygenase | Repetitive region 1 |
| mRNA transport regulator | Repetitive region 2 |
| DNA helicase | Hypothetical protein |
| Ribosome biogenesis | Region of no homology 1 |
| Polyketide synthase | Region of no homology 2 |
| <i>Forkhead-associated domain</i> | Region of no homology 3 |
| Hypothetical protein | |
| mRNA splicing | |
| Hypothetical protein | |
| <i>Hypothetical protein</i> | |

From the squamule library, ten fungal genes and three algal genes were chosen for bisulfite sequencing. Of the sequences tested, two were confirmed to be methylated by bisulfite sequencing. These two sequences were both algal in origin. From the podetial library, none of the nine sequences showed methylation by bisulfite sequencing. Therefore, this strategy was also deemed unsuccessful.

3.3.4 Whole genome bisulfite sequencing

Finally, bisulfite genome sequencing of total DNA from the fungus growing alone, from the alga growing alone, from cleaned podetia, from squamules and from soredia was bisulfite treated and prepared for Illumina sequencing, then subjected to two rounds of bisulfite treatment before the final steps of library preparation. The bisulfite conversion procedure performed twice successively as done in this work is

reported by the manufacturers to have a conversion rate of greater than 99.9 percent. However, even a 0.1% error rate over a 40 million basepair genome could lead to ten thousand unconverted cytosines. These errors are likely to be distributed randomly throughout the data. For this reason, only bases at which more than four reads reported cytosine were considered in this analysis. Appendix J shows the position of all the bases in the mycobiont genome for which more than four reads demonstrated methylation. Figures 8 and 9 in Appendix I summarize these results graphically.

DNA from the mycobiont grown in culture without the alga showed essentially no DNA methylation, as expected. The possible exception was a single base in a thioredoxin gene, for which 5 of 133 reads (0.88%) reported methylation (Fred Dietrich, personal communication). For most sites, the reads reporting methylation make up an even smaller percentage of all reads spanning the site. Additionally, cytosines neighboring a methylcytosine are almost always unmethylated (Appendix F). In the mycobiont, methylated cytosines are found in every possible sequence context with no preference for typically methylated contexts such as CG or CHG. This low level of methylation and the pattern of its occurrence are consistent with the manufacturer's reported false positive rate and other reported false positive rates (Cokus *et al.* 2008).

Contrary to expectations, the DNA from the squamules and the podetia also showed essentially no DNA methylation. Again, methylated cytosines are found in every possible sequence context. Likewise, the reads reporting methylation at any given

site make up only a small percentage of the total number of reads spanning the site. And in the majority of cases, neighboring cytosines are unmethylated (Appendix F; Figure 9).

However, in the symbiotic state eight regions demonstrated levels of methylation deemed not artifactual. These regions had more than one unconverted cytosine at neighboring or nearby positions, and also more than 10 reads indicating methylation for each putatively methylated site. Data on these regions are shown in Figures 8 and 9 in Appendix I and in Appendix J. Searching the sequence of these bisulfite reads against the *Cladonia* genome revealed that these methylated regions corresponded to ribosomal DNA repeats (Scaffold 000284) and to repetitive regions. Of all these regions, two were found in at least one lichen tissue but not in the other lichen tissue or in the negative control library from the mycobiont (Table 9 and Appendix I).

Repetitive regions such as the telomeric and/or centromeric repeats are often the targets of methylation in diverse fungi. However, telomeres and centromeres are difficult to assemble due to their repetitive nature and are probably missing from the *C. grayi* genome assembly. Nevertheless, these sequences should be present in the raw sequencing files of the bisulfite-treated lichen and mycobiont samples. Therefore, if there is methylation in these regions, it is possible to search the methylated reads from the bisulfite libraries of various lichen tissues directly against the library of bisulfite-

treated mycobiont DNA. Doing so did not reveal any additional methylated regions that could be attributed to the fungus beyond those already discussed.

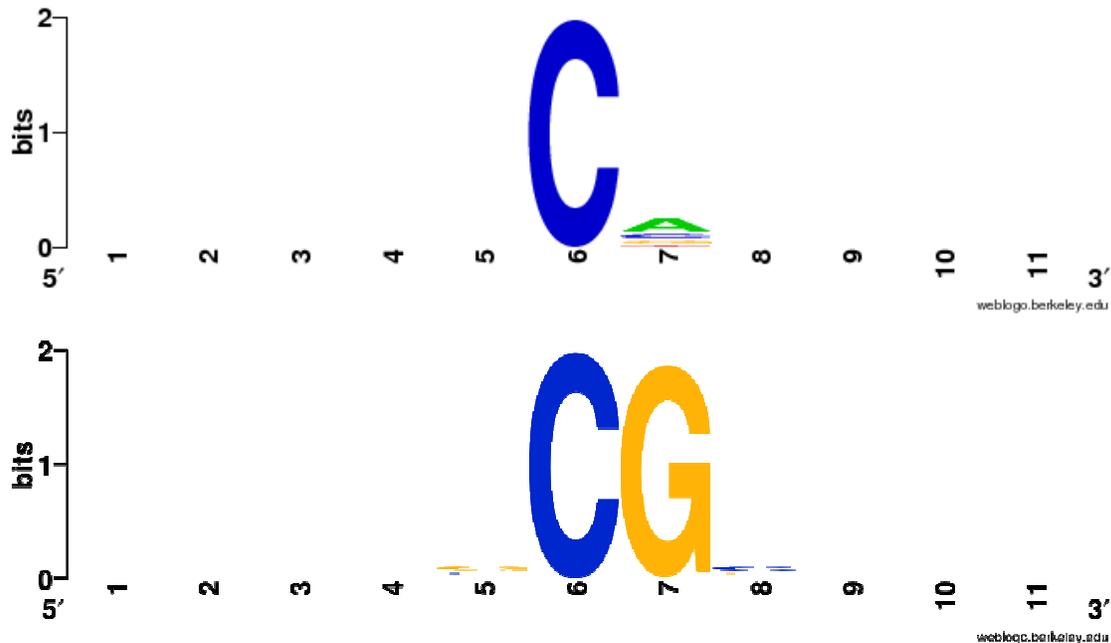


Figure 9. Consensus sequence of methylation sites in the genomes of the lichen *Cladonia grayi*. To generate each sequence logo, 10,000 11-basepair sequences each centered on a methylated cytosine were aligned using the online tool Weblogo (Crooks et al, 2004). **Upper panel:** Consensus methylation site in the fungus *Cladonia grayi* from mycobiont tissue. **Lower panel:** Consensus methylation site in the alga *Asterochloris sp.* from squamules.

While DNA methylation is very low in the symbiotically growing fungus and essentially absent in the aposymbiotically growing fungus, DNA methylation in the alga is much more extensive. DNA methylation in the alga is restricted to CG sites (Figure 7). Of the 1,760,151 CG sites in the *Asterochloris sp.* alga genome assembly hosted at JGI, 1,373,991 CG sites were shown to have methylation in the squamule library (78%) and

454, 402 CG sites were shown to have methylation in the podetial library (26%). The podetia are largely fungal in origin, so the algal reads in the podetial library likely represent alga from the soredia remaining on the podetia at the time of DNA extraction even after the stringent cleaning. As a result, coverage of the algal genome from this library is very low (~4x). Thus, it is likely that many sites in the genome were covered by three or fewer reads, meaning that even if 100% of reads at a site reported methylation, these sites would have failed to reach the minimum threshold and would have been excluded from the analysis. The proportion of methylated sites in the podetial library is therefore likely a gross underestimate of methylation in the soredia.

In general, CG sites in the DNA of the alga are methylated in exons, introns and intergenic regions. An analysis of the methylation status of CG sites in the aposymbiotic alga shows that these regions are methylated at nearly 100 percent. There do, however, appear to be regions of DNA largely devoid of methylation at CG sites. These regions tend to be short, on the order of 100 – 200 basepairs, and correspond to the promoter region and first exon of genes. A plot of methylated CG sites versus unmethylated CG sites is shown for the nuclear-encoded small subunit of RUBISCO (Figure 10). Similar results were obtained for plots of a flagellar-associated protein and GTP-binding protein LepA (data not shown).

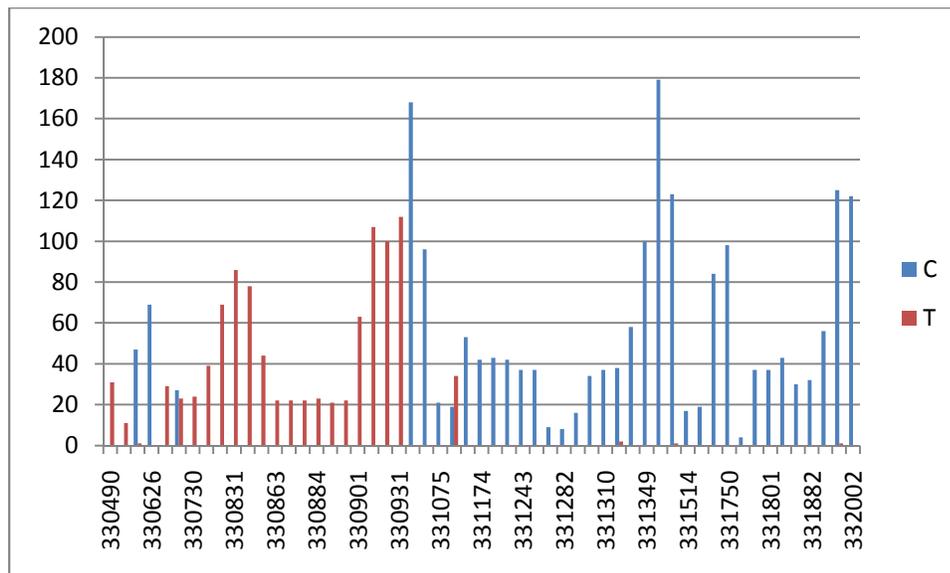


Figure 10: Plot of bisulfite-transformed cytosine and methylcytosine on a representative algal gene from the aposymbiotic alga. Unmethylated cytosines are transformed by bisulfite treatment into thymine, whereas methylcytosines are not. Translational start site is at 330490, translational stop site is at 331930. Cytosines are methylated in most regions except near the start site.

In addition, scattered unmethylated CG sites can also be found in the DNA from the aposymbiotic and symbiotic alga. These sites tend to be single sites separated from others by hundreds of basepairs. An analysis was performed of sites differentially methylated between the alga growing alone and the symbiotic alga from the podetia. The analysis was performed on sites which were unmethylated in the symbiotic state and methylated in the aposymbiotic state, and vice versa.

In all, 9869 sites were found to have five or more methylated reads in the symbiotic alga from the podetia and to contain more unmethylated reads at the site of interest in the free-living alga. However, by virtue of the higher coverage of the algal genome, this analysis generated many false positive results which did indeed have more

unmethylated reads than did the podetia, but also had an even greater number of methylated sites. A second filter requiring at least three unmethylated sites within 100 basepairs and a greater proportion of methylated sites in the podetia than in the alga narrowed the analysis down considerably. Sixteen representative regions were chosen for further study. Each consisted of 20 – 1500 basepairs with 5000 basepairs on either side for context. The approximately 10 kb regions were blasted against the data in Genbank to determine if any pattern could be found in the differentially methylated regions. Of the 16 regions, one was intergenic (between converging terminators), two were in introns, one was in an exon, one was in an intergenic region between divergent promoters, five were in intergenic regions, and six were in regions of low homology such that it was not possible to determine if any features were present or not. Interestingly, two of the sites were within 3 kb of a transposon-like sequence.

Fewer sites were found to have five or more unmethylated sites in the podetia and more methylated sites in the algal. Twenty representative regions were chosen. Each consisted of 20 – 100 basepairs with 5 kb on either side for context. Of the 20 regions, one was intergenic (between converging terminators), five were in introns, five were in exons, and nine were in regions of low homology such that it was not possible to determine if any features were present or not. Most differentially methylated regions were restricted to a single CG separated from others by hundreds or thousands of basepairs. The coverage over these sites is generally good, suggesting that they were not

artificial. It remains to be seen if these sites correlate with, for instance, transcriptional start sites, splice sites or other interesting features.

3.4 Discussion

The strongest piece of evidence supporting the idea that methylation of fungal DNA is specific to the symbiotic state in lichens generally and in *C. grayi* specifically was demonstrated by Armaleo and Miao, who showed differential cutting with restriction enzyme EcoRI (known to be sensitive to DNA methylation) of DNA isolated from the symbiotic structures versus DNA of fungal cultures. Three polyketide synthase genes were shown by Southern blotting to be methylated at one EcoRI site. They also demonstrated differential restriction of fungal DNA from various lichen structures with HpaII (cut site CCGG), similar to results presented here. Thus, it is surprising and perplexing that the various cloning and sequencing strategies undertaken here failed to uncover the predicted massive amounts of DNA methylation in the symbiotic state.

These seemingly contradictory results could be reconciled by invoking a modification other than C5 methylation of cytosine that would block the cutting of HpaII but not MspI and would not be modified by bisulfite treatment or recognized by an antibody to C5 methylation. For example, methylation of the nitrogen on cytosine (N4) blocks cutting by HpaII but not MspI and would not be recognized by an antibody to C5 methylation of cytosine and would not be subject to bisulfite treatment. N4

methylation of cytosine is known from bacteria. The modification is performed by enzymes such as PvuII from *Proteus vulgaris* and M.SmaI from *Serratia marcesens* (Butkus *et al.* 1987; Adams & Blumenthal 1997). These proteins are part of the N4 and A5 methyltransferase family. A search of Genbank including 130 fungal genomes shows that only bacteria have homologues of this gene. The genome of *C. grayi* lacks a homologue in this gene family. Likewise, the genomes of the eight other lichenizing fungi also lack obvious homologues in this gene family.

Hydroxymethylation of C5 could also interfere with the action of HpaII but not MspI. However, the Tet-domain proteins which make this modification are currently only known from metazoa (Iyer *et al.*, 2010). Related enzymes are found in green algae, basidiomycetes, and a few heteroloboseans and stramenopiles, but not in ascomycetes (Iyer *et al.*, 2010). Additionally, bisulfite sequencing cannot distinguish between hydroxymethylation and methylation of C5 (Huang *et al.*, 2009), meaning that were this modification responsible for the differential cutting by HpaII and MspI, bisulfite sequencing should have detected it and no discrepancy between the restriction experiments and bisulfite sequencing would have been apparent. Thus, for several reasons hydroxymethylation of C5 is unlikely to be the source of the discrepancy detected here.

It remains possible that DNA methylation is critical not in the maintenance but in the establishment of the lichen symbiosis. In this scenario, DNA methylation of the

genome would be a transient phenomenon during one or more stages of development. Were this to be the case, it is unclear how the phenomenon could be effectively studied, as the resynthesis of the mature lichen symbiosis has only sporadically been achieved under laboratory conditions and for only a few taxa such as *Cladonia cristatella* and *Endocarpon pusillum* (Ahmadjian 1966; Ahmadjian & Heikkilä 1970). Using fluorescent in situ hybridization (FISH) on developing thalli harvested from the wild could be a possibility, although attempts by lichen biochemists to get FISH working on intracellular targets in lichen thalli have been unsuccessful (R. Honegger, personal communication). Most importantly, this scenario could not account for the discrepancy between this work and that of Armaleo and Miao, both of which examined mature thalli.

Other possible explanations for the noted discrepancy could be that DNA methylation in lichens is sensitive to environmental conditions and is unusually labile. For instance, DNA methylation could be seasonal, such that DNA is methylated (i.e. protected) in the winter and unmethylated during the active growing season. Or, DNA could be differentially methylated when desiccated or hydrated. In this scenario, the discrepancy could be accounted for by the researchers harvesting the lichens at different times of the year, different times of day or in different weather. One weakness with this scenario is that DNA methylation is not easily reversed. To date, no fungal DNA demethylase is known, although enzymes which can modify 5-methylcytosine to 5-hydroxymethylcytosine are known in basidiomycetes, as previously mentioned (Iyer *et al.*

2011). Without enzymes to actively demethylate DNA, the only way to lose DNA methylation is passively, through the failure to transfer the DNA methylation mark onto new daughter strands of DNA during DNA synthesis. However, lichens are not known for fast cell replication to say the least, so passive loss of DNA methylation would not be fast enough to respond to changing weather in the span of a day. Another weakness of this scenario is that once in the lab, the lichens were treated essentially the same by the two groups (Armaleo and Miao, and McDonald and Armaleo), that is, first desiccated, then submerged in water and vortexed to remove soredia, then desiccated again.

Before embarking on any extension of this project, it could be prudent for future lichen epigenetics researchers to examine the expression of the DNA methyltransferase. Homologues of DNA methyltransferases were found in seven out of eight lichen genomes. No DNA methyltransferase was found in the genome of *Graphis scripta*, so it cannot be said that DNA methylation is strictly required for each and every lichen symbiosis. However, the prevalence of this gene in fungi of diverse lichenized lineages gives modest support to the idea that DNA methylation may play some role in the life of the lichenizing fungus, although what role that is remains unclear. When and where the DNA methyltransferase is expressed could give a clue to its function in the fungus and perhaps in the symbiosis.

It is important to note that epigenetic gene silencing need not involve DNA methylation at all. *S. cerevisiae*, *S. pombe* and *A. nidulans* all have silent chromatin even

though they lack DNA methyltransferases. In fact, DNA methylation is no longer believed to be strictly a mark of silent chromatin. Two recent papers demonstrate that although the last common ancestor of eukaryotes likely had a DNA methyltransferase, neither the methylation machinery nor the function of DNA methylation has been preserved throughout the evolution of eukaryotes (Feng *et al.* 2010; Zemach *et al.* 2010). Specifically, while DNA methylation is generally a mark of silencing in animals and fungi, in plants it is also found in actively transcribed regions of the genomes. This was found to be true of the green algae *Chlorella* sp. and *Volvox carteri*, as well as of *Oryza sativa*, but not the moss *Physcomitrella patens* or the spikemoss *Selaginella moellendorffii* (Zemach *et al.* 2010). Increased DNA methylation in actively transcribed gene regions was also found in *Arabidopsis thaliana*, *Populus trichocarpa*, and *Chlamydomonas reinhardtii* (Feng *et al.* 2010). Similar results for the alga *Asterochloris* sp. were obtained in this work. Of course, in fungi what little methylation there is does seem to be associated with silent regions. Most notably, in the yeast *Candida albicans*, the modicum of DNA methylation there is (<0.5%), was found in the bodies of genes, particularly those genes involved in phenotypic switches, i.e. from white to opaque cells or from yeast to hyphal growth forms (Mishra *et al.* 2011). Nevertheless, the data from plants argues that care should be taken when making a link between gene silencing or transcriptional repression and DNA methylation.

Methylation of H3K9 and the subsequent recruitment and deposition of HP1 homologues remain the strongest and most consistent mark of silent chromatin across species. While *S. cerevisiae* lacks H3K9 methylation, both *S. pombe* and *N. crassa* have it. Likewise, *C. grayi* and all seven lichen genomes have homologues of the H3K9-specific histone methyltransferase. It is therefore possible that histone methylation, particularly H3K9 trimethyl or dimethyl, could be solely responsible for marking silent chromatin domains in *C. grayi*. However, initial unsuccessful ChIP for H3K9 trimethyl from *C. grayi* casts doubt on this idea, suggesting that H3K9 dimethylation or no H3K9 mark at all is characteristic of silent chromatin. In the absence of new and compelling evidence to reopen one of these lines of inquiry, the current dataset would suggest that gene silencing is not important in any lichen tissue on the fungal side.

Examining DNA methylation in the alga, however, may prove more interesting and more fruitful. Of the methylated reads from the squamule and podetia bisulfite genome sequencing data, all with more than one unconverted cytosine in a CG context matched the algal genome with high homology. Alignment of these reads to the algal genome is underway.

Bisulfite treatment of DNA from the lichen soredia and from the alga growing alone in culture has been performed and results are forthcoming. When these data become available, it will be clear if there are any regions in the genome of the alga that are differentially methylated in the symbiosis. With this data in hand, it will be possible

to definitively determine if the lichenizing fungus *C. grayi* uses epigenetic silencing to facilitate the interaction with its symbiont in the lichen symbiosis.

Appendix A: Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences generated for or used in this study

Table 9: Accession numbers of ammonium transporter/ammonia permease genes of sequences generated for this study and included in the phylogenetic tree

| Organism, gene abbreviation | GenBank accession number | Primers (genome) | Taxonomy | Source (fungal culture or specimen number) | Clade |
|---|--------------------------------|----------------------------------|---------------------------------|--|--------------|
| Green algae | | | | | |
| <i>Asterochloris</i> sp., <i>amt1.1</i> | JF833069 | Genome | Trebouxiophyceae/Trebouxiaceae | <i>Cgr/DA1pho</i> | AMT-Euk |
| <i>Asterochloris</i> sp., <i>amt1.2</i> | JF833067 | Genome | Trebouxiophyceae/Trebouxiaceae | <i>Cgr/DA1pho</i> | AMT-Euk |
| <i>Asterochloris</i> sp., <i>amt1.3</i> | JF833068 | Genome | Trebouxiophyceae/Trebouxiaceae | <i>Cgr/DA1pho</i> | AMT-Euk |
| Fungi | | | | | |
| <i>Cladonia grayi</i> , <i>mep1a</i> | JF833071 | Genome | Lecanoromycetes/Cladoniaceae | <i>Cgr/DA2myc/ss</i> | MEP α |
| <i>Cladonia grayi</i> , <i>mep1b</i> | JF833070 | Genome | Lecanoromycetes/Cladoniaceae | <i>Cgr/DA2myc/ss</i> | MEP α |
| <i>Cladonia grayi</i> , <i>mep2</i> | JF833072 | Genome | Lecanoromycetes/Cladoniaceae | <i>Cgr/DA2myc/ss</i> | MEP γ |
| <i>Cladonia grayi</i> , <i>mep3</i> | JF833073 | Genome | Lecanoromycetes/Cladoniaceae | <i>Cgr/DA2myc/ss</i> | MEP γ |
| <i>Ramalina</i> sp., A | JF833078 | 1, 7, 10, 14 | Lecanoromycetes/Ramalinaceae | Rasp_TRM1470_PS1 | MEP α |
| <i>Ramalina</i> sp., B | JF833079 | 1, 7, 10, 14 | Lecanoromycetes/Ramalinaceae | Rasp_TRM1470_PS1 | MEP α |
| <i>Laurera megasperma</i> , A | JF833075 | 3, 4, 5, 6, 7, 10, 11, 12, 13 | Dothidiomycetes/Trypetheliaceae | Lame_DE9725_TRMPS1 | MEP α |
| <i>Laurera megasperma</i> , B | JF833074 | 3, 4, 5, 6, 7, 10, 11, 12, | Dothidiomycetes/Trypetheliaceae | Lame_DE9725_TRMPS1 | MEP α |

Table 9 (continued): Accession numbers of ammonium transporter/ammonia permease genes of sequences generated for this study and included in the phylogenetic tree

| Organism, gene abbreviation | GenBank accession number | Primers (genome) | Taxonomy | Source (fungal culture or specimen number) | Clade |
|--|---|----------------------------------|---------------------------------|---|--------------|
| | | 13 | | | |
| <i>Usnea mutabilis</i> , A | JF833082 | 1, 9, 10 | Lecanoromycetes/Parmeliaceae | McDonald 1468 | MEP α |
| <i>Usnea mutabilis</i> , B | JF833083 | 1, 9, 10 | Lecanoromycetes/Parmeliaceae | McDonald 1468 | MEP α |
| <i>Parmelia michauxianum</i> | JF833076 | 1, 3, 7, 9, 10, 11, 14 | Lecanoromycetes/Parmeliaceae | McDonald 1472 | MEP α |
| <i>Pyrenula cruenta</i> | JF833077 | 3, 4, 5, 6, 7, 10, 11, 12, 13 | Eurotiomycetes/Pyrenulaceae | Pycr_EGB_TRMPS1 | MEP α |
| <i>Trypethelium virens</i> | JF833081 | 3, 4, 5, 6, 7, 10 | Dothidiomycetes/Trypetheliaceae | Lame_DE9725_TRMPS1 | MEP α |
| <i>Stereocaulon tennesseense</i> | JF833080 | 1, 7, 10, 11, 14 | Lecanoromycetes/Cladoniaceae | McDonald 1471 | MEP α |

Table 10: Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|--|---|----------------------|--------------------|--------------|
| EUKARYOTES | | | | |
| Green plants (Viridiplantae; green algae and land plants) | | | | |
| <i>Arabidopsis thaliana</i> , AtAMT1;1 | GI:240256243 | 7858183..7859918 | GI:30682507 | AMT-Euk |
| <i>Arabidopsis thaliana</i> , AtAMT1;2 | GI:240254678 | 16039475..16042383 | GI:30697099 | AMT-Euk |
| <i>Arabidopsis thaliana</i> , AtAMT1;3 | GI:240254678 | (8805631..8807388) | GI:240255695 | AMT-Euk |
| <i>Arabidopsis thaliana</i> , AtAMT1;4 | GI:67633763 | 14161681..14163195 | GI:18417246 | AMT-Euk |
| <i>Arabidopsis thaliana</i> , AtAMT1;5 | N/A | N/A | GI:18404127 | AMT-Euk |
| ¹¹⁷ <i>Arabidopsis thaliana</i> , AtAMT2 | GI:240254678 | (16039475..16042383) | GI:42569729 | MEP α |
| <i>Brassica napus</i> , BnAMT1;2 | N/A | N/A | GI:11066959 | AMT-Euk |
| <i>Chlamydomonas reinhardtii</i> , AMT1A | GI:159478960 | (233330..238153) | GI:159478830 | AMT-Euk |
| <i>Chlamydomonas reinhardtii</i> , AMT1B | GI:159469368 | 851249..855814 | GI:159469030 | AMT-Euk |
| <i>Chlamydomonas reinhardtii</i> , AMT1C | GI:159487175 | 327409..336562 | GI:159487103 | AMT-Euk |
| <i>Chlamydomonas reinhardtii</i> , AMT1D | GI:159471504 | 1197959..1202240 | GI:159470638 | AMT-Euk |
| <i>Chlamydomonas reinhardtii</i> , AMT1E | GI:159474331 | (327551..334514) | GI:159474179 | AMT-Euk |
| | | | GI:159474181 | |
| <i>Chlamydomonas reinhardtii</i> , AMT1F* | GI:159488156 | (309559..315097) | GI:159488146 | N/A |
| <i>Chlamydomonas reinhardtii</i> , AMT1G | GI:159484046 | 797825..801801 | GI:159483300 | AMT-Euk |
| | | | GI:159483298 | |
| <i>Chlamydomonas reinhardtii</i> , AMT1H | GI:159469871 | 1189305..1193334 | GI:159469543 | AMT-Euk |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|--|---|-----------------|---------------------------------------|--------------|
| <i>Chlorella</i> sp. NC64A, 1.1 ^I | N/A | N/A | 1 56592 estExt_fgenes3_pg.C_10342JGI | AMT-Euk |
| <i>Chlorella</i> sp. NC64A, 1.2 ^I | N/A | N/A | 1 136742 IGS.gm_17_00146 | AMT-Euk |
| <i>Chlorella</i> sp. NC64A, 1.3 ^I | N/A | N/A | 1 36096 estExt_Genewise1Plus.C_140283 | AMT-Euk |
| <i>Chlorella</i> sp. NC64A, 1.4 ^I | N/A | N/A | 1 141357 IGS.gm_33_00067chlorella | AMT-Euk |
| <i>Chlorella</i> sp. NC64A, 1.5 ^I | N/A | N/A | 1 58614 estExt_fgenes3_pg.C_180083 | AMT-Euk |
| <i>Coccomyxa</i> sp. C-169, 1.1 ^I | scaffold 8 | 1173108-1176954 | N/A | AMT-Euk |
| <i>Coccomyxa</i> sp. C-169, 1.2 ^J | scaffold 1 | 2847728-2852827 | N/A | AMT-Euk |
| <i>Cryptomeria japonica</i> | N/A | N/A | GI:114841288 | MEP α |
| <i>Lotus corniculata</i> , LcAMT1;1 | N/A | N/A | (Huang and Peng, 2005) | AMT-Euk |
| <i>Lotus corniculata</i> , LcAMT1;2 | N/A | N/A | (Huang and Peng, 2005) | AMT-Euk |
| <i>Lotus japonicus</i> , LjAMT1;1 | N/A | N/A | GI:10952509 | AMT-Euk |
| <i>Lotus japonicus</i> , LjAMT1;2 | N/A | N/A | GI:31322043 | AMT-Euk |
| <i>Lotus japonicus</i> , LiAMT1;3 | N/A | N/A | GI:46409003 | AMT-Euk |
| <i>Lotus japonicus</i> , LiAMT2 | N/A | N/A | GI:15799271 | MEP α |
| <i>Medicago truncatula</i> , MtAMT2 | N/A | N/A | GI:22900885 | MEP α |
| <i>Micromonas</i> sp. RCC299, 1 | GI:255084674 | 237467..239735 | GI:255083670 | AMT-Euk |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|----------------------|--------------------|--------------|
| <i>Micromonas</i> sp. RCC299, 2 | GI:255081797 | 579223..580970 | GI:255079969 | MEP β |
| <i>Micromonas</i> sp. RCC299, * | GI:255081791 | (518206..520125) | GI:255070868 | N/A |
| <i>Oryza sativa</i> , OsAMT1;1% | GI:15705367 | N/A | GI:115459359 | AMT-Euk |
| <i>Oryza sativa</i> , OsAMT1;2 | GI:15705369 | N/A | GI:115447350 | AMT-Euk |
| <i>Oryza sativa</i> , OsAMT1;3 | GI:15705371 | N/A | GI:115447352 | AMT-Euk |
| <i>Oryza sativa</i> , OsAMT2;1 | GI:28875520 | N/A | GI:19909967 | MEP α |
| <i>Oryza sativa</i> , OsAMT2;2 | GI:115442598 | (35590381..35592277) | GI:28875522 | MEP α |
| <i>Oryza sativa</i> , OsAMT2;3 | GI:20160632 | 109596..111823 | GI:20160632 | MEP α |
| <i>Oryza sativa</i> , OsAMT3;1 | GI:115442598 | 37720693..37731081 | N/A | MEP α |
| <i>Oryza sativa</i> , OsAMT3;2 | GI:28269486 | 77711..81059 | GI:32985620 | MEP α |
| <i>Oryza sativa</i> , OsAMT3;3 | GI:46390229 | 104575..107816 | N/A | MEP α |
| <i>Ostreococcus lucimarinus</i> CCE9901, 1.1 | GI:145326699 | (76446..77958) | GI:145346880 | AMT-Euk |
| <i>Ostreococcus lucimarinus</i> CCE9901, 1.2 | GI:145327252 | (549105..550753) | GI:145348151 | AMT-Euk |
| <i>Ostreococcus lucimarinus</i> CCE9901, 3 | GI:145327766 | (334523..336339) | GI:145348948 | MEP β |
| <i>Ostreococcus lucimarinus</i> CCE9901, 4 | GI:145335532 | 34720..37072 | GI:145356297 | MEP β |
| <i>Physcomitrella patens</i> subsp. patens, 1.1 | GI:168068395 | 17676..19852 | GI:168068393 | AMT-Euk |
| <i>Physcomitrella patens</i> subsp. patens, 1.2 | GI:168067178 | (142647..144721) | GI:168067174 | AMT-Euk |
| <i>Physcomitrella patens</i> subsp. patens, 1.3 | GI:168012056 | 2305674..2307998 | GI:168011824 | AMT-Euk |
| <i>Physcomitrella patens</i> subsp. patens, 1.4 | GI:168012056 | 2284542..2286380 | GI:168011816 | AMT-Euk |
| <i>Physcomitrella patens</i> subsp. patens, 1.5 | GI:168012056 | 2290593..2293134 | GI:168011818 | AMT-Euk |
| <i>Physcomitrella patens</i> subsp. patens, 1.6 | GI:168012056 | (114438..118446) | GI:168011840 | AMT-Euk |
| <i>Physcomitrella patens</i> subsp. patens, 2.1 | GI:168004612 | 1799359..1802286 | GI:168004232 | MEP α |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|----------------------|--------------------|--------------|
| <i>Physcomitrella patens</i> subsp. <i>patens</i> , 2.2 | GI:168058452 | (46510..48246) | GI:168058414 | MEP α |
| <i>Physcomitrella patens</i> subsp. <i>patens</i> , 2.3 | GI:168035123 | (956773..959548) | GI:168052163 | MEP α |
| <i>Physcomitrella patens</i> subsp. <i>patens</i> , 2.4 | GI:168052203 | (228592..232717) | GI:168035109 | MEP α |
| <i>Physcomitrella patens</i> subsp. <i>patens</i> , 2.5 | GI:167999745 | (1405866..1407695) | GI:167999617 | MEP α |
| <i>Physcomitrella patens</i> subsp. <i>patens</i> , 2.6 | GI:168020770 | 823947..825718 | GI:168020544 | MEP α |
| <i>Physcomitrella patens</i> subsp. <i>patens</i> , 2.7 | GI:168003095 | (2804682..2807626) | GI:168003073 | MEP α |
| <i>Picea sitchensis</i> | N/A | N/A | GI:148909475 | MEP α |
| <i>Populus trichocarpa</i> , <i>PtrAMT1</i> ;1 | GI:116256305 | 5575440..5576936 | GI:224107548 | AMT-Euk |
| <i>Populus trichocarpa</i> , <i>PtrAMT1</i> ;2 | GI:116256300 | (2769026..2770821) | GI:224145860 | AMT-Euk |
| <i>Populus trichocarpa</i> , <i>PtrAMT1</i> ;3 | GI:116256303 | 11704820..11706966 | GI:224099996 | AMT-Euk |
| <i>Populus trichocarpa</i> , <i>PtrAMT1</i> ;4 | GI:116256317 | (23601578..23603065) | GI:224070026 | AMT-Euk |
| <i>Populus trichocarpa</i> , <i>PtrAMT1</i> ;5 | GI:116256317 | 23599283..23600819 | GI:224065518 | AMT-Euk |
| <i>Populus trichocarpa</i> , <i>PtrAMT1</i> ;6 | GI:116256304 | (7746820..7748217) | GI:224106262 | AMT-Euk |
| <i>Populus trichocarpa</i> , <i>PtrAMT2</i> ;1 | GI:116256301 | (7169018..7171902) | GI:224091027 | MEP α |
| <i>Populus trichocarpa</i> , <i>PtrAMT2</i> ;2 | GI:116256311 | (11521306..11523746) | GI:224140458 | MEP α |
| <i>Populus trichocarpa</i> , <i>PtrAMT3</i> ;1 | GI:116256298 | (26343416..26346328) | GI:224060484 | MEP α |
| <i>Populus trichocarpa</i> , <i>PtrAMT4</i> ;1 | GI:116256317 | (3007991..3009863) | GI:224066328 | MEP α |
| <i>Populus trichocarpa</i> , <i>PtrAMT4</i> ;2 | GI:116256313 | 5044547..5046163 | GI:224141842 | MEP α |
| <i>Populus trichocarpa</i> , <i>PtrAMT4</i> ;3 | GI:116256320 | 14994950..14996903 | GI:224082701 | MEP α |
| <i>Populus trichocarpa</i> , <i>PtrAMT4</i> ;4 | GI:116256308 | 2765333..2767047 | GI:224123403 | MEP α |
| <i>Populus trichocarpa</i> , <i>PtrAMT4</i> ;5 | GI:116265052 | (87989..89560) | GI:224133945 | MEP α |
| <i>Selaginella moellendorffii</i> , 1.1 ^U | scaffold 8 | 3112678-3114033 | N/A | AMT-Euk |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|--|---|-------------------|--------------------------|--------------|
| <i>Selaginella moellendorffii</i> , 1.2 [†] | scaffold 0 | (2254617-2256011) | N/A | AMT-Euk |
| <i>Selaginella moellendorffii</i> , 2.1 [†] | scaffold 37 | (1469550-1470973) | N/A | MEP α |
| <i>Selaginella moellendorffii</i> , 2.2 [†] | scaffold 62 | (1059095-1060581) | N/A | MEP α |
| <i>Selaginella moellendorffii</i> , 2.3 [†] | scaffold_28 | 1038805-1040650 | N/A | MEP α |
| <i>Selaginella moellendorffii</i> * [†] | scaffold 22 | 764553:766289 | N/A | N/A |
| <i>Selaginella moellendorffii</i> * [†] | scaffold 6 | 374638-376190 | N/A | N/A |
| <i>Selaginella moellendorffii</i> * [†] | scaffold 14 | 1451079-1452493 | N/A | N/A |
| <i>Solanum lycopersicum</i> , AMT1;1 | N/A | N/A | X92854 (EMBL) | AMT-Euk |
| <i>Solanum lycopersicum</i> , AMT1;2 | N/A | NA | X95098 (EMBL) | AMT-Euk |
| <i>Solanum lycopersicum</i> , AMT1;3 | N/A | N/A | GI:9992843 | AMT-Euk |
| 121 <i>Vitis vinifera</i> , 1.1 | gi157336920 | 2280597..2281994 | N/A | AMT-Euk |
| <i>Vitis vinifera</i> , 1.2 | GI:157336199 | 7949274..7950782 | N/A | AMT-Euk |
| <i>Vitis vinifera</i> , 2.1 | gi157357314 | 580580..583260 | N/A | MEP α |
| <i>Vitis vinifera</i> , 2.2 | gi157351961 | 1965726..1969581 | N/A | MEP α |
| <i>Vitis vinifera</i> , 2.3 | GI:157339171 | 4545828..4548058 | N/A | MEP α |
| <i>Vitis vinifera</i> , 2.4 | gi157347782 | 2402288..2404788 | N/A | MEP α |
| <i>Vitis vinifera</i> , 2.5 | gi157351961 | 1965726..1969581 | N/A | MEP α |
| <i>Vitis vinifera</i> , 2.6 | gi157329134 | 88960..90484 | N/A | MEP α |
| <i>Vitis vinifera</i> , 2.7 | gi157344876 | 468779..471094 | N/A | MEP α |
| <i>Vitis vinifera</i> , 2.8 | gi157344876 | 487711..495387 | N/A | MEP α |
| <i>Vitis vinifera</i> , 2.9* | GI:157344876 | (387252..389393) | N/A | N/A |
| <i>Volvox carteri</i> f. nagariensis* [†] | scaffold 90 | 149010-154761 | jgi Volca1 84428 estExt_ | N/A |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|-----------------|--|---------|
| <i>Volvox carteri</i> f. <i>nagariensis</i> * ¹ | scaffold 55 | (276244-288704) | Genewise1Plus.C_900018 jgi JGI_CBHO29025.fwd (partial) | N/A |
| <i>Volvox carteri</i> f. <i>nagariensis</i> , 1.1 ¹ | scaffold 6 | 1983780-1987620 | jgi Volca1 87801 fgenesH 4_pg.C_scaffold_6000238 | AMT-Euk |
| <i>Volvox carteri</i> f. <i>nagariensis</i> , 1.2 ¹ | scaffold 99 | 273062-280987 | jgi Volca1 69712 e_gw1. 99.26.1 | AMT-Euk |
| <i>Volvox carteri</i> f. <i>nagariensis</i> , 1.4 ¹ | scaffold_49 | 676000-681368 | jgi Volca1 65690 e_gw1. 49.57.1 | AMT-Euk |
| <i>Volvox carteri</i> f. <i>nagariensis</i> , 1.5 ¹ | scaffold 22 | 1100000 1104000 | jgi Volca1 61118 e_gw1. 22.94.1 | AMT-Euk |
| <i>Volvox carteri</i> f. <i>nagariensis</i> , 1.6 ¹ | scaffold_19 | 420000-423583 | jgi Volca1 41671 gw1.19. 140.1 | AMT-Euk |
| <i>Volvox carteri</i> f. <i>nagariensis</i> , 1.7 ¹ | scaffold 76 | 59023-63486 | jgi Volca1 77636 estExt_ Genewise1.C_760002 | AMT-Euk |
| <i>Volvox carteri</i> f. <i>nagariensis</i> , 1.8 ¹ | scaffold 67 | 468024-471998 | jgi Volca1 67734 e_gw1. 67.86.1 | AMT-Euk |
| <i>Volvox carteri</i> f. <i>nagariensis</i> , 1.9 ¹ | scaffold 26 | 541019-544444 | N/A | AMT-Euk |
| <i>Volvox carteri</i> f. <i>nagariensis</i> , 1.10 ¹ | scaffold 30 | 1150194-1154551 | jgi Volca1 75428 estExt_ Genewise1.C_300157 | AMT-Euk |
| Red algae (Rhodophyta) | | | | |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|--|--|------------------------|--------------------|--------------|
| <i>Cyanidioschyzon merolae</i> , 2 ^T | CMT526C | c20f0010_103914-105623 | N/A | AMT-Euk |
| <i>Cyanidioschyzon merolae</i> , AMT1 ^T | CMK126C | c11f0001_339165-340283 | N/A | AMT-Euk |
| <i>Cyanidioschyzon merolae</i> , 2 ^T | CMT526C | c20f0010_103914-105623 | N/A | AMT-Euk |
| <i>Galdiera sulphuraria</i> ^M | Supercontig 18 (Genome build 3.0 May 2007) | 44293-45576 | N/A | AMT-Euk |
| 123 Fungi | | | | |
| <i>Ajellomyces dermatitidis</i> , <i>mep3</i> | GI:154282232 | 1764968..1766804 | GI:154281506 | MEP γ |
| <i>Aspergillus fumigatus</i> , <i>meaA</i> | GI:71025129 | 1654974..1656856 | GI:70989734 | MEP γ |
| <i>Aspergillus fumigatus</i> , <i>mep2</i> | GI:71025128 | (2853321..2854904) | GI:70995403 | MEP γ |
| <i>Aspergillus fumigatus</i> , <i>mepA</i> | GI:71025132 | (2822229..2823889) | GI:70997608 | MEP γ |
| <i>Aspergillus nidulans</i> , <i>meaA</i> | GI:18693012 | 1478..>3330 | N/A | MEP γ |
| <i>Aspergillus nidulans</i> , <i>mepA</i> | GI:18693014 | 1815..3504 | N/A | MEP γ |
| <i>Batrachochytrium dendrobatidis</i> ^B | Supercontig 1 | 359666-361321 | N/A | MEP γ |
| <i>Candida albicans</i> , MEP1 | GI:68468731 | (19689..21293) | GI:68468505 | MEP γ |
| <i>Candida albicans</i> , MEP2 | GI:68485553 | 54760..56202 | GI:68485539 | MEP γ |
| <i>Candida albicans</i> , 3* | GI:68484429 | 66692..68434 | GI:68484407 | N/A |
| <i>Coccidioides immitis</i> , 1 | GI:119185453 | (3257148..3258822) | GI:119184212 | MEP γ |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|--------------------|--------------------|--------------|
| <i>Coccidioides immitis</i> , 2 | GI:119197146 | 2610936..2613096 | GI:119193249 | MEP γ |
| <i>Coprinopsis cinerea</i> okayama7#130, 1 | GI:169866184 | 17197..18998 | GI:169866180 | MEP γ |
| <i>Coprinopsis cinerea</i> okayama7#130, 2 | GI:169866153 | 6758..8559 | GI:169866149 | MEP γ |
| <i>Coprinopsis cinerea</i> okayama7#130, 3 | GI:169844956 | 27679..29664 | GI:169844550 | MEP γ |
| <i>Coprinopsis cinerea</i> okayama7#130, 4 | GI:169850323 | (130995..133174) | GI:169850235 | MEP γ |
| <i>Cryptococcus neoformans</i> var. <i>neoformans</i> JEC21, AMT1 | GI:58271746 | (585837..588247) | GI:58258402 | MEP γ |
| <i>Cryptococcus neoformans</i> var. <i>neoformans</i> JEC21, AMT2 | GI:58271747 | (549599..551680) | GI:58259901 | MEP γ |
| <i>Fusarium graminearum</i> PH-1, 1 | GI:46562332 | 1934463..1936246 | GI:46107473 | MEP γ |
| <i>Fusarium graminearum</i> PH-1, 2 | GI:46562332 | (6866192..6867738) | GI:46110424 | MEP γ |
| <i>Fusarium graminearum</i> PH-1, 3 | GI:46562332 | 1609324..1611001 | GI:46107291 | MEP γ |
| <i>Fusarium graminearum</i> PH-1, 4 | GI:46562343 | (765169..766784) | GI:46114441 | MEP γ |
| <i>Fusarium oxysporum</i> f. sp. <i>lycopersici</i> 4286, 1 | GI:144922572 | (54000-55395) | N/A | MEP γ |
| <i>Fusarium oxysporum</i> f. sp. <i>lycopersici</i> 4286, 2 | GI:144922350 | 57760-59300 | N/A | MEP γ |
| <i>Fusarium oxysporum</i> f. sp. <i>lycopersici</i> 4286, 3 | GI:144921882 | 33840-35337 | N/A | MEP γ |
| <i>Fusarium oxysporum</i> f. sp. <i>lycopersici</i> 4286, 4 | GI:144922346 | (65941-67381) | N/A | MEP α |
| <i>Fusarium oxysporum</i> f. sp. <i>lycopersici</i> | GI:144922346 | 66000-67170 | N/A | N/A |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|--|---|-------------------|--------------------|--------------|
| 4286, 5 | | | | |
| <i>Gibberella moniliformis</i> 7600, 1 | GI:116139414 | 155125-156539 | N/A | MEP γ |
| <i>Gibberella moniliformis</i> 7600, 2 | GI:116139519 | (424078-425576) | N/A | MEP γ |
| <i>Gibberella moniliformis</i> 7600, 3* | GI:116139519 | ~97480~99000 | N/A | N/A |
| <i>Gibberella moniliformis</i> 7600, 4 | GI:116139410 | 132661-134281 | N/A | MEP α |
| <i>Gibberella moniliformis</i> 7600 * | GI:116139361 | 59055-60699 | N/A | N/A |
| <i>Glomus intraradices</i> , <i>Gintamt1</i> | N/A | N/A | GI:124516978 | MEP γ |
| <i>Hebeloma cylindrosporum</i> , <i>AMT1</i> | N/A | N/A | GI:20501862 | MEP γ |
| <i>Hebeloma cylindrosporum</i> , <i>AMT2</i> | N/A | N/A | GI:15042692 | MEP γ |
| <i>Hebeloma cylindrosporum</i> , <i>AMT3</i> | N/A | N/A | GI:15042694 | MEP γ |
| 125 <i>Laccaria bicolor</i> , <i>AMT1.1a</i> | GI:170092085 | (1141763..1143937 | GI:170091535 | MEP γ |
| <i>Laccaria bicolor</i> , <i>AMT1.1b</i> | GI:170115426 | 153134..155099 | GI:17011531 | MEP γ |
| <i>Laccaria bicolor</i> , <i>AMT1.2</i> | GI:170103048 | (848838..851040) | GI:170103012 | MEP γ |
| <i>Laccaria bicolor</i> , <i>AMT2.1</i> | GI:170086374 | (798076..800048 | GI:170085394 | MEP γ |
| <i>Laccaria bicolor</i> , <i>AMT2.2</i> | GI:170110243 | 25357..26969 | GI:170109925 | MEP γ |
| <i>Laccaria bicolor</i> , <i>AMT2.3a</i> | GI:170112646 | (379312..382864) | GI:170112632 | MEP γ |
| <i>Laccaria bicolor</i> , <i>AMT2.3b</i> | GI:170112646 | (368498..370363) | GI:170112630 | MEP γ |
| <i>Laccaria bicolor</i> , <i>AMT2.3c</i> | GI:170115279 | 60031..61896 | GI:170115107 | MEP γ |
| <i>Magnaporthe grisea</i> 70-15, 1 | GI:38104209 | 14334..16198 | N/A | MEP γ |
| <i>Magnaporthe grisea</i> 70-15, 2 | GI:38108755 | (22223..23821) | N/A | MEP γ |
| <i>Magnaporthe grisea</i> 70-15, 3 | GI:38101984 | (9146..10992 | N/A | MEP γ |
| <i>Microbotryum violaceum</i> , <i>mepA</i> | N/A | N/A | GI:5230671 | MEP γ |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|----------------------|--------------------|--------------|
| <i>Neurospora crassa</i> , 1 | GI:164426245 | 631321..633365 | GI:164426202 | MEP γ |
| <i>Neurospora crassa</i> , 2 | GI:157070728 | 349091..350779 | GI:85107332 | MEP γ |
| <i>Neurospora crassa</i> , 3 | GI:164427336 | (221738..223494) | GI:85112787 | MEP γ |
| <i>Neurospora crassa</i> , 4 | GI:164426245 | 631321..633365 | GI:164426202 | MEP γ |
| <i>Penicillium chrysogenum</i> , 1 | GI:256353024 | (8082430..8084162) | GI:255937472 | MEP γ |
| <i>Penicillium chrysogenum</i> , 2 | GI:256353024 | (26853194..26855042) | GI:255948191 | MEP γ |
| <i>Penicillium chrysogenum</i> , 3 | GI:211588786 | (2625848..2627524) | GI:255954772 | MEP γ |
| <i>Penicillium chrysogenum</i> * | GI:256353024 | (21836862..21838538) | GI:255954772 | N/A |
| <i>Penicillium marneffei</i> , 1 | GI:212544102 | (1035187..1037338) | GI:212542532 | MEP γ |
| <i>Penicillium marneffei</i> , 2 | GI:212536833 | (165518..167568) | GI:212534103 | MEP γ |
| <i>Penicillium marneffei</i> , 3 | GI:212530903 | 3680024..3682106 | GI:212528861 | MEP γ |
| <i>Penicillium marneffei</i> , 4 | GI:212530903 | (4869932..4871615) | GI:212529771 | MEP α |
| <i>Penicillium marneffei</i> , 5 | GI:212534004 | 212731..214581 | GI:212531034 | MEP γ |
| <i>Pyrenophora tritici-repentis</i> , 1 | GI:189192591 | 2218689..2220719 | GI:189189627 | MEP γ |
| <i>Pyrenophora tritici-repentis</i> , 2 | GI:189206496 | (1468324..1469966) | GI:189206414 | MEP γ |
| <i>Pyrenophora tritici-repentis</i> , 3 | GI:189192591 | 5135420..5137396 | GI:189191563 | MEP γ |
| <i>Pyrenophora tritici-repentis</i> , 4 | GI:189196879 | 1673492..1675071 | GI:189196725 | MEP γ |
| <i>Pyrenophora tritici-repentis</i> , 5 | GI:189198580 | 1673492..1675071 | GI:189197910 | MEP γ |
| <i>Rhizopus oryzae</i> , 1 ^B | Supercontig 3 | (407915-410079) | N/A | MEP γ |
| <i>Rhizopus oryzae</i> , 2 ^B | Supercontig 9 | (212386-214521) | N/A | MEP γ |
| <i>Rhizopus oryzae</i> , 3 ^B | Supercontig 10 | 1263026-1264373 | N/A | MEP γ |
| <i>Rhizopus oryzae</i> , 4 ^B | Supercontig 10 | 1726208-1727707 | N/A | MEP γ |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|--------------------|--------------------|--------------|
| <i>Rhizopus oryzae</i> , 5 ^B | Supercontig 4 | (668186-669408) | N/A | MEP γ |
| <i>Saccharomyces cerevisiae</i> , MEP1 | GI:162949218 | (731454..732932) | N/A | MEP γ |
| <i>Saccharomyces cerevisiae</i> , MEP2 | GI:117937805 | 357455..358954 | N/A | MEP γ |
| <i>Saccharomyces cerevisiae</i> , MEP3 | GI:50593503 | (810980..812449) | N/A | MEP γ |
| <i>Schizosaccharomyces pombe</i> , amt1 | GI:63054406 | 2069806..2071299 | GI:68037248 | MEP γ |
| <i>Schizosaccharomyces pombe</i> , amt2 | GI:162312575 | 1731627..1733165 | GI:67999505 | MEP γ |
| <i>Schizosaccharomyces pombe</i> , amt3 | GI:162312575 | (2924893..2926446) | GI:68000034 | MEP γ |
| <i>Sclerotinia sclerotiorum</i> , 1 | GI:156053526 | (748453..750262) | GI:156052866 | MEP γ |
| <i>Sclerotinia sclerotiorum</i> , 2 | GI:156058194 | (1380869..1382769) | GI:156057540 | MEP γ |
| <i>Sclerotinia sclerotiorum</i> , 3 | GI:156056511 | 145159..146698 | GI:156055065 | MEP γ |
| 127 <i>Talaromyces stipitatus</i> , 1 | GI:242771011 | (4296843..4298562) | GI:242768291 | MEP γ |
| <i>Talaromyces stipitatus</i> , 2 | GI:242797504 | (198426..200031) | GI:218717854 | MEP γ |
| <i>Talaromyces stipitatus</i> , 3 | GI:242771011 | (2855574..2857925) | GI:242765292 | MEP γ |
| <i>Talaromyces stipitatus</i> , 4 | GI:242771011 | 1387342..1389023 | GI:242762291 | MEP α |
| <i>Talaromyces stipitatus</i> , 5 | GI:218721414 | 259064..260882 | GI:242771613 | MEP γ |
| <i>Tuber borchii</i> , AMT1 | N/A | N/A | GI:17221149 | MEP γ |
| <i>Ustilago maydis</i> 521, mep1 | GI:49083521 | (110635..112196) | GI:71023612 | MEP γ |
| <i>Ustilago maydis</i> 521, ump2 | GI:49083411 | 45715..47301 | GI:71020878 | MEP γ |
| Animals | | | | |
| <i>Anopheles gambiae</i> str. PEST | N/A | N/A | GI:31323878 | AMT-Euk |
| <i>Caenorhabditis briggsae</i> , amt-1 | N/A | N/A | GI:268579152 | AMT-Euk |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|--|---|----------------------|--------------------|---------|
| <i>Caenorhabditis briggsae</i> , amt-2 | N/A | N/A | GI:268568449 | AMT-Euk |
| <i>Caenorhabditis briggsae</i> , amt-3 | N/A | N/A | GI:268530299 | AMT-Euk |
| <i>Caenorhabditis briggsae</i> , amt-4 | N/A | N/A | GI:268579150 | AMT-Euk |
| <i>Caenorhabditis elegans</i> , amt-1 | GI:193211360 | 4571421..4573867 | GI:71982374 | AMT-Euk |
| <i>Caenorhabditis elegans</i> , amt-2 | GI:212659372 | 13042410..13046826 | GI:133931099 | AMT-Euk |
| <i>Caenorhabditis elegans</i> , amt-3 | GI:212645681 | (8369089 .. 8363126) | GI:193204786 | AMT-Euk |
| <i>Caenorhabditis elegans</i> , amt-4 | GI:193211360 | 4565276..4569533 | GI:86564584 | AMT-Euk |
| <i>Ciona intestinalis</i> , Amt1 | N/A | N/A | GI:46095018 | AMT-Euk |
| <i>Ciona intestinalis</i> , Amt2 | N/A | N/A | GI:46095022 | AMT-Euk |
| <i>Ciona savignyi</i> , Amt1 | N/A | N/A | GI:5917780 | AMT-Euk |
| <i>Ciona savignyi</i> , Amt2 | N/A | N/A | GI:74136060 | AMT-Euk |
| <i>Drosophila melanogaster</i> CG6499 % | GI:56411841 | (11072969..11074859) | GI:221379345 | AMT-Euk |
| <i>Drosophila yakuba</i> strain Tai18E2 | N/A | N/A | GI:195501395 | AMT-Euk |
| <i>Strongylocentrotus purpuratus</i> , Amt1* | GI:115768329 | 2436..18327 | GI:115768328 | AMT-Euk |
| <i>Strongylocentrotus purpuratus</i> , Amt2 | GI:115812275 | (191051..208427) | GI:115812240 | AMT-Euk |
| <i>Strongylocentrotus purpuratus</i> , Amt3 | GI:115965798 | 617..15309 | GI:115965789 | AMT-Euk |
| Slime molds (Amoebozoa) | | | | |
| <i>Dictyostelium discoideum</i> AX4, amtA% | GI:269316049 | (8113998-8112427) | GI:31323872 | AMT-Euk |
| <i>Dictyostelium discoideum</i> AX4, amtB | GI:269316126 | 276805-278197 | GI:22293681 | AMT-Euk |
| <i>Dictyostelium discoideum</i> AX4, amtC | GI:269316041 | 1034624-1036114 | GI:31323876 | grade |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|--|---|---------------|--------------------|---------|
| Chromalveolates | | | | |
| Alveolates | | | | |
| <i>Paramecium tetraurelia</i> , 1.1 | N/A | N/A | GI:145536026 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.2 | N/A | N/A | GI:145499883 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.3 | N/A | N/A | GI:145499228 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.4 | N/A | N/A | GI:145541595 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.5 | N/A | N/A | GI:145551996 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.6 | N/A | N/A | GI:145545836 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.7 | N/A | N/A | GI:145513123 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.8 | N/A | N/A | GI:145539663 | AMT-Euk |
| ¹²⁹ <i>Paramecium tetraurelia</i> , 1.9 | N/A | N/A | GI:145479558 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.10 | N/A | N/A | GI:145534018 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.11 | N/A | N/A | GI:145549086 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.12 | N/A | N/A | GI:145483072 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.13 | N/A | N/A | GI:145540729 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.14 | N/A | N/A | GI:145479880 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.15 | N/A | N/A | GI:145540001 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.16 | N/A | N/A | GI:145552095 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.17 | N/A | N/A | GI:145547943 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.18 | N/A | N/A | GI:145476398 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.19 | N/A | N/A | GI:145515630 | AMT-Euk |
| <i>Paramecium tetraurelia</i> , 1.20 | N/A | N/A | GI:145506207 | AMT-Euk |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|-------------------|--------------------|---------|
| <i>Paramecium tetraurelia</i> , 1.21* | N/A | N/A | GI:145549312 | N/A |
| <i>Perkinsus marinus</i> ATCC 50983, 1.1 | GI:239890294 | 158917-160404 | N/A | AMT-Euk |
| <i>Perkinsus marinus</i> ATCC 50983, 1.2 | GI:239879718 | 20760-25116 | N/A | AMT-Euk |
| <i>Perkinsus marinus</i> ATCC 50983, 1.3 | GI:239890061 | 257746-251045 | N/A | AMT-Euk |
| <i>Perkinsus marinus</i> ATCC 50983, 1.4 | GI:239866251 | (963-306) | N/A | AMT-Euk |
| <i>Perkinsus marinus</i> ATCC 50983, 1.5 | GI:239879718 | 34129-35281 | N/A | AMT-Euk |
| <i>Perkinsus marinus</i> ATCC 50983, 1.6 | GI:239879784 | 666714-668240 | N/A | AMT-Euk |
| <i>Perkinsus marinus</i> ATCC 50983, 1.7 | GI:239890294 | 156592-158112 | N/A | AMT-Euk |
| <i>Perkinsus marinus</i> ATCC 50983, 1.8* | GI:239866277 | (1132-173) | N/A | N/A |
| Stramenopiles (Heterokonts) | | | | |
| 130 <i>Aureococcus anophagefferens</i> , 1.1 ^l | scaffold_12 | 1045276-1045419 | N/A | AMT-Euk |
| <i>Aureococcus anophagefferens</i> , 1.2 ^l | scaffold 2 (cDNA) | (523655-522243) | N/A | AMT-Euk |
| <i>Aureococcus anophagefferens</i> , 1.3 ^l | scaffold_25 | 523671-52529 | N/A | AMT-Euk |
| <i>Aureococcus anophagefferens</i> , 1.4 ^l | scaffold 10 | 247160-248532 | N/A | AMT-Euk |
| <i>Aureococcus anophagefferens</i> , 1.5 ^l | scaffold 2 | 3112-165-3113127 | N/A | AMT-Euk |
| <i>Aureococcus anophagefferens</i> , 1.6 ^l | scaffold 6 | 2161438-2162691 | N/A | AMT-Euk |
| <i>Aureococcus anophagefferens</i> , 1.7 ^l | scaffold 1 | (1200981-1202402) | N/A | AMT-Euk |
| <i>Aureococcus anophagefferens</i> , 1.8 ^l | scaffold 29 | (145699-146665) | N/A | AMT-Euk |
| <i>Aureococcus anophagefferens</i> , 1.9 ^l | scaffold 29 | (147244-145717) | N/A | AMT-Euk |
| <i>Aureococcus anophagefferens</i> , 1.10 ^l | scaffold 2 (genomic) | (523655-522243) | N/A | AMT-Euk |
| <i>Aureococcus anophagefferens</i> , 1.11 ^l | scaffold 45 | 20487-21639 | N/A | AMT-Euk |
| <i>Cylindrotheca fusiformis</i> , AMT1 | N/A | N/A | GI:56068113 | AMT-Euk |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|--|--|--------------------------------|---------------------------|--------------|
| <i>Cylindrotheca fusiformis</i> , AMT2a* | N/A | N/A | GI:56068115 | N/A |
| <i>Phytophthora infestans</i> T30-4, 1 ^J | >gnl WGS:AATU c ont1.466 | 35000-39000 | N/A | MEP β |
| <i>Phytophthora infestans</i> T30-4, 2 ^J | >gnl WGS:AATU c ont1.2760 | 2830-4800 | N/A | MEP β |
| <i>Phytophthora infestans</i> T30-4, 3 ^J | >gnl WGS:AATU c ont1.8908 | 1-700 | N/A | MEP β |
| <i>Thalassiosira pseudonana</i> CCMP1335, 1 ^J | Chromosome 7 | (503000-505000) | N/A | AMT-Euk |
| <i>Thalassiosira pseudonana</i> CCMP1335, AMT2 ^J | Chromosome 2 | 1429802-1431446 | N/A | AMT-Euk |
| ¹⁵¹ <i>Thalassiosira pseudonana</i> CCMP1335, AMT3 (1) ^J | Chromosome 9 | 391755-393784 | N/A | AMT-Euk |
| <i>Thalassiosira pseudonana</i> CCMP1335, AMT3 (2) ^J | Chromosome 9 | (393784-395145) | N/A | AMT-Euk |
| <i>Thalassiosira pseudonana</i> CCMP1335, AMT4 (1) ^J | Chromosome 2 | (671116-672906) | N/A | AMT-Euk |
| <i>Thalassiosira pseudonana</i> CCMP1335, AMT4 (2) ^J | Chromosome 2 | 669044-666139 | N/A | AMT-Euk |
| <i>Thalassiosira pseudonana</i> CCMP1335, AMT4 (3) ^J | Chromosome 2 | (671116-672906) (duplicate) | N/A | AMT-Euk |
| <i>Thalassiosira pseudonana</i> CCMP1335, AMT4 (4) ^J | Chromosome 2 | 669044-666139 (duplicate) | N/A | AMT-Euk |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|--|-----------------------------|---------------------------|--------------|
| <i>Thalassiosira pseudonana</i> CCMP1335, AMT4 (5) ^J | Chromosome 2 | 669044-666139 (duplicate) | N/A | AMT-Euk |
| <i>Thalassiosira pseudonana</i> CCMP1335, AMT4 (6) ^J | Chromosome 2 | (671116-672906) (duplicate) | N/A | AMT-Euk |
| <i>Thalassiosira pseudonana</i> CCMP1335, AMT4 (7) ^J | Chromosome 2 | 669044-666139 (duplicate) | N/A | AMT-Euk |
| <i>Thalassiosira pseudonana</i> CCMP1335, AMT6 (1) ^J | Chromosome 4 | 2293792-2295870 | N/A | AMT-Euk |
| <i>Thalassiosira pseudonana</i> CCMP1335, AMT6 (2) ^J | Chromosome 4 | (2306493-2308491) | N/A | AMT-Euk |
| ¹³ <i>Thalassiosira pseudonana</i> CCMP1335, THAPSDRAFT_36263 ^J | Chromosome 9 | 393783-395190 | N/A | AMT-Euk |
| <i>Thalassiosira pseudonana</i> CCMP1335, THAPSDRAFT_40537 (1) ^J | Chromosome 4 | 2306899-2308341 | N/A | AMT-Euk |
| <i>Thalassiosira pseudonana</i> CCMP1335, THAPSDRAFT_40537 (2) ^J | Chromosome 4 | (293942-2295300) | N/A | AMT-Euk |
| Excavates | | | | |
| Euglenozoa | | | | |
| <i>Trypanosoma cruzi</i> strain CL Brener | GI:70876681 | 73022-74566 | GI:71421943 | grade |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|------------------|--------------------|--------------|
| Heterolobosea | | | | |
| <i>Naegleria gruberi</i> , 1.1 [†] | scaffold_10 | 386054-387543 | N/A | AMT-Euk |
| <i>Naegleria gruberi</i> , 1.2 [†] | scaffold_30 | 105142-111073 | N/A | AMT-Euk |
| <i>Naegleria gruberi</i> , amtA [†] | scaffold_4 | 165684-167083 | N/A | grade |
| PROKARYOTES | | | | |
| Archaea | | | | |
| <i>Archaeoglobus fulgidus</i> , amt-1 | GI:11497621 | 879899..881074 | N/A | MEP α |
| <i>Archaeoglobus fulgidus</i> , amt-2 | GI:11497621 | 1568078..1569484 | N/A | grade |
| <i>Archaeoglobus fulgidus</i> , amt-3 | GI:11497621 | 1570632..1571798 | N/A | MEP α |
| ¹³¹ <i>Caldivirga maquilingsensis</i> IC-167 | GI:159040592 | 715687..717117 | N/A | MEP α |
| <i>Ferroplasma acidarmanus</i> | GI:126008996 | 221620..222921 | N/A | MEP α |
| <i>Ferroplasma acidarmanus</i> | GI:126009238 | (48378..49760) | N | grade |
| <i>Methanosarcina acetivorans</i> C2A, amtB (1) | GI:20088899 | 4815556..4816764 | N/A | grade |
| <i>Methanosarcina acetivorans</i> C2A, amtB (2) | GI:20088899 | 4817101..4818315 | N/A | MEP β |
| <i>Methanosarcina acetivorans</i> C2A, amt | GI:20088899 | 5130958..5132367 | N/A | grade |
| <i>Methanosarcina barkeri</i> Meth_1922, 2 | GI:23052039 | (985..2202) | N/A | grade |
| <i>Methanosarcina barkeri</i> Meth_1922, 1 | GI:23052039 | (2418..3617) | N/A | MEP β |
| <i>Methanosarcina barkeri</i> Meth_1922, 3 | GI:23051535 | 16186..17595 | N/A | grade |
| <i>Methanosarcina mazei</i> Go1 1 % | GI:21226102 | 881714..882913 | N/A | MEP β |
| <i>Methanosarcina mazei</i> Go1 2 | GI:21226102 | 1137526..1138935 | N/A | grade |
| <i>Methanosarcina mazei</i> , amtB1* | GI:18496419 | 537..1730 | N/A | N/A |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|--|---|------------------|--------------------|--------------|
| <i>Methanosphaera stadtmanae</i> DSM 3091 | GI:84488831 | 774206-775429 | N/A | MEP α |
| <i>Methanothermobacter thermautotrophicus</i> str. Delta H, 1 | GI:15678031 | 588943-590166 | N/A | MEP α |
| <i>Methanothermobacter thermautotrophicus</i> str. Delta H, 2 | GI:15678031 | 590781-592019 | N/A | MEP α |
| <i>Picrophilus torridus</i> DSM 9790 | GI:48477072 | 1460613-1461911 | N/A | MEP α |
| <i>Sulfolobus tokodaii</i> str. 7 % | GI:24473558 | 643486-645042 | N/A | MEP α |
| <i>Sulfolobus solfataricus</i> | GI:15896971 | 914084..915613 | N/A | MEP α |
| Bacteria | | | | |
| ¹³¹ <i>Acidimicrobium ferrooxidans</i> DSM 10331 | GI:256370824 | 2108845-2110290 | N/A | MEP α |
| <i>Acidithiobacillus caldus</i> ATCC 51756, <i>amt-1</i> | ACVD01000125.1 | 109..834 | N/A | MEP α |
| <i>Acidithiobacillus caldus</i> ATCC 51756, <i>amt-2*</i> | ACVD01000080.1 | 19056..20363 | N/A | N/A |
| <i>Acidithiobacillus caldus</i> ATCC 51756, <i>amtB</i> | ACVD01000076.1 | 13063..14280 | N/A | grade |
| <i>Acidithiobacillus ferrooxidans</i> ATCC 53993 <i>amt-2</i> | GI:198282148 | 2491440..2492756 | N/A | MEP β |
| <i>Acidithiobacillus ferrooxidans</i> ATCC 53993, <i>amt-1</i> | GI:198282148 | 2497431-2498735 | N/A | MEP α |
| <i>Acidithiobacillus ferrooxidans</i> ATCC 53993, <i>amtB</i> | GI:198282148 | 1554881-1556098 | N/A | grade |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|--------------------|--------------------|--------------|
| <i>Acidithiobacillus ferrooxidans</i> ATCC 23270 <i>amt-2</i> | GI:218665024 | 2589616-2590932 | N/A | MEP β |
| <i>Acidithiobacillus ferrooxidans</i> ATCC 23270 <i>amt-1</i> | GI:218665024 | 2595607-2596911 | N/A | MEP α |
| <i>Acidithiobacillus ferrooxidans</i> ATCC 23270 <i>amtB</i> | GI:218665024 | 1673774-1674991 | N/A | grade |
| <i>Agrobacterium tumefaciens</i> str. C58 (Cereon), <i>amtB</i> | GI:159184118 | 2757775..2758980 | N/A | MEP β |
| <i>Anabaena variabilis</i> ATCC 29413 | GI:75699950 | 3662100..3663656 | N/A | grade |
| <i>Aquifex aeolicus</i> VF5 | GI:15282445 | 69169-70440 | N/A | MEP α |
| ¹³ <i>Azoarcus</i> sp. BH72, <i>amtB</i> % | GI:19851906 | N/A | N/A | MEP β |
| <i>Azorhizobium caulinodans</i> , <i>amtB</i> | GI:2980822 | N/A | N/A | MEP β |
| <i>Azospirillum brasilense</i> , <i>amtB</i> | GI:3136067 | N/A | N/A | MEP β |
| <i>Azotobacter vinelandii</i> DJ, <i>amtB</i> | GI:226942170 | 4846081-4847397 | N/A | MEP β |
| <i>Azotobacter vinelandii</i> , <i>amtB</i> | GI:29359385 | 165528..166787 | N/A | MEP β |
| <i>Bacillus cereus</i> ATCC 14579 | GI:30018278 | (1135908..1137140) | N/A | MEP β |
| <i>Bacillus halodurans</i> C-125 | GI:57596592 | (3958228-3956963) | N/A | grade |
| <i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168 <i>amtB</i> | GI:255767013 | 3756790..3758004 | N/A | MEP β |
| <i>Bacteroides thetaiotaomicron</i> VPI-5482 | GI:29345410 | (670232..671689) | N/A | grade |
| <i>Bifidobacterium longum</i> DJO10A, <i>amtB</i> | GI:29251702 | (91351..92583) | N/A | MEP β |
| <i>Bifidobacterium longum</i> NCC2705, <i>amtP</i> | GI:58036264 | (1382167..1383462) | N/A | MEP β |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|--|---|--------------------|--------------------|-------------|
| <i>Bordetella bronchiseptica</i> RB50, <i>amtB</i> | GI:33575370 | (93085..94323) | N/A | MEP β |
| <i>Bordetella parapertussis</i> | GI:33573592 | (10300..11538) | N/A | MEP β |
| <i>Bradyrhizobium japonicum</i> USDA 110 chromosome* | GI:27375111 | 652211..653227 | N/A | N/A |
| <i>Bradyrhizobium japonicum</i> USDA 110 chromosome* | GI:27375111 | 652842..653513 | N/A | N/A |
| <i>Bradyrhizobium japonicum</i> USDA 110 chromosome, <i>amtB</i> | GI:27375111 | 655678..657210 | N/A | MEP β |
| <i>Brevibacterium linens</i> BL2 | GI:62422953 | (16634..17854) | N/A | MEP β |
| ¹³ <i>Brevibacterium linens</i> = <i>Burkholderia vietnamensis</i> G4 (2) | GI:134137285 | (1214773..1215966) | N/A | grade |
| <i>Brevibacterium linens</i> = <i>Burkholderia vietnamensis</i> G4 (1) | GI:134137285 | (3282231..3283730) | N/A | MEP β |
| <i>Brucella melitensis</i> | GI:17986284 | 161196..162497 | N/A | MEP β |
| <i>Brucella suis</i> 1330, <i>amt</i> | GI:56968325 | (1831408..1832709) | N/A | MEP β |
| <i>Brucella suis</i> ATCC 23445 chromosome I | GI:163673000 | 1647375-1648676 | N/A | MEP β |
| <i>Burkholderia fungorum</i> Bcep_273, 1 | GI:28873745 | 45938..47218 | N/A | MEP β |
| <i>Burkholderia fungorum</i> Bcep_282, 2 | GI:28873773 | (1955..3454) | N/A | grade |
| <i>Burkholderia</i> sp. 383 (= <i>Burkholderia cepacia</i> R-18194) (2) | GI:77965403 | (1256821..1258014) | N/A | grade |
| <i>Burkholderia</i> sp. 383 (= <i>Burkholderia cepacia</i> R-18194) (1) | GI:77965403 | (3341542..3343056) | N/A | MEP β |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|--------------------|--------------------|-------------|
| <i>Carboxydotherrmus hydrogenoformans</i> Z-2901 | GI:77994731 | 62220-63536 | N/A | MEP β |
| <i>Caulobacter crescentus</i> CB15 | GI:16124256 | (1492580..1494112) | N/A | MEP β |
| <i>Chlorobium chlorochromatii</i> CaD3 (= <i>Chlorochromatium aggregatum</i>) | GI:78170183 | 2533520..2534854 | N/A | MEP β |
| <i>Chlorobium tepidum</i> TLS, <i>amt-1</i> | GI:21672841 | (511031..512581) | N/A | grade |
| <i>Chlorobium tepidum</i> TLS, <i>amt-2</i> | GI:21672841 | (933265..934590) | N/A | grade |
| <i>Chlorobium tepidum</i> TLS, <i>amt-3</i> | GI:21672841 | 128241..12957 | N/A | MEP β |
| <i>Chloroflexus aurantiacus</i> J-10-fl, <i>amtB</i> | GI:163845603 | 1388274..1389689 | N/A | grade |
| 137 <i>Chromobacterium violaceum</i> ATCC 12472, <i>amtB</i> | GI:34495455 | 4329876..4331186 | N/A | MEP β |
| <i>Clostridium acetobutylicum</i> , <i>nrgA</i> | GI:15893970 | (787791..789008) | N/A | MEP β |
| <i>Clostridium thermocellum</i> ATCC 27405 | GI:28836975 | (37261..39015) | N/A | grade |
| <i>Corynebacterium efficiens</i> YS-314, 2 | GI:25028257 | (1790505..1791824) | N/A | MEP β |
| <i>Corynebacterium efficiens</i> YS-314, <i>amtP</i> | GI:25028524 | (2073802..2075127) | N/A | MEP β |
| <i>Corynebacterium glutamicum</i> ATCC 13032, 2 | GI:58036263 | 1675265-1676623 | N/A | MEP β |
| <i>Corynebacterium glutamicum</i> ATCC 13032, 1 | GI:58036263 | (2172151..2173467) | N/A | MEP β |
| <i>Cyanothece</i> sp. PCC 7822 | GI:196258658 | 210450..211856 | N/A | MEP β |
| <i>Cyanothece</i> sp. ATCC 51142 <i>amt1</i> | GI:172034917 | 3375443-3376957 | N/A | grade |
| <i>Cytophaga hutchinsonii</i> , <i>amtB</i> , 1 | GI:110636427 | (3786660..3787988) | N/A | grade |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|--------------------|----------------------|--------------|
| <i>Cytophaga hutchinsonii</i> , amtB, 2 | GI:110636427 | (4151178..4152431) | N/A | grade |
| <i>Cytophaga hutchinsonii</i> , amtB, 3 | GI:110636427 | 1047840..1049228 | N/A | grade |
| <i>Cytophaga hutchinsonii</i> , amtB, 4 | GI:110636427 | (4121951..4123291) | N/A | grade |
| <i>Dechloromonas aromatica</i> RCB, 1 | GI:71845263 | 62898..64397 | N/A | MEP β |
| <i>Dechloromonas aromatica</i> RCB, 2 | GI:71845263 | (3980601..3981803) | N/A | grade |
| <i>Deinococcus radiodurans</i> | GI:15805720 | 708354..709673 | N/A | MEP β |
| <i>Desulfitobacterium hafniense</i> | GI:30471518 | 19330..20676 | N/A | MEP β |
| <i>Desulfococcus oleovorans</i> Hxd3, 2 | GI:158520017 | (633089-634375) | N/A | MEP α |
| <i>Desulfococcus oleovorans</i> Hxd3, 1 | GI:158520017 | (634397-635680) | N/A | MEP β |
| <i>Desulfococcus oleovorans</i> Hxd3, 3 | GI:158520017 | (616751-618085) | N/A | grade |
| ¹³ <i>Desulfovibrio desulfuricans</i> G20, 1 | GI:28877455 | 62208..63413 | N/A | MEP β |
| <i>Desulfovibrio desulfuricans</i> G20, amtB | GI:28877458 | (394101..395486) | N/A | grade |
| <i>Desulfovibrio desulfuricans</i> , 2 | GI:78217452 | (2328899..2330104) | N/A | MEP β |
| <i>Desulfuromonas acetoxidans</i> , DSM 684 ctg55, 1 | GI:95133343 | 70185..71510 | N/A | MEP β |
| <i>Desulfuromonas acetoxidans</i> , DSM 684 ctg62, 2 | GI:95134209 | (94607..95920) | N/A | grade |
| <i>Desulfuromonas acetoxidans</i> , DSM 684 ctg65* | GI:95928471 | (24390..25790) | N/A | grade* |
| <i>Desulfuromonas acetoxidans</i> , DSM 684 ctg66* | GI:95928205 | 176919..17873 | N/A | grade* |
| <i>Desulfuromonas acetoxidans</i> , 3 | N/A | N/A | (#260 from Huang and | grade |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|--------------------|---|-------------|
| <i>Desulfuromonas acetoxidans</i> , 4 | N/A | N/A | Peng, 2005) (#259 from Huang and Peng, 2005) | grade |
| <i>Enterococcus faecium</i> | GI:22992526 | 3209..4423 | N/A | grade |
| <i>Escherichia coli</i> B str. REL606* | GI:253972022 | 445033-446319 | N/A | N/A |
| <i>Escherichia coli</i> str. K-12 substr. MG1655 | GI:49175990 | 472256..473473 | N/A | MEP β |
| <i>Exiguobacterium sibiricum</i> 255-15 | GI:171988566 | (737565..738860) | N/A | grade |
| <i>Geobacter metallireducens</i> GS-15, 2 | GI:23053086 | 93981..95207 | N/A | grade |
| <i>Geobacter metallireducens</i> GS-15, 1 | GI:23054753 | 7603..9069 | N/A | MEP β |
| <i>Geobacter sulfurreducens</i> PCA, 1 | GI:39995111 | 1009771..1011240 | N/A | MEP β |
| <i>Geobacter sulfurreducens</i> PCA, 2 | GI:39995111 | (1324618..1325838) | N/A | grade |
| <i>Gloeobacter violaceus</i> * | GI:37519569 | (2798589..2799443) | N/A | N/A |
| <i>Gloeobacter violaceus</i> | GI:37519569 | 3253283..3254719 | N/A | MEP β |
| <i>Gluconacetobacter diazotrophicus</i> , amtB2* | GI:209542188 | 1454476..1455855 | N/A | N/A |
| <i>Gluconacetobacter diazotrophicus</i> , amtB1 | GI:209542188 | 669752..671143 | N/A | MEP β |
| <i>Haemophilus somnus</i> 129PT, amtB | GI:28301892 | 19208..20506 | N/A | MEP β |
| <i>Lactobacillus casei</i> , str. Zhang, amtB | GI:300437532 | (592027-593343) | N/A | MEP β |
| <i>Lactobacillus plantarum</i> WCFS1, amtB | GI:28269805 | (318549..319865) | N/A | MEP β |
| <i>Lactococcus lactis</i> subsp cremoris MG1363, amtB | GI:124491690 | 876626..877867 | N/A | MEP β |
| <i>Lactococcus lactis</i> subsp. lactis II1403, amtB | GI:13400022 | (1635861..1637102) | N/A | MEP β |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|--|---|--------------------|--------------------|--------------|
| <i>Leptospira interrogans</i> serovar lai str. 56601, <i>amtB</i> * | GI:294827553 | (3577147..3579492) | N/A | N/A |
| <i>Leptospirillum rubarum</i> LeptoII_Scaffold_8241 | GI:124516098 | 100910-102217 | N/A | MEP α |
| <i>Leptospirillum</i> sp. Group II '5-way CG' Scaffold11277 | GI:206602886 | 108230-109537 | N/A | grade |
| <i>Leuconostoc mesenteroides</i> subsp. <i>mesenteroides</i> ATCC 8293 | GI:28876869 | 51528..52724 | N/A | MEP β |
| <i>Listeria innocua</i> | GI:16414035 | 2144..3362 | N/A | MEP β |
| <i>Listeria monocytogenes</i> EGD-e | GI:1641081 | 138688..139906 | N/A | MEP β |
| ¹⁴⁰ <i>Magnetococcus</i> sp. MC-1 | GI:23000293 | 19671..20981 | N/A | MEP β |
| <i>Magnetospirillum magnetotacticum</i> , <i>amb-1</i> | GI:23013338 | 5418..6665 | N/A | grade |
| <i>Mesorhizobium loti</i> , <i>amtB</i> | GI:57165207 | (3387960..3389315) | N/A | MEP β |
| <i>Methanocaldococcus jannaschii</i> , (1289834..1291096) | GI:15668172 | (1289834..1291096) | N/A | grade |
| <i>Methanocaldococcus jannaschii</i> , <i>amtB</i> * | GI:15668172 | (57588..58763) | N/A | N/A |
| <i>Methanopyrus kandleri</i> AV19 | GI:20093440 | 51991..53284 | N/A | grade |
| <i>Methylobacillus flagellatus</i> KT, 1 | GI:91774356 | 2693851-2695236 | N/A | MEP β |
| <i>Methylobacillus flagellatus</i> KT, 2 | GI:91774356 | 442257-443561 | N/A | MEP β |
| <i>Methylobacillus flagellatus</i> KT, 3 | GI:91774356 | (1890161-1891474) | N/A | grade |
| <i>Mycobacterium tuberculosis</i> H37Rv, <i>amt</i> | GI:41352756 | (112036..113469) | N/A | MEP β |
| <i>Nautilia profundicola</i> AmH, <i>amt</i> (1) | GI:224372070 | 185564-186844 | N/A | MEP α |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|--------------------|--------------------|-------------|
| <i>Nautilia profundicola</i> AmH, amt (2) | GI:224372070 | 351534-352706 | N/A | grade |
| <i>Nautilia profundicola</i> AmH (duplicate)* | GI:223588326 | 185564-186844 | N/A | N/A |
| <i>Neisseria meningitidis</i> MC58, amtB | GI:7735869 | (645072..646376) | N/A | MEP β |
| <i>Neisseria meningitidis</i> Z2491, amtB | GI:15793034 | (801446..802750) | N/A | MEP β |
| <i>Nostoc punctiforme</i> PCC 73102 (1.1) | GI:186680550 | 5780355-5781815 | N/A | AMT-Euk |
| <i>Nostoc punctiforme</i> PCC 73102 (1) | GI:186680550 | 4091213-4092625 | N/A | MEP β |
| <i>Nostoc punctiforme</i> PCC 73102* | GI:30581866 | 99651..101111 | N/A | N/A |
| <i>Nostoc punctiforme</i> PCC 73102* | GI:30581855 | (31482..32906) | N/A | N/A |
| <i>Nostoc</i> sp. PCC 7120, 3 | GI:17227497 | 1156507..1158003 | N/A | grade |
| <i>Nostoc</i> sp. PCC 7120* | GI:17227497 | 1158426..1159982 | N/A | N/A |
| ¹⁴¹ <i>Nostoc</i> sp. PCC 7120, 2 | GI:17227497 | 1160381..1161793 | N/A | MEP β |
| <i>Nostoc</i> sp. PCC 7120, amt1 | GI:17227497 | 5582484..5583356 | N/A | AMT-Euk |
| <i>Novosphingobium aromaticivorans</i> , DSM 12444 | GI:87198026 | 2449396..2450709 | N/A | MEP β |
| <i>Novosphingobium aromaticivorans</i> , DSM 12444* | GI:87198026 | 733046..734404 | N/A | N/A |
| <i>Oceanobacillus iheyensis</i> | GI:23097455 | (1260475..1261845) | N/A | grade |
| <i>Oenococcus oeni</i> MC | GI:29378884 | (34354..35700) | N/A | MEP β |
| <i>Photorhabdus luminescens</i> subsp. laumondii TTO1, amtB | GI:37524032 | (4529787..4531106) | N/A | MEP β |
| <i>Prochlorococcus marinus</i> str. MIT 9313, amt1 | GI:33862273 | 1960347..1961825 | N/A | grade |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|--------------------|--------------------|-------------|
| <i>Prochlorococcus marinus</i> subsp. <i>marinus</i> str. CCMP1375, <i>amtB</i> | GI:33239452 | 286257..287765 | N/A | grade |
| <i>Prochlorococcus marinus</i> , subsp. <i>pastoris</i> str. PCC 9511, <i>amt</i> | GI:6538781 | 252232-254193 | N/A | grade |
| <i>Pseudomonas aeruginosa</i> PA01, <i>amtB</i> | GI:110645304 | (5951115..5952443) | N/A | MEP β |
| <i>Pseudomonas aeruginosa</i> PA01, 2 | GI:9949144 | N/A | N/A | grade |
| <i>Pseudomonas fluorescens</i> PfO-1, 2 | GI:28880299 | 36..1295 | N/A | grade |
| <i>Pseudomonas fluorescens</i> PfO-1, 1 | GI:28880281 | 2..1336 | N/A | MEP β |
| <i>Pseudomonas fluorescens</i> PfO-1, 2 | GI:253992019 | (4493353..4494561) | N/A | grade |
| <i>Pseudomonas fluorescens</i> PfO-1, 1 | GI:253992019 | (6170031..6171368) | N/A | MEP β |
| <i>Pseudomonas putida</i> KT2440, <i>amtB</i> | GI:26986745 | (5966352..5967683) | N/A | MEP β |
| <i>Pseudomonas putida</i> KT2440, 2 | GI:26986745 | 2403371..2404579 | N/A | grade |
| <i>Pseudomonas stutzeri</i> , <i>amtB1</i> | GI:19033132 | 1848..3164 | N/A | MEP β |
| <i>Pseudomonas stutzeri</i> , <i>amtB2</i> | GI:19033132 | 3281..4534 | N/A | grade |
| <i>Pseudomonas syringae</i> pv. <i>syringae</i> B728a, 1 | GI:66043271 | (200642..201979) | N/A | MEP β |
| <i>Pseudomonas syringae</i> pv. <i>syringae</i> B728a, 2 | GI:66043271 | 2643672..2644997 | N/A | MEP β |
| <i>Pseudomonas syringae</i> pv. <i>tomato</i> str. DC3000, <i>amt-1</i> | GI:28867243 | 240728..242065 | N/A | MEP β |
| <i>Pseudomonas syringae</i> pv. <i>tomato</i> str. DC3000, <i>amt-2</i> | GI:28867243 | 2856588..2857913 | N/A | MEP β |

161

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|--------------------|--------------------|-------------|
| <i>Psychrobacter arcticus</i> 273-4, <i>amtB</i> | GI:71037566 | 2514104..2515366 | N/A | grade |
| <i>Ralstonia eutropha</i> JMP134, 1 | GI:72117119 | 324070..325611 | N/A | MEP β |
| <i>Ralstonia eutropha</i> JMP134, 2 | GI:72117119 | (680963..682167) | N/A | grade |
| <i>Ralstonia metallidurans</i> , <i>amtB</i> (= <i>Cupriavidus metallidurans</i> CH34) | GI:93352797 | 256874..258406 | N/A | MEP β |
| <i>Ralstonia metallidurans</i> , 2 (=Cupriavidus <i>metallidurans</i> CH34) | GI:288237308 | 1393967..1395169 | N/A | grade |
| <i>Ralstonia solanacearum</i> , <i>amtB</i> | GI:17544719 | 374899..376209 | N/A | MEP β |
| <i>Rhizobium etli</i> , <i>amtB</i> | GI:6599354 | 787..2211 | N/A | MEP β |
| <i>Rhodobacter sphaeroides</i> , <i>amtB</i> | GI:77461965 | (2636500..2637852) | N/A | MEP β |
| ¹⁴³ <i>Rhodopirullela baltica</i> SH1, <i>amtB</i> | GI:32397972 | (147355..148944) | N/A | MEP β |
| <i>Rhodopirullela baltica</i> SH1, 2 | GI:32397972 | 137662..139158 | N/A | grade |
| <i>Rhodopseudomonas palustris</i> , <i>amtB1</i> | GI:39652705 | 300666..301997 | N/A | MEP β |
| <i>Rhodopseudomonas palustris</i> , <i>amtB2</i> | GI:39652705 | 302683..304125 | N/A | MEP β |
| <i>Rhodospirillum rubrum</i> ATCC 11170, 2 | GI:83574254 | (655212..656531) | N/A | grade |
| <i>Rhodospirillum rubrum</i> ATCC 11170, 1 | GI:83574254 | (1330513..1331808) | N/A | MEP β |
| <i>Saccharophagus degradans</i> (=Microbulbifer <i>degradans</i>) <i>amtB1</i> | GI:89949249 | (2643236..2644531) | N/A | MEP β |
| <i>Saccharophagus degradans</i> (=Microbulbifer <i>degradans</i>) | GI:89949249 | 293080..294315 | N/A | grade |
| <i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhi, <i>amtB</i> | GI:16501496 | 262755..264041 | N/A | MEP β |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|--------------------|--------------------|--------------|
| <i>Shewanella oneidensis</i> MR-1, <i>amt</i> | GI:24371600 | (775008..776234) | N/A | grade |
| <i>Shigella flexneri</i> 2a str. 301, <i>amtB</i> | GI:24111450 | 408059..409345 | N/A | MEP β |
| <i>Sinorhizobium meliloti</i> , <i>amtB</i> | GI:15963753 | 3471607..3472962 | N/A | MEP β |
| <i>Staphylococcus aureus</i> subsp. <i>aureus</i> Mu50, <i>nrgA</i> | GI:57634611 | (2162727..2163977) | GI:83574254 | MEP β |
| <i>Staphylococcus aureus</i> subsp. <i>aureus</i> MW2, <i>nrgA</i> | GI:47118312 | (2114899..2116149) | N/A | MEP β |
| <i>Staphylococcus epidermidis</i> ATCC 12228 | GI:27466918 | (1696434..1697684) | N/A | MEP β |
| <i>Streptococcus mutans</i> UA159, <i>nrgA</i> | GI:24378532 | 1575845..1577080) | N/A | MEP β |
| ††1 <i>Streptomyces avermitilis</i> MA-4680, <i>amtB1</i> | GI:162960844 | (3248960..3250300) | N/A | MEP β |
| <i>Streptomyces avermitilis</i> MA-4680, <i>amtB2</i> | GI:162960844 | (6766506..6767816) | N/A | MEP β |
| <i>Streptomyces coelicolor</i> | GI:24413886 | 105031..106377 | N/A | MEP β |
| <i>Sulfurihydrogenibium</i> sp. YO3AOP1 | GI:188931022 | (55106..56395) | N/A | MEP α |
| <i>Synechococcus elongatus</i> PCC 7942, <i>amt1%</i> | GI:81167692 | 431458..432951 | N/A | grade |
| <i>Synechococcus elongatus</i> PCC 7942, 1 | GI:81167692 | 2347614..2349026 | N/A | MEP β |
| <i>Synechococcus</i> sp. PCC 7002, <i>amtA</i> | GI:170076636 | 1096636..1098366 | N/A | grade |
| <i>Synechococcus</i> sp. strain WH8102, <i>amt1</i> | GI:33632062 | (253620..255092) | N/A | grade |
| <i>Synechococcus</i> sp. WH 8102, <i>amt1</i> | GI:33632062 | (253620..255092) | N/A | grade |
| <i>Synechocystis</i> sp. PCC 6803, <i>amt1</i> | GI:16329170 | (2969677..2971200) | N/A | grade |
| <i>Synechocystis</i> sp. PCC 6803, <i>amt2</i> | GI:16329170 | (400720..402048) | N/A | grade |
| <i>Synechocystis</i> sp. PCC 6803, <i>amt3</i> | GI:16329170 | (3241089..3242714) | N/A | grade |
| <i>Thermobifida fusca</i> , <i>amtB</i> | GI:72160406 | 764337..765656 | N/A | MEP β |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|--------------------|--------------------|--------------|
| <i>Thermosynechococcus elongatus</i> BP-1, 1 | GI:22297544 | (169488..171011) | N/A | grade |
| <i>Thermosynechococcus elongatus</i> BP-1, 2 | GI:22297544 | 1710518..1711210 | N/A | grade |
| <i>Thermotoga maritima</i> | GI:15642775 | 422169..423476 | N/A | grade |
| <i>Trichodesmium erythraeum</i> IMS101, 1 | GI:113473942 | (2730734..2732224) | N/A | grade |
| <i>Trichodesmium erythraeum</i> IMS101, 2 | GI:113473942 | 6910533..6911999 | N/A | grade |
| <i>Ureaplasma urealyticum</i> , amt-1 | GI:13357558 | 260648..262180 | N/A | grade |
| <i>Ureaplasma urealyticum</i> , amt-2 | GI:13357558 | 262345..263910 | N/A | grade |
| <i>Vibrio parahaemolyticus</i> RIMD 2210633, 2 | GI:28896774 | 1634487..1635770 | N/A | grade |
| <i>Vibrio parahaemolyticus</i> RIMD 2210633, 1 | GI:28896774 | (2628155..2629384) | N/A | grade |
| <i>Vibrio vulnificus</i> CMCP6, 1 | GI:117956319 | 1631042..1632271 | N/A | grade |
| <i>Vibrio vulnificus</i> CMCP6, 2 | GI:117956319 | 2637559..2638779 | N/A | grade |
| <i>Vibrio vulnificus</i> YJ016, 2 | GI:37678184 | (1744530..1745750) | N/A | grade |
| <i>Vibrio vulnificus</i> YJ016, 1 | GI:37678184 | (2795993..2797237) | N/A | grade |
| <i>Wolinella succinogenes</i> , amtB | GI:34482500 | (51759..53072) | N/A | MEP α |
| <i>Xanthomonas campestris</i> pv. <i>campestris</i> str. ATCC 33913, amtB | GI:21166373 | 243559..245034 | N/A | MEP β |
| <i>Xanthomonas citri</i> , amtB | GI:5917780 | N/A | N/A | MEP β |
| <i>Xylella fastidiosa</i> 9a5c | GI:57014152 | 1759686..1761164 | N/A | MEP β |
| <i>Xylella fastidiosa</i> Dixon | GI:22994217 | 20672..22057 | N/A | MEP β |
| <i>Xylella fastidiosa</i> Temecula1, amtB | GI:28197945 | (1222493..1223947) | N/A | MEP β |
| <i>Yersinia pestis</i> CO92, amtB | GI:16120353 | (3501231..3502526) | N/A | MEP β |

Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|----------------|---------------------------------|-------|
| OUTGROUP | | | | |
| EUKARYOTES | | | | |
| Green algae | | | | |
| <i>Chlamydomonas reinhardtii</i> , RHP1 | N/A | N/A | GI:159474701 ref XM_001695412.1 | Rh |
| <i>Chlamydomonas reinhardtii</i> , RHP2 | N/A | N/A | GI:159474703 ref XM_001695413.1 | Rh |
| Animals | | | | |
| ¹⁴ <i>Drosophila melanogaster</i> , Rh50 | N/A | N/A | GI:221379345 | Rh |
| <i>Homo sapiens</i> , RHAG % | N/A | N/A | GI:156627564 | Rh |
| <i>Homo sapiens</i> , RHBG | N/A | N/A | GI:224548944 | Rh |
| <i>Homo sapiens</i> , RHCG | N/A | N/A | GI:7706682 | Rh |
| <i>Homo sapiens</i> , RHD | N/A | N/A | GI:77157792 | Rh |
| Slime molds (Amoebozoa) | | | | |
| <i>Dictyostelium discoideum</i> , rhgA | GI:90970616 | (30188..31863) | GI:66814323 | Rh |
| <i>Dictyostelium discoideum</i> , rhgB | GI:90970667 | 505388..50719 | N/A | Rh |
| Excavates | | | | |
| Heterolobosea | | | | |
| <i>Naegleria gruberi</i> , 3 ^l | N/A | N/A | GI:290995577 | Rh |

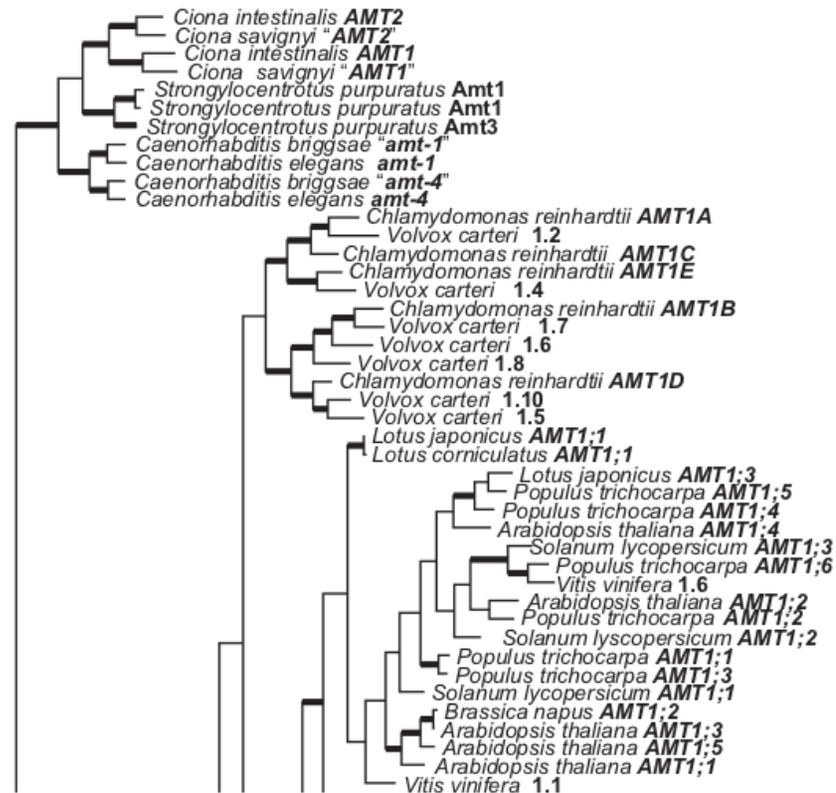
Table 10 (continued): Accession numbers and genome coordinates of ammonium transporter/ammonia permease genes of sequences included in the phylogenetic tree

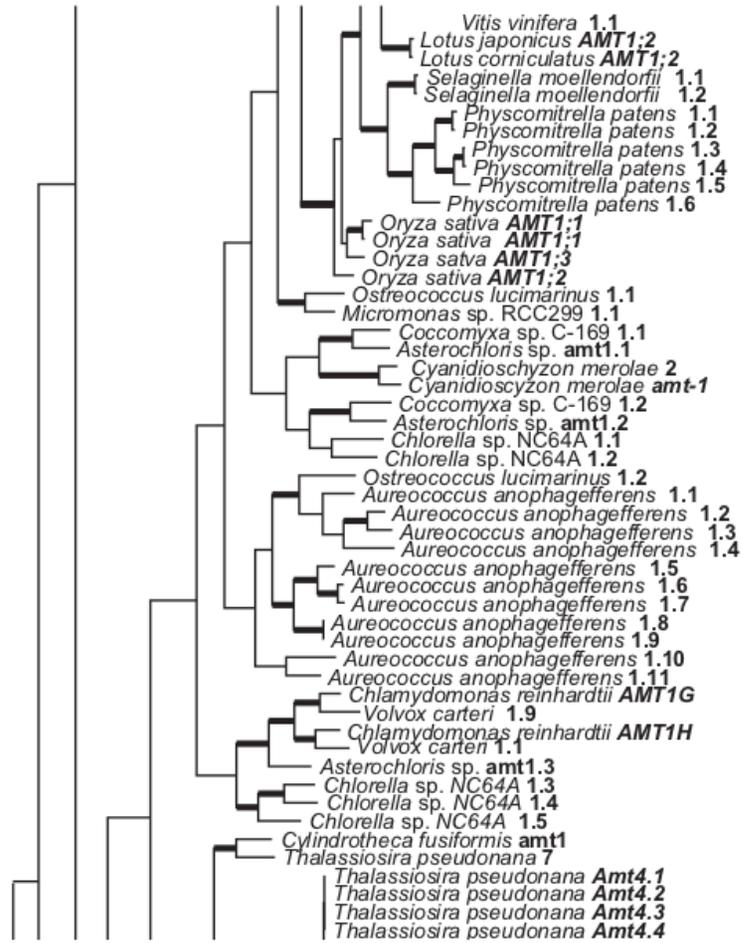
| Organism, gene abbreviation | GI number or GenBank accession number for DNA | Genome Region | GI number for mRNA | Clade |
|---|---|-------------------|-------------------------------|-------|
| PROKARYOTES | | | | |
| Bacteria | | | | |
| <i>Geobacter</i> sp. M21 | GI:253698656 | (1678563-1679774) | N/A | |
| <i>Geobacter</i> sp. M21* | GI:253698656 | (1329373-1328153) | N/A | Rh |
| <i>Nitrosomonas europaea</i> | GI:38570258 | N/A | GI:38570258 gb AY37792 3.1 | Rh |
| <i>Nitrospira multiformis</i> ATCC 25196 | GI:82409200 | (585829..587052) | N/A | Rh |
| <i>Strongylocentrotus purpuratus</i> , Rh50 | GI:115959933 | (29430..84508) | GI:115733129 | Rh |

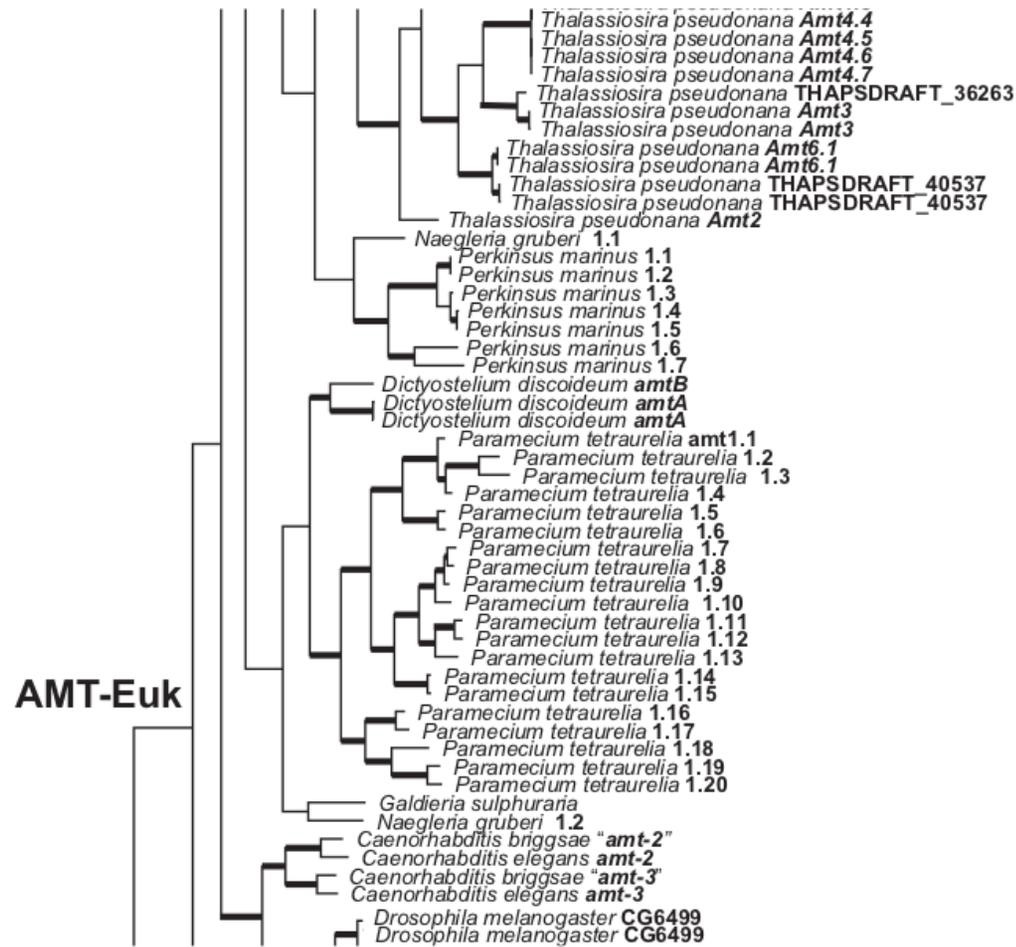
/FI

%= duplicate sequence in the analysis, * = excluded or omitted from the analysis, ^J = data from <http://www.jgi.doe.gov/>; ^B=data from <http://www.broadinstitute.org/>; ^M=data from <http://genomics.msu.edu/galdieria/>; ^T=data from <http://merolae.biol.s.u-tokyo.ac.jp/>.

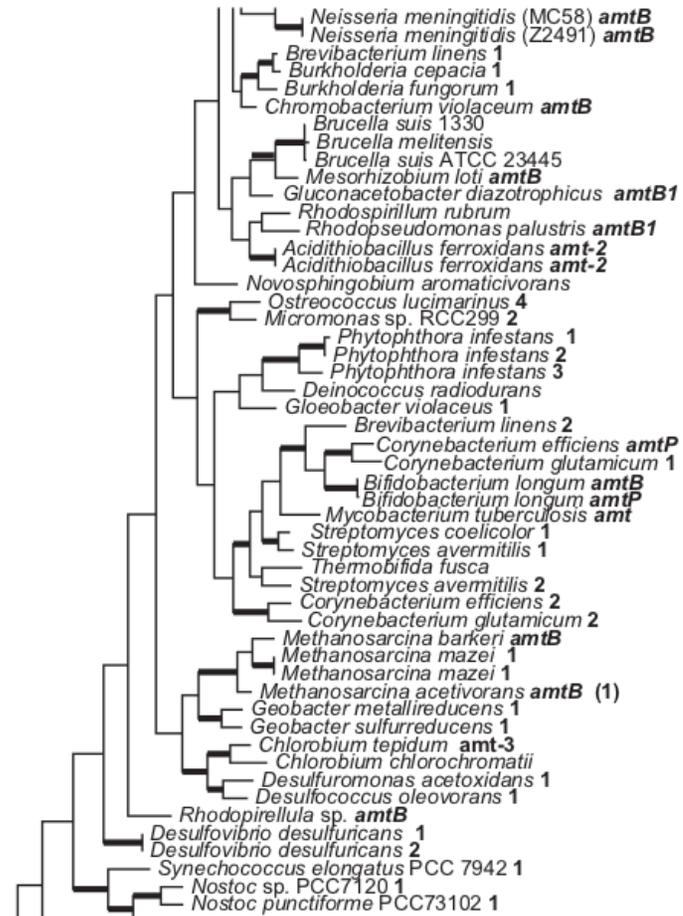
Appendix B: Evolution of ammonium transporters/ammonia permeases. Maximum likelihood analysis of 513 ammonium transporter/ammonia permease genes shows the presence of a eukaryotic clade (AMT-Euk) representing vertical transmission of AMTs into eukaryotes and a well-supported predominantly prokaryotic clade (MEP) in which eukaryotic lineages demonstrate horizontal gene transfer. AMT-Euk=eukaryotic ammonium transporters clade; MEP=methylammonium permease clade; MEP grade= prokaryotic ammonium transporter/ammonia permease grade; Rh=Rhesus family (outgroup). Thickened branches show bootstrap support of $\geq 70\%$

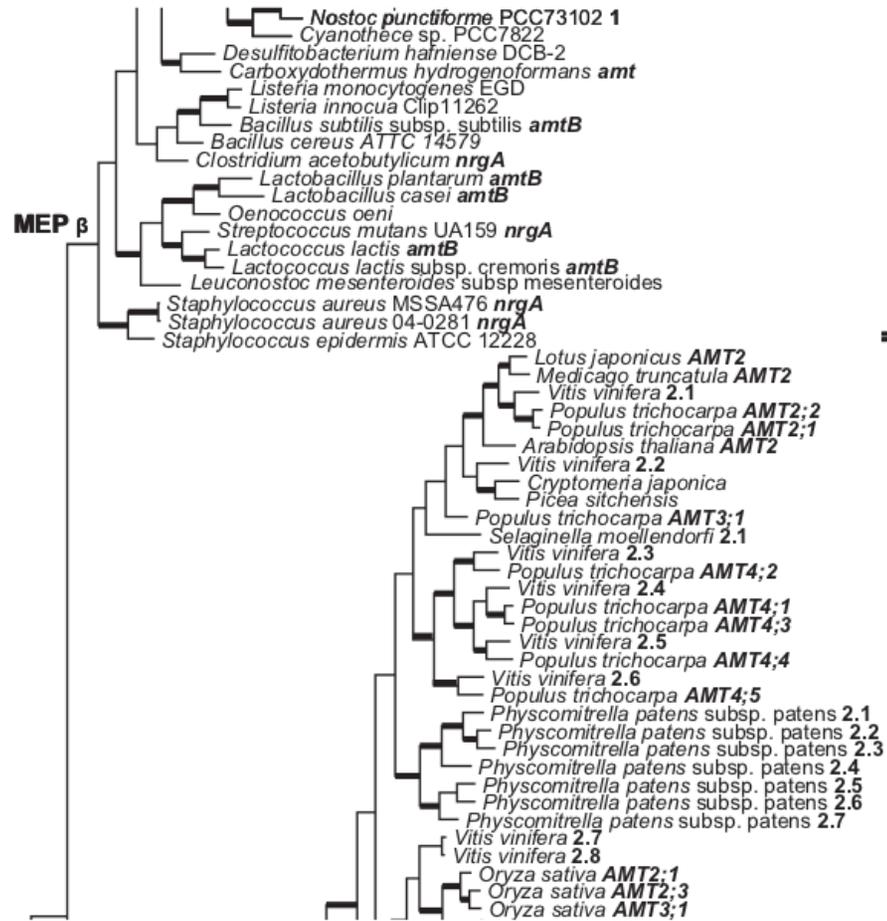


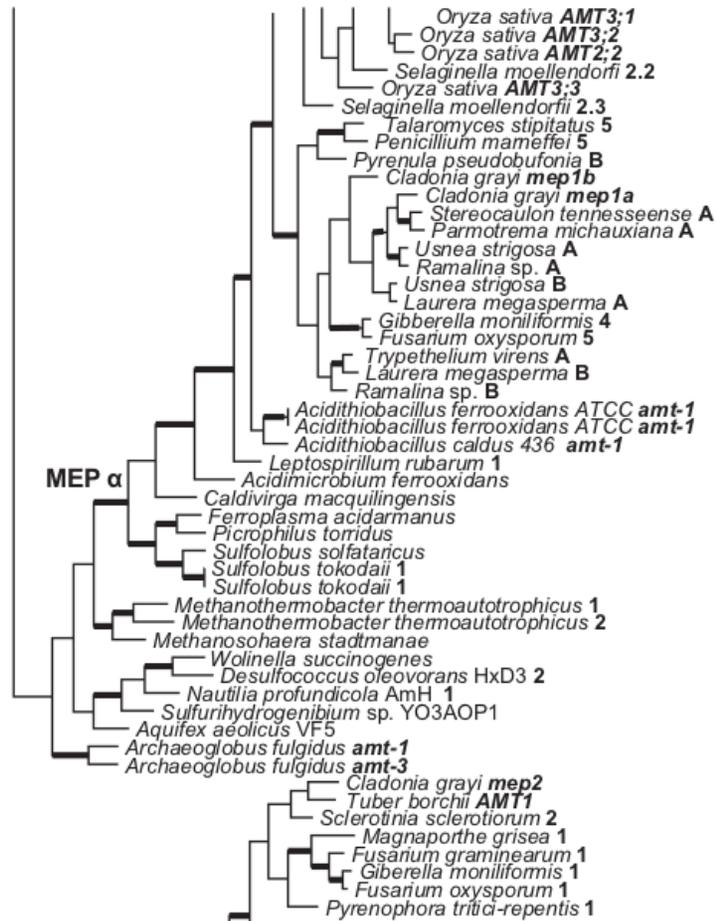


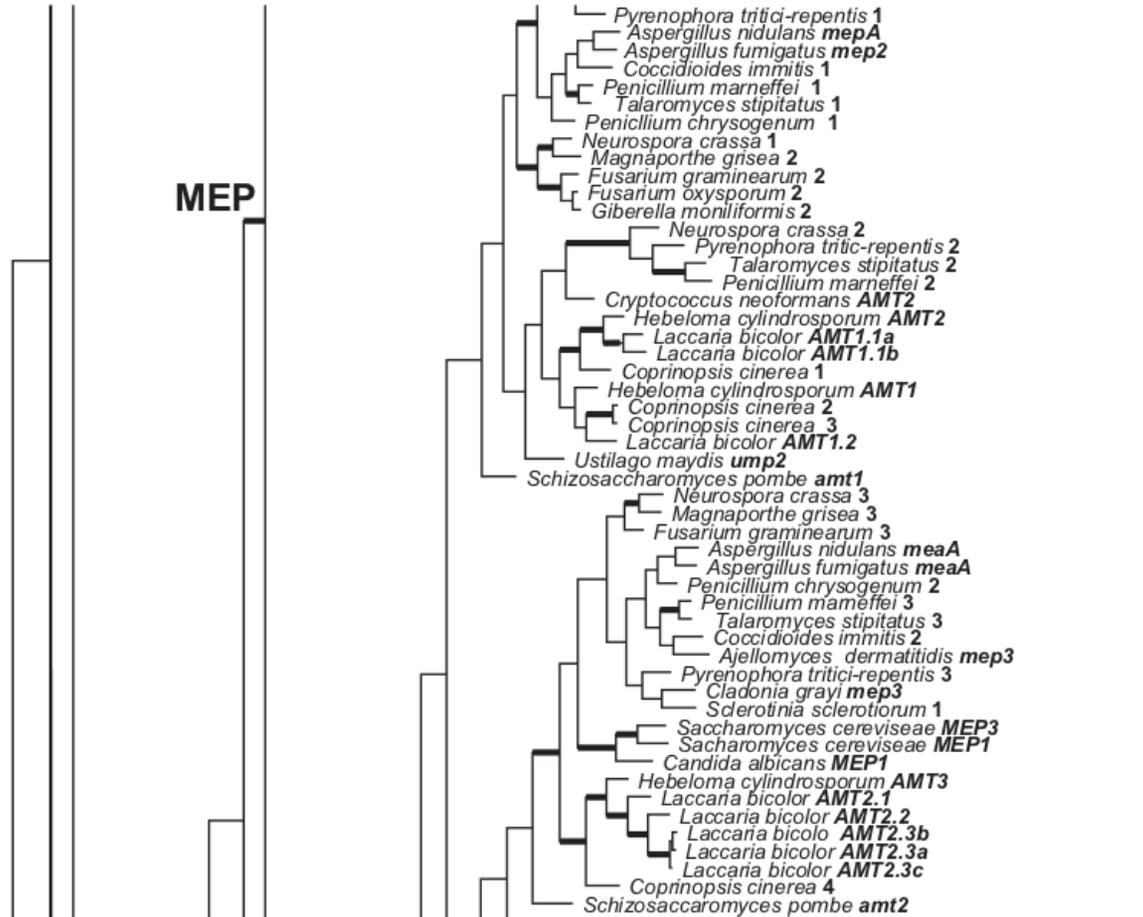






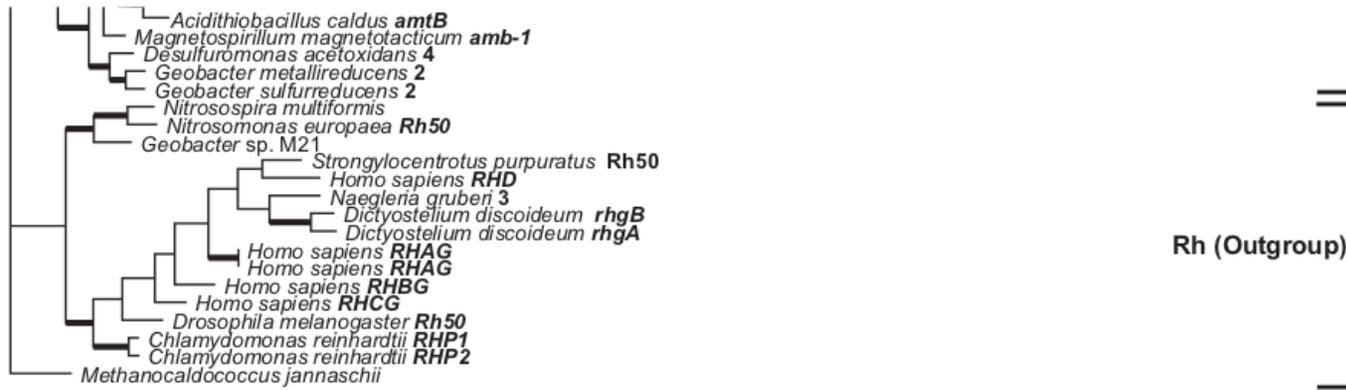








MEP grade



— 0.1 substitutions/site

Appendix C. Summary of assembly results for eight lichen genomes.

Six assemblies were performed on each genome, varying the proportion of the sequencing data used or the k-mer size. The short reads assembly program Velvet was used requiring a minimum overlap of either 27 basepairs or 31 basepairs to produce contigs (k-mer size), and using either all the reads, the first 100 million reads or the first 50 million reads. N/D – not done; bp = basepair.

Table 11: Summary of assembly results for eight lichen genomes

| <i>Acarospora strigata</i> | 27 bp k-mer size | | | 31 bp k-mer size | | |
|----------------------------|------------------|---------|--------|------------------|---------|--------|
| | All | Top 100 | Top 50 | All | Top 100 | Top 50 |
| #contigs | | | | | | |
| >5000 bp | 21 | 1037 | 1104 | N/D | 1192 | 945 |
| 2000 - 4999 | 1693 | 2874 | 1865 | N/D | 1819 | 1302 |
| 1000 - 1999 | 5966 | 2384 | 1353 | N/D | 1282 | 1025 |
| 500 - 999 | 8544 | 2433 | 1346 | N/D | 1414 | 1077 |
| 200 - 499 | 11026 | 3499 | 2265 | N/D | 2502 | 2215 |
| 100 - 199 | 8439 | 3812 | 3355 | N/D | 3738 | 3730 |
| 0 - 99 | 97145 | 39340 | 26494 | N/D | 24026 | 18569 |
| Total contigs | 132834 | 55379 | 37782 | N/D | 35973 | 28863 |
| Longest contig | 7696 | 23431 | 40528 | N/D | 40628 | 55327 |

| <i>Arthonia cf rubrocincta</i> | 27 bp k-mer size | | | 31 bp k-mer size | | |
|--------------------------------|------------------|---------|--------|------------------|---------|--------|
| | All | Top 100 | Top 50 | All | Top 100 | Top 50 |
| #contigs | | | | | | |
| >5000 bp | 0 | 974 | N/D | 3 | 1169 | N/D |
| 2000 - 4999 | 40 | 3222 | N/D | 802 | 2024 | N/D |
| 1000 - 1999 | 1771 | 2665 | N/D | 5160 | 1313 | N/D |
| 500 - 999 | 10122 | 2504 | N/D | 10777 | 1416 | N/D |
| 200 - 499 | 28448 | 3824 | N/D | 17382 | 2677 | N/D |
| 100 - 199 | 24530 | 4565 | N/D | 13004 | 4393 | N/D |
| 0 - 99 | 181094 | 34256 | N/D | 83984 | 18755 | N/D |
| Total contigs | 246005 | 52010 | N/D | 131112 | 31747 | N/D |
| Longest contig | 24602 | 20086 | N/D | 6169 | 43566 | N/D |

Table 11 (continued): Summary of assembly results for eight lichen genomes

| <i>Dibaeis baeomyces</i> | 27 bp k-mer size | | | 31 bp k-mer size | | |
|--------------------------|------------------|---------|--------|------------------|---------|--------|
| #contigs | All | Top 100 | Top 50 | All | Top 100 | Top 50 |
| >5000 bp | 14 | 918 | 1325 | 15 | 1158 | 1438 |
| 2000 - 4999 | 81 | 3680 | 3466 | 199 | 3460 | 3024 |
| 1000 - 1999 | 508 | 4389 | 3230 | 2660 | 8274 | 6722 |
| 500 - 999 | 6773 | 5669 | 3379 | 11932 | 4820 | 2360 |
| 200 - 499 | 46468 | 11097 | 5175 | 41100 | 8801 | 3915 |
| 100 - 199 | 103693 | 15648 | 7054 | 83402 | 14459 | 6633 |
| 0 - 99 | 585504 | 77779 | 43951 | 367592 | 51845 | 28844 |
| Total contigs | 743041 | 119180 | 67580 | 506900 | 92817 | 52936 |
| Longest contig | 11416 | 32275 | 44235 | 15040 | 32271 | 58795 |

| <i>Endocarpon pallidulum</i> | 27 bp k-mer size | | | 31 bp k-mer size | | |
|------------------------------|------------------|---------|--------|------------------|---------|--------|
| #contigs | All | Top 100 | Top 50 | All | Top 100 | Top 50 |
| >5000 bp | 3 | 1244 | 1252 | 148 | 1276 | 1172 |
| 2000 - 4999 | 1218 | 3216 | 2682 | 3400 | 2445 | 2167 |
| 1000 - 1999 | 6557 | 3122 | 2350 | 6769 | 2326 | 1639 |
| 500 - 999 | 12423 | 3493 | 2354 | 8157 | 2247 | 2163 |
| 200 - 499 | 19732 | 7545 | 6437 | 14473 | 9482 | 9109 |
| 100 - 199 | 21648 | 17584 | 17120 | 24901 | 24176 | 23673 |
| 0 - 99 | 354446 | 163539 | 140443 | 227203 | 122608 | 106125 |
| Total contigs | 416027 | 199743 | 172638 | 285051 | 164560 | 146048 |
| Longest contig | 5652 | 25014 | 39526 | 8940 | 40911 | 54863 |

| <i>Graphis scripta</i> | 27 bp k-mer size | | | 31 bp k-mer size | | |
|------------------------|------------------|---------|--------|------------------|---------|--------|
| #contigs | All | Top 100 | Top 50 | All | Top 100 | Top 50 |
| >5000 bp | 2 | 1362 | 1415 | 36 | 1497 | 1144 |
| 2000 - 4999 | 538 | 3733 | 2543 | 2484 | 2410 | 1557 |
| 1000 - 1999 | 5282 | 3477 | 2167 | 7680 | 2153 | 1447 |
| 500 - 999 | 14883 | 4081 | 2448 | 11306 | 2687 | 1982 |
| 200 - 499 | 27888 | 6390 | 4637 | 17237 | 4680 | 4120 |
| 100 - 199 | 22618 | 6737 | 5984 | 13731 | 5052 | 4989 |
| 0 - 99 | 172232 | 40079 | 27922 | 82120 | 21757 | 16833 |
| Total contigs | 243443 | 65859 | 47116 | 134594 | 40236 | 32072 |
| Longest contig | 5723 | 26289 | 52836 | 10337 | 38671 | 66205 |

Table 11 (continued): Summary of assembly results for eight lichen genomes

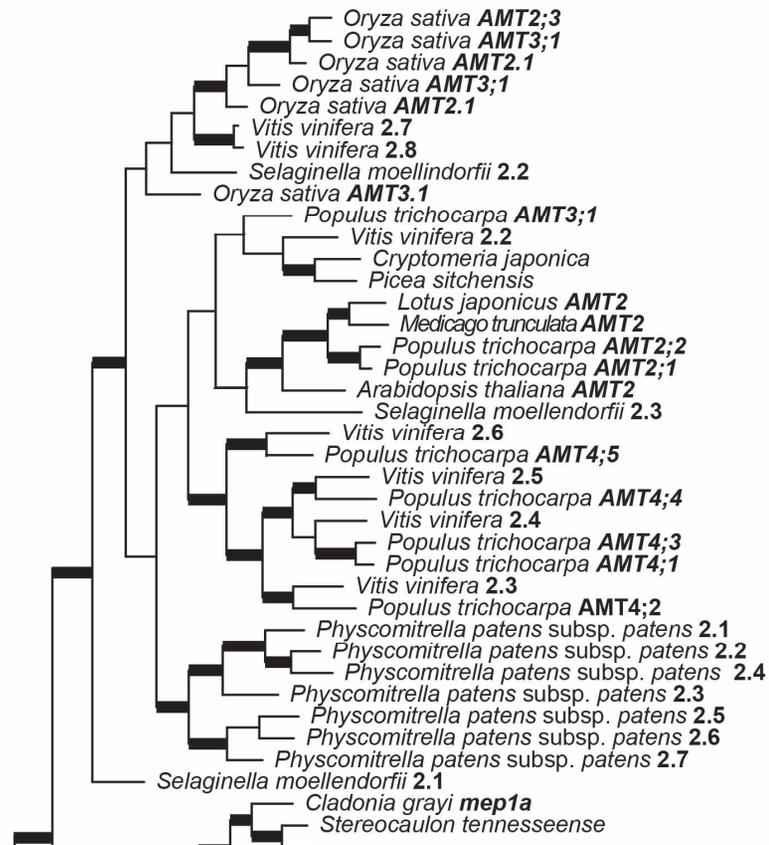
| <i>Leptogium</i> sp. Part 1 | 27 bp k-mer size | | | 31 bp k-mer size | | |
|-----------------------------|------------------|---------|--------|------------------|---------|--------|
| #contigs | All | Top 100 | Top 50 | All | Top 100 | Top 50 |
| >5000 bp | 703 | 1265 | 170 | N/D | 1388 | 270 |
| 2000 - 4999 | 3963 | 4005 | 2304 | N/D | 4150 | 1803 |
| 1000 - 1999 | 6004 | 4336 | 8044 | N/D | 4101 | 6829 |
| 500 - 999 | 14205 | 5175 | 14465 | N/D | 4604 | 15320 |
| 200 - 499 | 74340 | 16146 | 23374 | N/D | 14327 | 27969 |
| 100 - 199 | 290332 | 60715 | 36174 | N/D | 52834 | 37823 |
| 0 - 99 | 1270982 | 309678 | 201483 | N/D | 221728 | 153248 |
| Total contigs | 1660529 | 401320 | 286014 | N/D | 303132 | 243262 |
| Longest contig | 17820 | 29786 | 20013 | N/D | 54209 | 22817 |

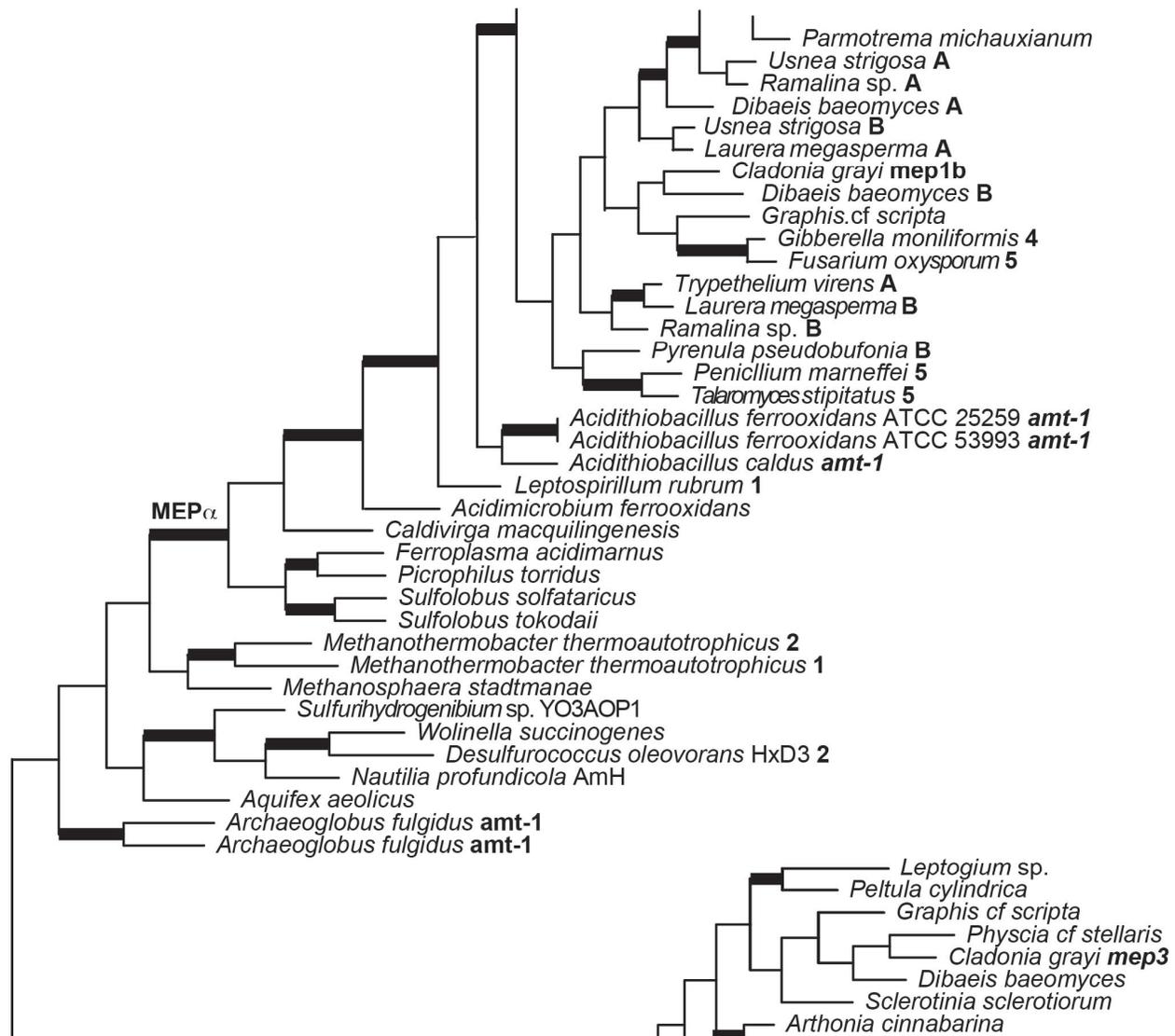
| <i>Peltula cylindrica</i> | 27 bp k-mer size | | | 31 bp k-mer size | | |
|---------------------------|------------------|---------|--------|------------------|---------|--------|
| #contigs | All | Top 100 | Top 50 | All | Top 100 | Top 50 |
| >5000 bp | 272 | 329 | 538 | N/D | 553 | 740 |
| 2000 - 4999 | 2593 | 3066 | 3430 | N/D | 3372 | 3617 |
| 1000 - 1999 | 9287 | 5283 | 4901 | N/D | 4706 | 4042 |
| 500 - 999 | 29305 | 9239 | 5758 | N/D | 6608 | 4357 |
| 200 - 499 | 126863 | 37786 | 12773 | N/D | 29512 | 9493 |
| 100 - 199 | 333273 | 82385 | 41397 | N/D | 75541 | 33825 |
| 0 - 99 | 1322311 | 243259 | 149995 | N/D | 182105 | 112502 |
| Total contigs | 1823904 | 381347 | 218792 | N/D | 302397 | 168576 |
| Longest contig | 16649 | 14530 | 17798 | N/D | 27360 | 27804 |

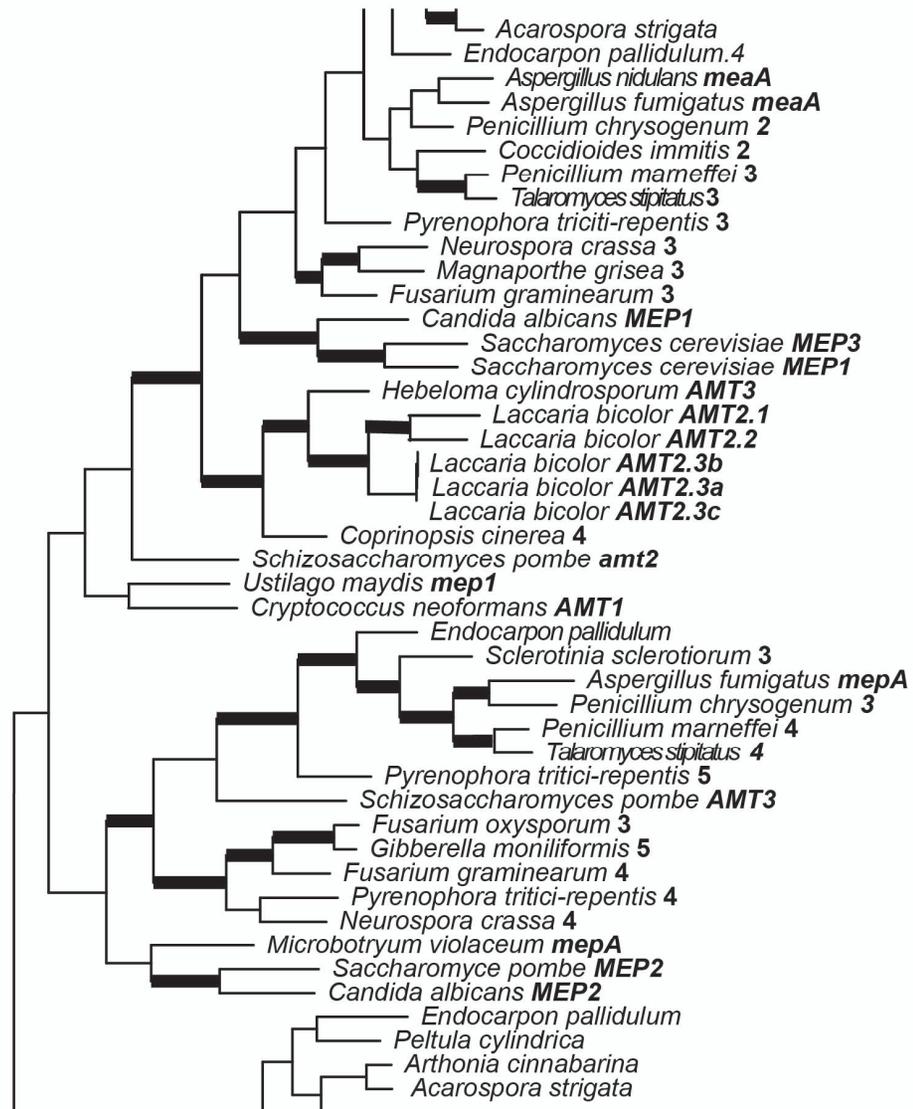
| <i>Physcia cf stellaris</i> | 27 bp k-mer size | | | 31 bp k-mer size | | |
|-----------------------------|------------------|---------|--------|------------------|---------|---------|
| #contigs | All | Top 100 | Top 50 | All | Top 100 | Top 50 |
| >5000 bp | 24 | 4 | 59 | N/D | 68 | 68 |
| 2000 - 4999 | 565 | 724 | 985 | N/D | 892 | 983 |
| 1000 - 1999 | 6411 | 2637 | 3253 | N/D | 3048 | 3534 |
| 500 - 999 | 33342 | 8435 | 9540 | N/D | 9729 | 10521 |
| 200 - 499 | 168822 | 40739 | 37582 | N/D | 44367 | 39322 |
| 100 - 199 | 410939 | 126877 | 77519 | N/D | 125194 | 777478 |
| 0 - 99 | 2756283 | 711884 | 421492 | N/D | 348334 | 376651 |
| Total contigs | 3376386 | 891300 | 550430 | N/D | 531632 | 1208557 |
| Longest contig | 12928 | 11546 | 19301 | N/D | 16789 | 19733 |

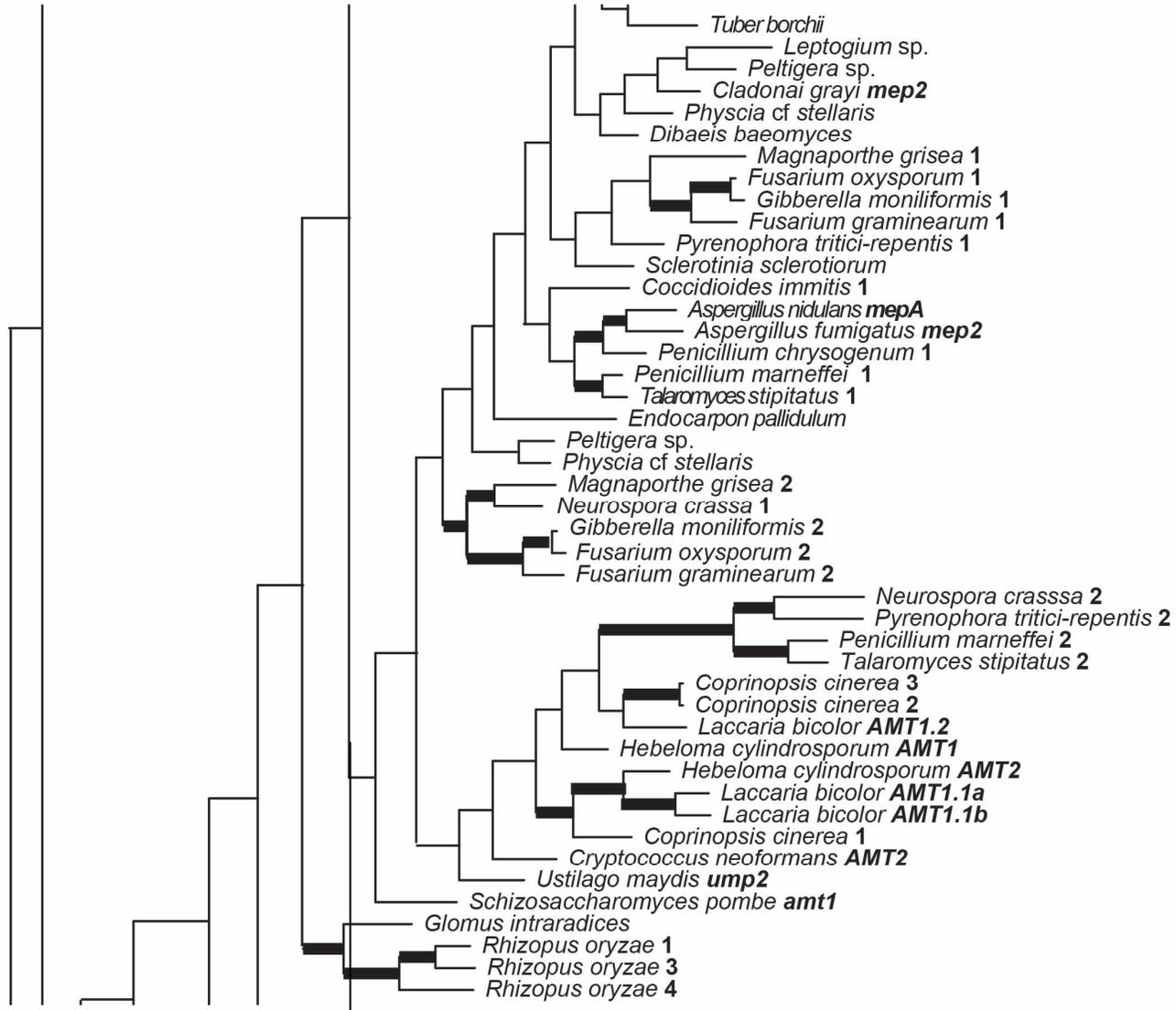
Appendix D. Phylogenetic placement of ammonium transporters/ammonia permeases from eight lichen genomes.

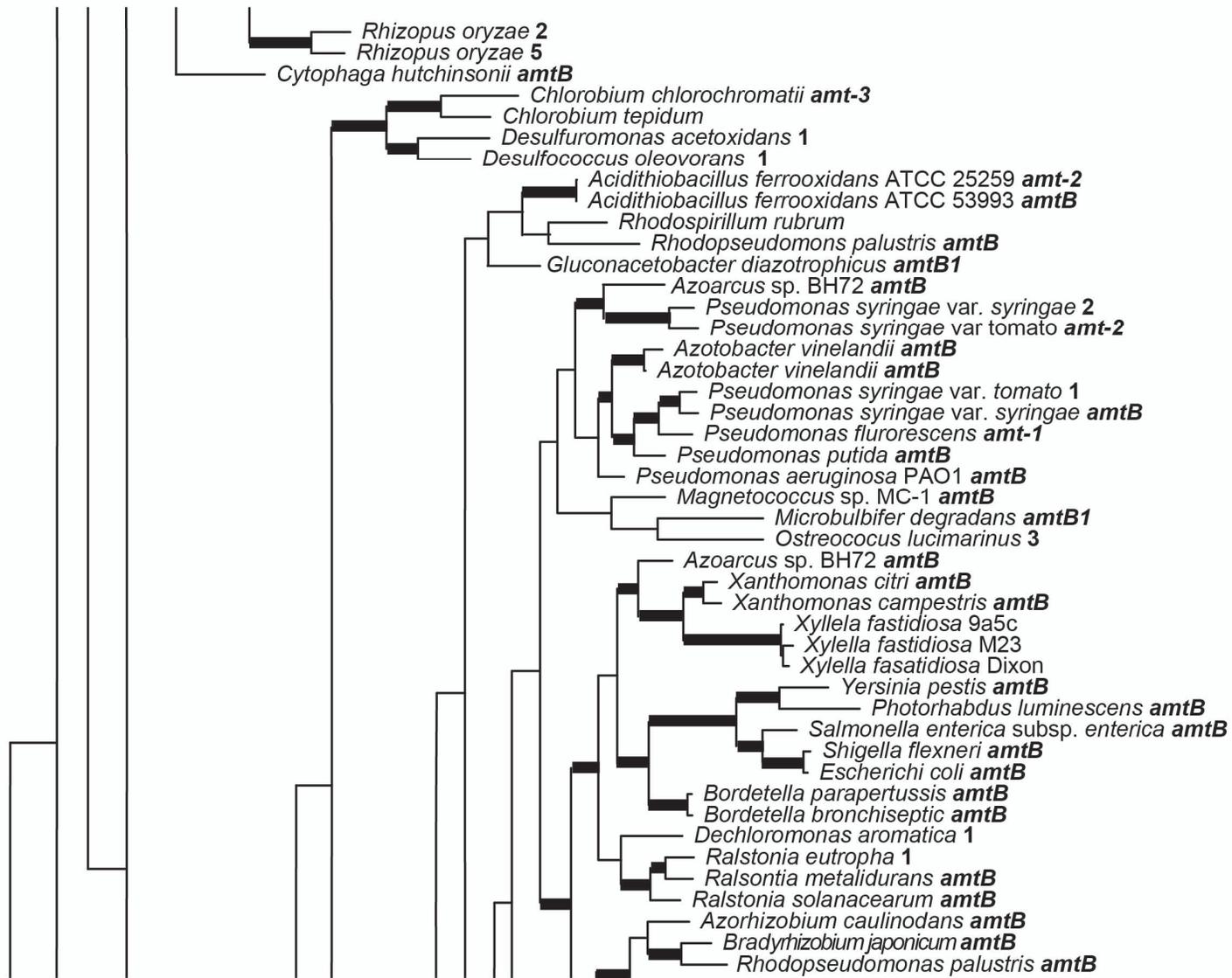
Maximum likelihood analysis of 300 ammonium transporter/ammonia permease genes details the phylogenetic placement of ammonium transporter/ammonia permease genes in the well-supported predominantly prokaryotic clade (MEP) in which eukaryotic lineages demonstrate horizontal gene transfer. MEP=methylammonium permease clade; MEP grade = prokaryotic ammonium transporter/ammonia. Thickened branches show bootstrap support of $\geq 70\%$.

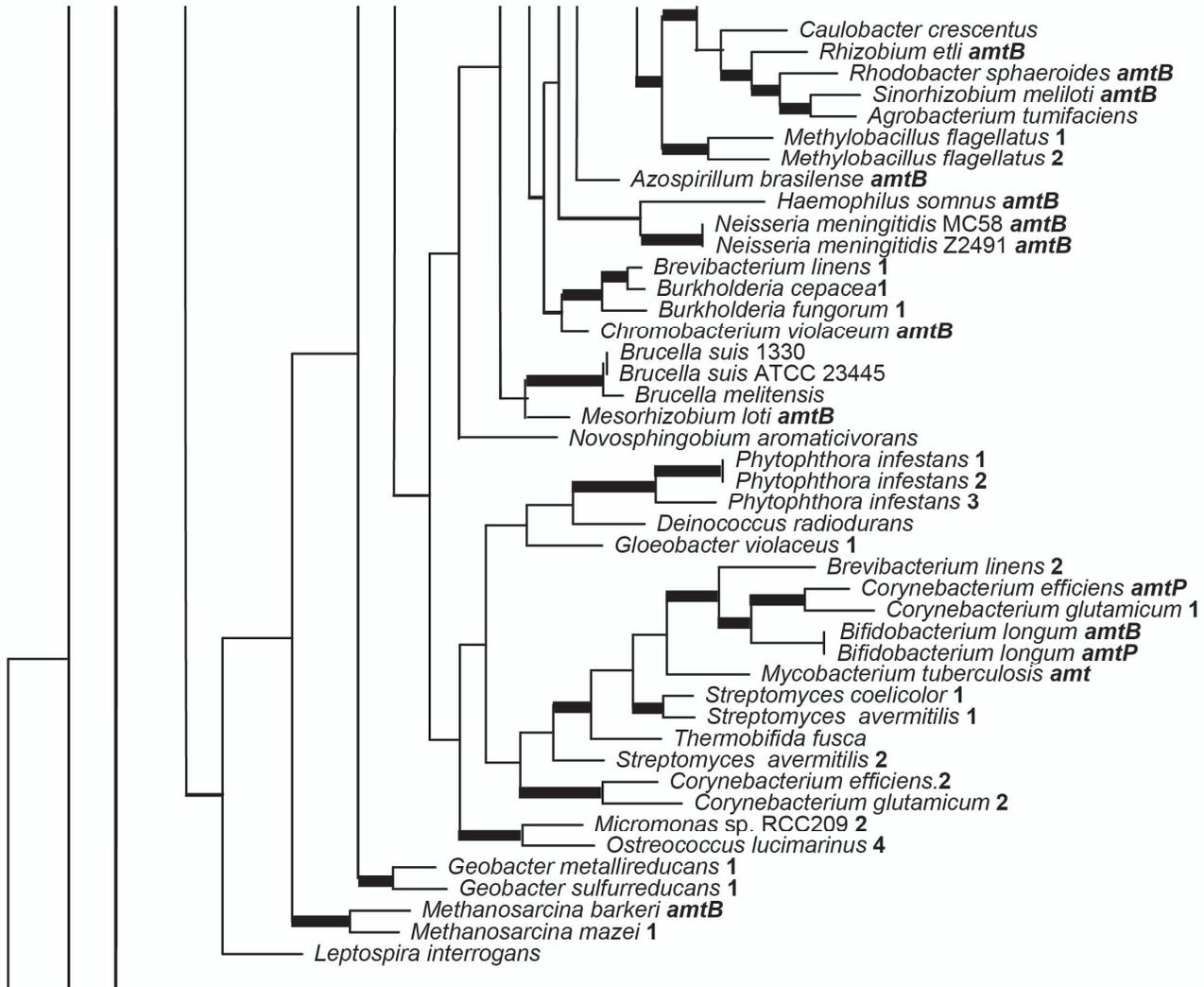


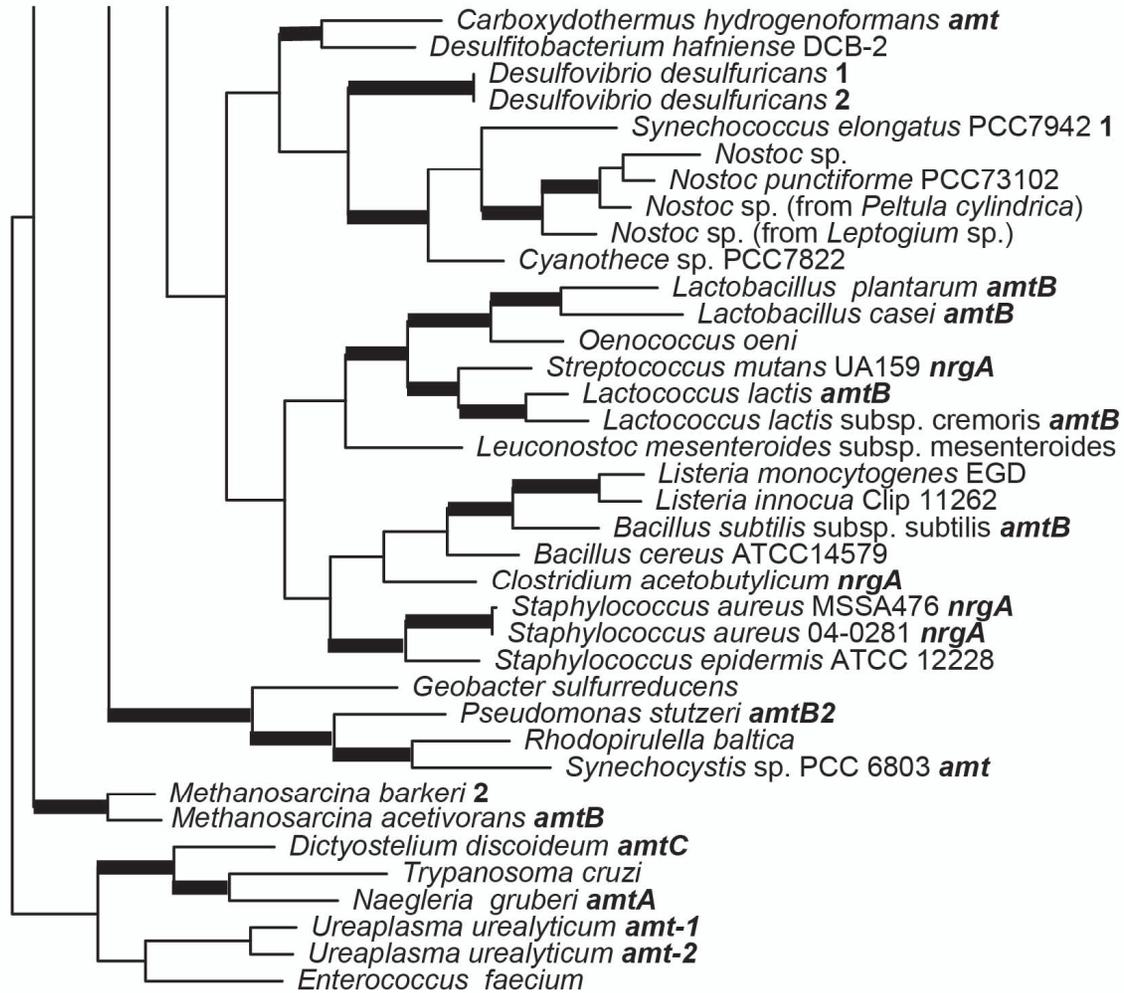












Appendix E. Primers used in Chapter 3

Table 12: Primers used in Chapter 3

| Primer name | Sequence (5' – 3') |
|---|--------------------------------|
| DNA methyltransferase primers | |
| Dim2_19_3_RaceNested | AGCCTAGAGCAAGGGCGACTGACCTAT |
| Dim2_3'_Xba | CTCTCTAGACCTCCTACTCAGCCTTTT |
| Dim2_3RNA_nested | CGTACATCCTGTAATCCCCCATCCAGAC |
| Dim2_5'_Xba | GATTCTAGATTCTTATAGCCCTCGAC |
| Dim2_end_R | AGTGCCATGTTACCTATGCGGTCAGAAG |
| H3K9 histone methyltransferase primers | |
| CHSH_ylxwkng | TAYCTNIIITGGAAIAAYGG |
| CHSHrev_CGY | TCRTARAARCGIAGCATYTT |
| CHSHrev_mrr | TCRTARAAYYKIAGCATYTT |
| Cwnrvv-AGY | TGYTGGAAYGRGTIGT |
| Cwnrvv-CTI | TGYTGGAAYCGIGTIGT |
| DLAFFA-CTI_Rev | GCRAARAAGCIAGRTC |
| DLAFFA-TTR_Rev | GCRAARAAGCYAARTC |
| ELTFDY-AGY | TARTCRAAIGTYAAYTC |
| ELTFDY-CTI | TARTCRAAIGTIAGYTC |
| H3K9_209 | TAGGCATCGTCACAGCTAACCTCAACG |
| H3K9_3'UTR_2R | AAACGATACCAGGCGTTTTG |
| H3K9_3'UTR_F | AAAGTCCCCTCAACCACGCACTACTCA |
| H3K9_368 | GCTGACAGTGTGATTGGGTTTAAGG |
| H3K9_438 | TCATTAGGCTCCAGAAAGCGCAACTTC |
| H3K9_481 | GTGATGATGAGGAGCCAGTCAC |
| H3K9_580 | GACAATCAAGAAATCCAGGTAAGAT |
| H3K9_613 | TATCATCCAAAAATTCAGCGTCTA |
| H3K9_742 | CCCAGAAAAGCGTCTAAACCGTAAC |
| H3K9_834 | CACCAGCTGAGTCTCGACGTATGTTCC |
| H3K9_GEIITS_F | GAGAGATTAATCACGAGCGAAGAAGCTGA |
| H3K9_gsp4 | CGAAGATGAGGTCGTTGAGGTTAGCTG |
| H3K9_gsp5 | TGGGTGTGAAATACCTTCAAAGCGAAG |
| H3K9_gsp5F | CTTCGCTTTGAAGGTATTTACACCCA |
| H3K9_gsp6 | ATTTCCAAACCGCTATCGCGCTTATCC |
| H3K9-DLaf | CTACGACCTGGCGTTCTTTGCTACCGAGCC |

Table 12 (continued): Primers used in Chapter 3

| Primer name | Sequence (5' – 3') |
|---|--------------------------------|
| H3K9 histone methyltransferase primers (continued) | |
| H3K9DLArev | GGCTCGGTAGCAAAGAACGCCAGGTCGTAG |
| H3K9RFVNrev | CAATTCGGCTCACACGAGTGATTCACGAAC |
| Preset_Eurot | TGCCDAACCGHGYSGTHGA |
| PreSETC | TGCTGGAATCGMGTCGTSCA |
| RFXNHSC-AGY | CGITTYITIAAYCAYAGYTG |
| RFXNHSC-TCI | CGITTYITIAAYCAYTCITG |
| Set2_new2 | GCAGTTGGGGGTGCAGGAGTGGTTCATGAA |
| Set2F | TTCATSAACCACTCSTGCAACCCCAACTGC |
| HP1 primers | |
| CYEKCPQK | CKNGGRCAYTITTCRTARCA |
| HP1_3flank_F | TCCAGTTCAAACCTCTTAAACGCAGATG |
| HP1_3flank_F | GACGTGAACTTCGAGAGCTATGAAGATGGT |
| HP1_5flank | AACCTTGTCGCTGATTCGATCAAGCTC |
| HP1_F1 | ATACAGTCGTTTCGTGAAAACGCCTGACG |
| HP1_F2 | CTATCGCACAAAAGCGAATCTCCCAAC |
| HP1_R1 | GTTCCCGCGCTCCACTATGATATCTGTC |
| HP1_R2 | TCGAGATTTTCTTCGACTTCCCAGGTTT |
| HP1_R3 | GGTCTCTCGCTTGGCTTCTCGTACC |
| MLKFYE | TCRTARAAYTKNAGCATCT |
| TLEPEEN | ACNCTIGARCCIGARGARAAYCT |
| First methylated DNA library primers | |
| DIM1F | ATTGCACGAAAGTTAGGATCCTGTGT |
| Dim1F1 | TGTTGTAGTAGATTTTGATTTTGGG |
| Dim1F2 | AGGAGTTTTTTGGTATTATAGAAA |
| Dim1F3 | TATTGAAGGTTTTGAGGTTTGTATG |
| DIM1R | TCTTTGTTTCTGATCAAGTCGACCTG |
| Dim1R1 | AATACTTTTCAATACTACCACCCTAC |
| Dim1R2 | TAAACATACCATTACCTAAAAAAA |
| Dim1R3 | ACAAAAAATTAATAAACAACAAC |
| DisulfIsom_F | GAAAAGGAGACAAAGAACGGCTACTG |
| DisulfIsom_R | CTCTTTATCCTTCTCTGCCGGCTTAT |
| DMP_Brr_F | TGGGTGCTCAAGCCTTCAT |

Table 12 (continued): Primers used in Chapter 3

| Primer name | Sequence (5' – 3') |
|---|-------------------------------|
| First methylated DNA library primers (continued) | |
| DMP_Brr_R | TCTAACGCACGCGATATTGG |
| DMP10F | GTTTCGGGTGGCTTTTCTAATGAT |
| DMP10R | ATATGGAAAATTTGGTGGGGATGAC |
| DMP15F | CAAACCATGGGTCTTGTGTATTGGT |
| DMP15R | ATAGCGAGGCAGTTTCAGACACAAC |
| DMP4F | CATTGAATTATCCAAGCCTGTCTGC |
| DMP6F | GCCATGGGTAATGAAACAGC |
| DMP6R | GGTGAATTGGCTCTGATGGT |
| DMPUrea_F | AACGTAAGGGTAGGGCGTCA |
| DMPUrea_R | CACTGTGACAGCCACTTCGT |
| LOS_BSS_F | TTTTAATTTAATATGTTGGAAGCGA |
| LOS_BSS_Meth_F | ATTGTTTTAATTTAATATGTTGGAAGTGA |
| LOS_BSS_Meth_R | CAAATAAAATTTTCACCCTATACAAA |
| LOS_BSS_R | AAAATAAAATTTTCACCCTATACGAA |
| PreMRNAspliceF | TGACAATAACCCCAAATTTCGAATG |
| PreMRNAsplicer | AGGTCGATCTCGATGTTGGGGTATT |
| RPB3F1 | AAAAAATTATACTTCAAATAAATCCTATC |
| RPB3F2 | TTGTTTTAATTTAATATGTTGGAAG |
| RPB3R1 | ATTAAGATTTAGATGAAATTAAGGT |
| RPB3R1nest | TCATAATAAACCATCATAATTCTAC |
| RPB3R2 | CCAAAATAAATTTTCACCCTATAC |
| Sth1F1 | TTTTAGATAATTAATTTGTTTTTTTT |
| Sth1F2 | TGGAAAAGGTATTAATATGGAAT |
| Sth1R1 | ATACTTTATTTATTACCCAAACAAC |
| Sth1R1 | TCATAACCTACAACCAATAAAAAAATCT |
| MeDIP adapters | |
| MeDIP 4 PCR | GACGATGAGTCCTGAGTAA |
| MseI-adaptor F (MeDIP) | GACGATGAGTCCTGAG |
| MseI-adaptor R (MeDIP) | TACTCAGGACTCAT |
| MeDIP podetia library primers | |
| BSS-unknown_F | GCGGYAGFCGGGGYYYYYAG |
| Conidial pigment methF | ATTATTTAATGTGGGTGTATGTTGT |

Table 12 (continued): Primers used in Chapter 3

| Primer name | Sequence (5' – 3') |
|--|--------------------------------|
| MeDIP podetia library primers (continued) | |
| Conidial pigment methR | AAAACATTTACTTTCCCATACTTATAC |
| GTPase methF | AAGTATGAAATTTTATAAAGGTATGT |
| GTPase methR | ACTACAAAACTAAAACAAAAA |
| LEA domain methF | TGTGTTTTATTTTAGGAAGGAGTT |
| LEA domain methR | CAAATTATCAACAAAATAACAATATC |
| Pod3 meth1F | TTGTGTATTGAGTTAGGAGATAGAGG |
| Pod3 meth1R | CATACTACCCAACAAAACCTTCATAC |
| Pod3 meth2F | TATGAAGTTTTTGTGGGTAGTATG |
| Pod3 meth2R | ATCAATTAAATACATCAAACCTAAC |
| Pod4 methF | TTTATTTGTTGTTTATAGATTTTTTTT |
| Pod4 methR | AATTATAAAACCCTTACATTTCAAC |
| MeDIP squamule library primers | |
| A1_F | TGTTTTGTATGTTGAGAGTTGATAT |
| A1_R | AACTTATCCCCACCCCTATCTTAT |
| A2_F | AATGGTATTTTGGTTAATGGTAGAG |
| A2_R | TAATCAAACAACACTACAACTCCCC |
| A3_F | GTGGATGGTATTTTATTTGTTTT |
| A3_R | AACTACATCATAACATATAACTTTAACTAA |
| CG10F | ATTAAGGGGATGGGAATAGTTTAGA |
| CG10R | AAACACTCICCTTATTTTAAAATCATAAAA |
| CG11F | TTTTTATTATTTAGAAATTTAGTTAGTTTT |
| CG11R | CACCTAAAAATACCCAACAAC |
| CG12F | TTGGAAATAGATTATTAAGAGAGAAAA |
| CG12nest | ACCAAAAAAATACCCAAAAAATATC |
| CG12R | ATAATAAACCAAAAAATAACAAAAC |
| CG13F | TAATTTATTTGTGTGATGTGTTAGT |
| CG13FSTD | TATTGCGGCATTGATAAGCA |
| CG13R | ATAATAACTATCTAAAATCCTTTTCCTATT |
| CG13RSTD | GAGATACCCCGTCATCAGGA |
| CG1-F | GATAAATATTATTTAATAGGAAAATTT |
| CG1-R | AATAACATAATAACTAAAATAAAAAC |
| CG2_F | TTTTGTTATAGTAGTTTTGGTTGTGTG |
| CG2-Fnest | TTTTGGTTGTGTGGTATTATTTGG |

Table 12 (continued): Primers used in Chapter 3

| Primer name | Sequence (5' – 3') |
|---|--------------------------------|
| MeDIP squamule library primers (continued) | |
| CG2-R | TTAATAAACACTTCTTAATACAAATACAAT |
| CG2-Rnest | AAACACTTCTTAATACAAATACAATACTTA |
| CG3-F | TGGGATAGTTTGATAAGTGGTAGAAG |
| CG3-R | CCAAACATACCCATATCTTCATCTT |
| CG3-Rnest | CAAATCACAATATTCCTCTCTTCAC |
| CG4-F | AGTTAATGTTAAGTTTAGAAGGTTG |
| CG4-R | ATCTCCCCTTAATAAAAATAATTAC |
| CG5_F | GTTTTTTAAAAGGGTATATTTATGATTAGA |
| CG5-R | AAAACAAAAAAAATTACATAAACATATTT |
| CG6_F | GGGATTTTTTGGAGAAAGTAATGT |
| CG6_R | AAATACTCTAAATTTTCAAACCTTATC |
| CG6_Rnest | CTAATTCCAAATCCTCCAACATCT\ |
| CG7_F | ATTAGTGTATTTGAAAAGGTATAAT |
| CG7_Fnest | TTTGTGTGTAATTATTATGGTAATATGAT |
| CG7_R | TAAAAAATAAAAAAATTAAAAAAAA |
| CG7_Rnest | AACCAACAAAAATAAACTAACTATAC |
| CG8_F | GAGAGGTAGTAGTTGTTGGTAGGAGAT |
| CG8_R | CAAATTTTAAAACCTTATCCACATC |
| CG9F | AATGTTTGGGGTTTTAATTTGATTA |
| CG9R | AACTTTACCTTTTTAATCCTAAACCC |

Appendix F. List of all contexts in which 5-methylcytosine is found in the fungus *Cladonia grayi*.

The methylated base is at position four, with three bases on either side as context.

| | | | | | | |
|----------|---------|---------|---------|----------------------|---------|---------|
| AAaCAAG | AGGCATC | CATCCAC | CTACCTC | GCTCCGG | TAGCAAG | TGTCCGT |
| AAACAGA | AGGCATG | CATCCAT | CTACCTG | GCTCCTA | TAGCACG | TGTCGAC |
| AAACAGT | AGGCATG | CATCCAT | CTACTGG | GCTCCTC | TAGCAGA | TGTCGCG |
| AAACCAT | AGGCATG | CATCCAT | CTCCCC | GCTCGTA | TAGCAGA | TGTCGCT |
| AAACCAT | AGGCCCG | CATCCAT | CTCCCCG | GCTCTGG | TAGCAGC | TGTCGGA |
| AAACCCCT | AGGCCCT | CATCCAT | CTGCAAC | GGACCAC | TAGCAGC | TGTCGGG |
| AAACCGA | AGGCCGT | CATCGAC | CTGCAAC | GGACCAC | TAGCAGC | TGTCTTT |
| AAACCTC | AGGCCTA | CATCGGC | CTGCATC | GGACCC | TAGCAGG | TGTCTTT |
| AAACGAT | AGGCCTA | CATCGTT | CTGCCCT | GGACCGA | TAGCCAC | TTACAAG |
| AAACTTC | AGGCCTA | CATCTAA | CTGCCGG | GGACCGA | TAGCCAT | TTACACA |
| AACCACG | AGGCCTA | CATCTAT | CTGCCGT | GGACGAA | TAGCCTA | TTACCTC |
| AACCAGC | AGGCCTA | CATCTTC | CTGCCTA | GGACGAT | TAGCTCC | TTACGAT |
| AACCAGG | AGGCCTG | CATCTTG | CTGCCTC | GGACGCG | TATCAAC | TTACTGC |
| AACCATC | AGGCCTT | CCACAAA | CTGCCTG | GGACGGG | TATCAGA | TTCCAAC |
| AACCATC | AGGCCTT | CCACACG | CTGCCTT | GGACGGG | TATCAGT | TTCCAAG |
| AACCATT | AGGCGAA | CCACAGA | CTGCGGT | GGCCACT | TATCATA | TTCCAGG |
| AACCCAC | AGGCGAG | CCACAGC | CTGCTGC | GGCCATG | TATCATG | TTCCATT |
| AACCCTC | AGGCGCT | CCACATC | CTGCTTT | GGCCCAA | TATCATG | TTCCCAG |
| AACCGGT | AGGCTAT | CCACCAC | CTTCAAC | GGCCCAC | TATCATT | TTCCCCT |
| AAGCAAA | AGGCTGT | CCACCAG | CTTCACT | GGCCCAC | TATCCAA | TTCCCGC |
| AAGCAAT | AGGCTTT | CCACCCA | CTTCAGG | GGCCCC | TATCCAA | TTGCAAC |
| AAGCAAT | AGTCAAC | CCACCCG | CTTCCA | GGCCCGC | TATCCAG | TTGCAAT |
| AAGCACA | AGTCAAG | CCACCCG | CTTCCCG | GGCCCGC | TATCCCT | TTGCAAT |
| AAGCACC | AGTCACG | CCACCCT | CTTCCGC | GGCCCGC | TATCCTG | TTGCACC |
| AAGCACG | AGTCAGA | CCACCCT | CTTCGAG | GGCCCGC | TATCCTT | TTGCAGA |
| AAGCAGG | AGTCAGC | CCACCCT | CTTCGGA | GGCCCTC | TATCGAT | TTGCAGT |
| AAGCAGG | AGTCAGT | CCACCGA | CTTCGGC | GGCCGCA | TATCGCG | TTGCCCA |
| AAGCAGG | AGTCATA | CCACCTA | CTTCGGG | GGGCAAT | TATCGGC | TTGCCCC |
| AAGCATA | AGTCATA | CCACCTC | CTTCGGG | GG _g CATG | TATCGTC | TTGCCGC |
| AAGCATG | AGTCATT | CCACTGA | CTTCTTG | GG _g CATG | TATCTTA | TTGCCTA |
| AAGCCAC | AGTCATT | CCCCAAG | GAACAAC | GGGCATG | TATCTTC | TTGCGCG |
| AAGCCCA | AGTCCAT | CCCCAGC | GAACCAT | GGGCATG | TATCTTG | TTGCGTT |
| AAGCCCA | AGTCCTC | CCCCCAA | GAACCCT | GGGCATT | TCACACC | TTGCTGG |
| AAGCCCA | ATACACA | CCCCCAC | GAACCTA | GG _g CCAG | TCACACG | TTTCAAC |
| AAGCCCG | ATACACG | CCCCCAC | GAACCTG | GGGCCCC | TCACAGG | TTTCAAG |
| AAGCCCT | ATACAGC | CCCCCCG | GAACTCT | GGGCCGC | TCACCAA | TTTCACG |

| | | | | | | |
|---------|---------|---------|---------|---------|----------|---------|
| AAGCCGA | ATACCAA | CCCCCGC | GAACTGG | GGGCCGT | TCACCGC | TTTCAGG |
| AAGCCGG | ATACCCC | CCCCCGG | GACCACC | GGgCGCG | TCACCGT | TTTCCAA |
| AAGCCTG | ATACCGA | CCCCCGG | GACCACG | GGgCTAA | TCACCTT | TTTCCAG |
| AAGCGAA | ATCCAAT | CCCCCTA | GACCACG | GGGCTGC | TCACTCC | TTTCCAT |
| AATCAGC | ATCCATC | CCCCCTT | GACCCCC | GGTCAA | TCCCAAC | TTTCGAA |
| AATCAGG | ATCCATT | CCCCCTT | GACCCGT | GGTCAAT | TCCCAAG | TTTCGCG |
| AATCATC | ATCCCTA | CCGCACG | GACCGCC | GGTCAGT | TCCCAGG | TTTCGGT |
| AATCATT | ATCCGAT | CCGCATG | GACCGGG | GGTCATC | TCCCATT | TTTCTAC |
| AATCCAA | ATGCAAT | CCGCATT | GACCGTA | GGTCATT | TCCCCAT | TTTCTGG |
| AATCCAT | ATGCACT | CCGCCAA | GACCTGG | GGTCCAT | TCCCCCG | TTTCTTG |
| AATCCCT | ATGCATC | CCGCCAA | GAGCACA | GGTCCCC | TCCCCGT | |
| AATCGGT | ATGCATG | CCGCCAA | GAGCATA | GGTCCGA | TCCCCCTT | |
| AATCGTT | ATGCCAT | CCGCCCC | GAGCATA | GGTCCGC | TCCCTGC | |
| ACACACA | ATGCCAT | CCGCCGA | GAGCATG | GGTCCGG | TCGCATG | |
| ACACACC | ATGCCCC | CCGCCTT | GAGCCAA | GGTCCTA | TCGCCAC | |
| ACACACT | ATGCCCT | CCTCAAG | GAGCCAC | GGTCCTC | TCGCCAG | |
| ACACCCC | ATGCCTA | CCTCAAT | GAGCCAC | GGTCGGG | TCGCCCC | |
| ACACCCC | ATGCCTA | CCTCATA | GAGCCCT | GGTCGGG | TCGCCGT | |
| ACACCGC | ATGCTAT | CCTCCAC | GAGCCGG | GGTCGTC | TCTCATC | |
| ACACCTA | ATGCTCT | CCTCCAG | GAGCCGT | GGTCTAG | TCTCATG | |
| ACACGCA | ATGCTGA | CCTCCCC | GAGCGCC | GGTCTGT | TCTCCGC | |
| ACCCAC | ATGCTGC | CCTCCGC | GAGCTAA | GTACACA | TCTCTTC | |
| ACCCAC | ATGCTTA | CCTCCGC | GAGCTAA | GTACATG | TGACAAA | |
| ACCCAG | ATGCTTG | CCTCTTC | GATCAAG | GTACATG | TGACACA | |
| ACCCAT | ATTATA | CGACCCC | GATCAGA | GTACATT | TGACACG | |
| ACCCCC | ATTATT | CGACCCC | GATCAGT | GTACATT | TGACAGC | |
| ACCCCT | ATTCCC | CGACCGC | gATCGCG | GTACCAG | TGACAGG | |
| ACCCCT | ATTCTT | CGACCTA | GATCGGT | GTACCGG | TGACAGT | |
| ACCCCGC | CAACAAA | CGACCTG | GATCTCC | GTACCTG | TGACCCC | |
| ACCCCGG | CAACAGG | CGACGAT | GATCTGT | GTACGAC | TGACCCC | |
| ACCCCTC | CAACAGT | CGCCAAT | GATCTTG | GTACGGT | TGACCCC | |
| ACCCGGG | CAACATC | CGCCACT | GCACAAT | GTCCAGG | TGACCTC | |
| ACGCATT | CAACCAG | CGCCAGC | GCACAGG | GTCCATA | TGACGAG | |
| ACGCATT | CAACCAT | CGCCATC | GCACCAC | GTCCATG | TGACTAC | |
| ACGCCAC | CAACCAT | CGCCATT | GCACCCA | GTCCATG | TGACTAC | |
| ACGCCAG | CAACCCT | CGCCATT | GCACCCC | GTCCCCA | TGACTCG | |
| ACGCCAT | CAACCTT | CGCCCCG | GCACCGC | GTCCGTA | TGCCATA | |
| ACGCCCC | CACCAA | CGCCCCT | GCACCGG | GTCCTTG | TGCCATC | |
| ACGCCGG | CACCAA | CGCCCGT | GCACGCG | GTGCACT | TGCCATG | |
| ACGCCTG | CACCCAC | CGGCAAC | GCACGCT | GTGCATC | TGCCATG | |

| | | | | | |
|---------|---------|---------|---------|----------|---------|
| ACGCCTT | CACCCAG | CGGCAAC | GCACGGG | GTGCATT | TGCCATT |
| ACGCGAG | CACCCAT | CGgCAGC | GCCCAAG | GTGCATT | TGCCCAA |
| ACTCAGT | CACCCCA | CGGCAGG | GCCCATT | GTGCGTA | TGCCCAC |
| ACTCCCA | CACCCCC | CGGCAGG | GCCCCGA | GTGCTAC | TGCCCCC |
| ACTCCCC | CACCCCT | CGGCCAA | GCCCCGC | GTGCTGT | TGCCCCC |
| ACTCCGT | CACCCGA | CGGCCAC | GCCCCGC | GTTCAAA | TGCCCCT |
| ACTCCTC | CACCCTA | CGGCCAT | GCCCCTT | GTTACC | TGCCCGA |
| ACTCCTT | CACCCTA | CGGCCCA | GCCCCTT | GTTACC | TGCCCTT |
| ACTCGCC | CACCCTG | CGGCCCA | GCCCGTC | GTTCAGC | TGCCCTT |
| ACTCTTC | CACCGCC | CGGCCGC | GCCCTTT | GTTCTT | TGCCGAG |
| AGACAAG | CACCGGT | CGGCCGG | GCGCAAG | GTTCTAA | TGCCGGG |
| AGACAGG | CACCTTG | CGGCCTC | GCGCAAG | GTTCTAT | TGCCTTC |
| AGACAGT | CAGCAAT | CGGCCTG | GCGCAC | GTTCTTG | TGGCAA |
| AGACCGC | CAGCACC | CGGCGAA | GCGCAGT | GTTCTTG | TGGCAAG |
| AGACCTG | CAGCAGC | CGGCTAT | GCGCAGT | TAACAAA | TGGCAGA |
| AGACCTG | CAGCATC | CGGCTTT | GCGCATA | TAACACT | TGGCAGA |
| AGACGTC | CAGCCAA | CGTCACT | GCGCATG | TAACACT | TGGCAGA |
| aGACGTG | CAGCCAA | CGTCACT | GCGCATG | TAACAGA | TGGCAGC |
| AGACTAA | CAGCCCC | CGTCATG | GCGCCAC | TAACAGG | TGGCAGG |
| AGCCAAA | CAGCCCC | CGTCATG | GCGCCAT | TAACCCA | TGGCAGG |
| AGCCAAC | CAGCCGT | CGTCATG | GCGCCAT | TAACCCC | TGGCATA |
| AGCCACG | CAGCCTC | CGTCGCT | GCGCCCT | TAACCTC | TGGCATG |
| AGCCATA | CAGCCTG | CGTCGGG | GCGCCGG | TAACCTAT | TGGCCAT |
| AGCCATT | CAGCGAA | CTACAGA | GCGCCGG | TACCACC | TGGCCCC |
| AGCCCAC | CAGCTGC | CTACCAA | GCGCCTG | TACCCGG | TGGCCGG |
| AGCCCAG | CATCAAT | cTACCAC | GCGCGCT | TACCCGT | TGGCCTA |
| AGCCTAG | CATCACT | CTACCAC | GCTCAAG | TACCGAC | TGGCGAG |
| AGGCAAT | CATCAGA | CTACCAG | GCTCACG | TACCGGG | TGTCAAG |
| AGGCAGC | CATCAGC | CTACCAT | GCTCATG | TACCGTC | TGTCACG |
| AGGCAGG | CATCATA | CTACCCC | GCTCCAA | TACCTAT | TGTCACT |
| AGGCATA | CATCATC | CTACCCG | GCTCCCC | TACCTGA | TGTCATT |
| AGGCATA | CATCATG | CTACCGG | GCTCCCC | TAGCAAG | TGTCCAC |

Appendix G. Amino acid alignment of DNA methyltransferase proteins identified in lichen genomes to *Neurospora crassa* DIM-2 DNA methyltransferase.

Ncrassa = *Neurospora crassa* DIM-2; *Cgrayi* = *Cladonia grayi* homologue; *Lsp* = *Leptogium* sp. homologue; *Dbaeomyces* = *Dibaeis baeomyces* homologue; *Astriga* = *Acarospora strigata* homologue; *Epallid* = *Endocarpon pallidulum* homologue; *Arubroc* = *Arthonia cf rubrocincta* homologue; *Pcylind* = *Peltula cylindrica* homologue; *Pstella* = *Physcia cf stellaris* homologue. **** = bromo-adjacent homology (BAH) domain; #### = DCM DNA methyltransferase domain.

Ncrassa MDSPDRSHGGMFIDVPAETMGFQEDYLDMFASVLSQGLAKEGDYAAHQPLPAGKEECLEP
Cgrayi -----
Lsp -----
Dbaeomy -----
Astriga -----
Epallid DAILVESAIEAVVSTSSITSIYSENSSDTISISSADRLDQKLPDQQFSITQTPDLFDRGS
Arubroc -----
Pcylind -----
Pstella -----

Ncrassa IAVATTITPSPDDPQLQLQLELEQQFQTESGLNGVDPAPAPESEDEADLPDGFSDESPDD
Cgrayi -----
Lsp -----QRPRCTTIQGI
Dbaeomy -----
Astriga -----
Epallid GEVL DHTASQEKERPQDLHFRSVCRLLEDAENFEDTGVRAGSRKDGSSQRDISCSEGVTSL
Arubroc -----
Pcylind -----
Pstella -----

Ncrassa DFVVQRSKHITVDLPVSTLINPRSTFQRIDENDNLVPPPQSTPERVAVEDLLKAAKAAGK
Cgrayi -----MLDGFSIYRTHQMRSQKKHKRPEE
Lsp SSARFGESKISLSRAFASPFNGGEDCLVRHTSRFIVPRSFNIARVTKSKEESKTKSEDE
Dbaeomy -----
Astriga -----HDCAEKASTRDRIRKHSTGS
Epallid SESSLEENGGEESLMDEVLSVLGVGGPTARKRAIHPSDTESSSTIPNDYSHSSPLSQNLEC
Arubroc -----
Pcylind -----LYQSQSEAELANAGRRSPSQDV
Pstella -----EQAGSEAQDSARRSFQGEQFLALDLDD

Ncrassa NKEDYIEFELHDFN-----FYVN-YAYHPQ-----
Cgrayi -----SSQAN-----
Lsp NCHESSTFIEADLEDFSIYLPHNYTSLGPNEFAEKTLSN-----
Dbaeomy -----
Astriga NSAPYLSFLLTDFS-----IYRTGEGQHPD-----
Epallid VSIPPLLVPQSKYIGWKHSLSTTDERATLSRILEAQAGNQEQPDSFHDILVDDFVIYRPG

Arubroc -----
Pcylind SEGGFISYRLEQFS-----IYKPGSGKYPD-----
Pstella FSIYRRHTAVKNFR-----RDGLEEKVRAN-----

Ncrassa -----EMRPIQLVATKVLHDKYYFDGVLKYG-NTKHVYTG--MQVLELPVGNYG-ASL
Cgrayi -----ELVSLHEINER-NVGRLCFDGWIIDG-EKARYVEN--VLFSIMAIGGYKELDR
Lsp -----ELTSLHGIYQR-GDRLRYMDGVLSTFG-GKRRYVQR--VPFDTSLIGGYEDKSL
Dbaeomy -----MAALHDL CVKPKNESYYIDGCLVASGSTRRYVEK--VPFKWLSIGGYGDPTV
Astriga -----EFTSLHNLHTYRGHHNYLLDGLLCHG-SSRRYVQK--VPFDILSIGGYEDTNR
Epallid YPTHYPWQVVTLDNIIISNRESCRYLLNGVLKCG-AIAEYVEAMEIDITISIDGFEDTEI
Arubroc -----
Pcylind -----EFAPLHHL CVKPANNYYFDGILEEG-QMRRYVQN--VAFKFLSIGNYCDTVQ
Pstella -----ELMSIHEVDRR-GSSQWVFDGFCVCHG-QHREYVER--VPCTNFSIGGFEDINW

Ncrassa HSVKQGIWVRSKHNAK----KEIYYLLKKPAFEYQRYYPFLWIADLGKHVVYDYG---T
Cgrayi PTVGSEIWIQSAQSQK----AGVWYRLRTPAPEYRRLHEPFLWIADLAKHVVD-----
Lsp HTIGSEIWIQSRGRK----MDVWYRLKAPAPEYRRYHEPFLWVAELAKHVID-----
Dbaeomy HSKNGMVWIQSIKSLQ----RNTWYRLGKPSPEYKRFHEPFLWISNFAKYFVN-----
Astriga HTVGADIWLQTKQSAR----FDVWYQLGKPAPEYERYHQPFLLWADLAKHVVD-----
Epallid HSVQDMVYLQTVVCARRRTALDCWYRLGEPGQYRDLHKRFLWVADLAKHVIDYANWKY
Arubroc -----DLSKHFVD-----
Pcylind HGVGDQLWIQSTRLER----SDVWYQLGNPAPEYRRYHTEFVWLANFAKHFTD-----
Pstella HTVGSSVWIQSQASEK----LNLWYCLRRPSAEYKRFHEPFLWIADLAKHLVD-----

Ncrassa RMVERKREVTLGCFKSDFIQWASKAHGKSKAFQNWRAQHP SDDFRTSVAANIGYIWKEIN
Cgrayi -YLHIHWQVTLNHFRRFRFHTWMQTLYGSNAHACNWLN-CGSRDFRRIVAAHANFLYCQAS
Lsp -YMTDHERVSI SFRSNFFGWLEEVHGSSDTFRTWH SKYKDADFRRVIACHSGFLYYQAM
Dbaeomy -FLDENECVSI SSFSEDFYSWLNLSLPQRRVGFQKWLHDYGD TDFRRVVTAQSEFLSKEAG
Astriga -FLDNHDNVSLQNFRTGFFSWINDLHEQDASFQTWLAQYPGTDFRRAIAA YPSFLFNQAR
Epallid HSEDHNMVRLKDFQQDFMTRIVDWHAGSLQFE EWLHVYNKPDFRIPLNRHREFIYSRAT
Arubroc -FLDNHDSVSLQDFRTNFFSWINDLHEQGFQMWLAQYSGTDFRRAIAAHPGFLFNQAR
Pcylind -FLHRHVDVHIHDFREKFYIELEELHGTDEAFQTLWKEYGDTDFCRVVA AHPEFLYKESV
Pstella -YINNHNEVSLIQLREHFCLWLKRTHDPNMIFQQWLKTYGDNDFRRIVA AHAEFLYNQAA

Ncrassa GVAGAKRAAGDQLFRELMI V KPGQYFRQEVPPGP-VVTEG---DRT-----
Cgrayi QVDH-----RLEDHPVWKEI HPRFLSAIEEQ---VERGTTLD--MYTLSTR
Lsp QLDP-----KYDLHPVWAETDPEALNAIPKR---PST-----NKSCS
Dbaeomy -----LLDDSYLIHP--IWDEVHPRGLNAIRNQ---PLV-----
Astriga DL SG-----AAFDKLPLSVEVHPPELNRIRPQ---EYV-----
Epallid GLGG-----HYSDDLWSDLMVRPAVMSDCNGLP-----
Arubroc DL SG-----TAFNQHPLSVEVHPPELNRVRPQ---KCV-----
Pcylind DVE-----RRNDGQPIWREVHPKSLTAVREQ---KVV-----
Pstella LLGK-----KYTLHPLWGEIDPFAMNAVPRQ---TKQIKTYEPIASTNHAK

Ncrassa -----VAATIVTPYIKECFGHMILGKVLRLAGEDAEKEKEVKLAKRLK IENKNATKADT
Cgrayi GSTTSSRRKTTVNPYVYKCFENLTWAKFLYSQTPSISVA-TRSSC-----
Lsp SNSASIVTKTVVTPYVFGCFKHLPFARFLDCKTPEPHVLRRLRNR-----
Dbaeomy -----EAKTLVTPYVYSCFKDMSWAHLKVGASLTGRHRESKSIDRVACPRVRPRIKRL
Astriga -----ETSTIVTPFVYECFKHISWARFLESRKPCA KVLKMRRGN-----
Epallid -----EDTLVTTYVKHCCRSMPWSYV LKAERMSESVRRQQRR-----

Arubroc -----ETSTIVTPYVYECFKHISWARFLEPRKPCAKVLKTRHEK-----
Pcylind -----EQLTVVTPFVYGCFKHLPWGKFLNPVEPDPRAFKLRRRREQAFHLTADSPVHMT
Pstella LSEPTSRPKTVVTPFVYECFRHLPWAKFLDLQHPSTSHPKRSYSR-----

Ncrassa KDDMKNDTATESLPTPLRSLPVQVLEATPIESDIVSIVSSDLPPSENN-----PPPLTN
Cgrayi -----
Lsp -----
Dbaeomy VDHCNP-----
Astriga -GR-----
Epallid -----
Arubroc -GR-----
Pcylind PGQGEIVAHNKVMRNSTPAKNLRSKSTSRVPETSEMLHEFFDELDPDYEINERQERLTA
Pstella -----SSP

Ncrassa GSVKPKAKANPKPKPSTQPLHAAHVKYLSQELVNKIKVGDVISTPRDDSSNTDTKWKPTD
Cgrayi ---LNGGPGLPSGLSEVHVYAQVS--SPSEIKSRSIHKGDVVALPRDKES----PWKT--
Lsp ---LAASSFYPSGAKVATPMKKQDELDAFPAKPVQVSVGDVVAINSKQS----SWKH--
Dbaeomy -----SILVEVGDVIKLAHDKKS----IWK--
Astriga ----APKQTIETGMTTSRRVHQGP-----LAPESINVGDVVGVSKDNDT----AWKC--
Epallid --ATAMGFFLPSLHHHSLGLDEIPRTAALLEVAAKSGLTSAISADEASGRFAIVSWKR--
Arubroc ----APKQTIETDRTTDRRRHQGP-----LPPEHIKVGDVIGVLKDNNT----AWKG--
Pcylind RAEIAKKKFIGVITASTRIHRDQEGNLLLQRPDSIQVGDVVGIERDQDT----VWKG--
Pstella STSFRPISKFGREVTEVEDPHSPSRE-----ERQTVQIGDVVAI-----

Ncrassa TDDHRWFGLVQRVHTAKTKSSGRGLNSKSFVDVIWFYRPEDTPCCAMKYKWRNELFLSNHC
Cgrayi -IDTHWYGIVHGIKSSKGP-----LELLWLRYRPSDTECRKMHYPFKELFLSDHC
Lsp -EDRLWYGYVHETQEKVKGTS-----LKI IWLYRPTDTACQRMKYAYANELFISDHC
Dbaeomy TTDAYWYAYVQGIENRCEQN-----LKILWLYKPADTICSESFYPWKNELFLSNHC
Astriga -NADMWFAYVQDMRSDRIGGT-----LLDVLWLYAPSDTTCSTGRYPFKNELFFSDNC
Epallid SQGRTTYHDYSYVYIQKVVKS-----ELKVIFVYLPSETICLDGHYPHQNELFLSECC
Arubroc -DADVWFAYVQALQSNRKGKT-----VLDVLWLYAPSDTTCSTGRYPFKNELFFSDNC
Pcylind -NADLWFAYVQDVRTSARGQS-----SLKVIWLYAPSDTSCSTMHYPIHNELFFSDHC
Pstella -----
*****BAH DOMAIN*****

Ncrassa TCQEGHARVKGNEVLAVHPVDWFGTPESNKGEFFVVRQLYESEQRRWITLQKDHLTCYHN
Cgrayi NCGD---APILVQEVTHRPRVAFFGTPDTEVDVDFVVRQTYIQGHGAWESLKDSDFSCRCY
Lsp NCND---RRIYASEVISKPKVSFFAGPDT-DADFFVRQKYVGEDAAWVTLQESDFECPCK
Dbaeomy ECTR---TRIRAEIDGKASVTFYGSPSSS-TEFFVVRQRYADET--FMDLRPEHFKCLCH
Astriga NCED---ATLKLTEVICVSVAFFSHPRETAAEYFVRQKYTH-EETFVTLRKSDFCVHR
Epallid NCSAD-CAPIFLSDIIRLVGVT-IGDEIVGLQDFVVRQKYVHNEDAISRLKSDFLCPCL
Arubroc NCED---AVLNLADVICKVSVEFFSHPQETAADYFVRQKYTH-EETFVTLRKSDFCVHR
Pcylind NCKD---STLPASDVVCKVTVDFFRGPGEAGAEFIVRQKYKIGDAAFVTLKKSDFRCVHY
Pstella -----

Ncrassa QPPK-----PPTAPYKPGDVTLATLSPSDKFS----DPYEVVEYFTQGEKETAFVR-LRK
Cgrayi KKEE-----FSQYSAGDTLLVR-----IRGILEPVVFI EEDYRDLQGKFRVRRLLR
Lsp TDSI-----PLQYSVGDTLLVKN--PPTGSHQMLEAVEAIEKAPDGNYNHLRIRRRLLR
Dbaeomy RPS-----PTLGYQPGNTVLVEKVISGRVAT--LEPVEIVDLPLRTGKELVKVR-LLR
Astriga SALSKEIEELMEDYRVGDTVLVLG-LGDKDDHERLEPAEIVAFSNG-----VQIRRRLLR
Epallid RPPD-----VASTDEKYDQVQTEP-----
Arubroc SASGKSEIEELMEDYRVGDTVLMMG-LGDKDNDNRLEPAEIIAFSKT-----VQVRRLLR
Pcy lind SGNGKSEMEEFKERYHVGDTVLIRKRIGQKEDG--LEPVEIVEFRVEGSTQLLLLVRRLLR
Pstella -----

Ncrassa LLRRRKVDRQDAPANELVYTEDLVDVRAERIVGKCIMRCFRP----DERVPSPYDRGGTG
Cgrayi -MRDYGHMD--AEPNELVISNKFQSIARTDVDRECHVRFYTEENKTLGQIPPPYCRGSS
Lsp KSRHCGTDD--AEPNELMYTDIFEVISVKDVVRACHVRFYTAEDKVHCRI PCQYRRGGTG
Dbaeomy RKRR-DYGDMLARLNELVYTNKFEYLPISDLKRSCHVRFSSQKSLNNGDIPVPYSRDGIG
Astriga RGRDFGSEHCNARNPNELVYTDILLIPASAVDRRCHVRFYTV EELNASQIPAPYNRDGNA
Epallid -----
Arubroc RGQDFGHEHRDARNPNELVYTDILLSIPASAVDRRCHVRFYTV EELNSSQIPAPYNRDGNA
Pcy lind RSRDFPNEK-DARNPNELVYTDDEMFTI PAEKVERRCDIRFYTEDDKAKRRI PAPYNRDGTT
Pstella -----

Ncrassa NMFFI THRQDH--GRCVP---LDTLPPTLRQGFNPLGNL GKPK--LRGMDLYCGGGNFGR
Cgrayi DHYYI ILEDLD-DSELGPRPFSSTSLKLINQGWDP CNMLLRP--MRGLDLYCGGGNFGR
Lsp DFFYIMHH-----RPWPGS---MKEGWDPTTEVQDTPPV MKGLDIFCGGGNLGR
Dbaeomy DAYYIVNEEDE--HHQILRELQEPYPSSLVQGF DPLRNSTMPK--LKILDLCGGGNFGR
Astriga DAYYIAYRQTE---AKTLEPLHIPFPATLIQGF DPRGPPVKSV--MNGMDLFCGGGNFGR
Epallid -----LSGLGLFAGCGNFDL
Arubroc DAYYITYRQTE---TKTLEPLRTPFPATLIQGF DPRVPPVKPV--MNGMDLFCGGGNFGR
Pcy lind DAYYITCRQVDGSLSPRLEPLLEFPNTLIQGF DPSAPSIRRR--LNGMDLYCGGGNFGR
Pstella -----

#####

Ncrassa GLEEGGVVEMRWANDIWDKAIHTYMANTPDP--NKTNPFLG SVDDLLRLALEGKFS DNVP
Cgrayi GLEEGGAVKMEWAVDWDYNEAIIHTYKANMVDQ--DAVKLFRGSVNDYLTQAMQ GKGRGLVA
Lsp GLEEGGAVEFAYAVDYFKEAIIHTYRANLSNP--DKTDLFYGSVNK YLSQAMRGDPS--VA
Dbaeomy GIEEGKVAEVRWAIDIDSNALHTYRANLPSP--ENVGLYLGSVNDYLSQALDGKFSELIA
Astriga GLEEGGSVLNKWAVDWDNTAMHTYSANLGEAAPGVTQLYNGSVNDYLAQAMDGSSAKEIA
Epallid GLEAFGAVKFVAAVEISETALKTYAANRSMG---FHGLIFDSVNPTLLKILQGLD----
Arubroc GLEEGGSVLNKWAVDWDNTAMHAYSANLREADSEDT HLYNGSVNDYLAQAMGGSSAKEIA
Pcy lind GLEEGGAVHNKWAVDYDRDAIHTYHANLKYP--QDTALYFGSVNDLLAQAMKG RYTKYVP
Pstella -----

#####

Ncrassa RPGEVDFIAAGSPCPGFSLLTQDKKVLNQVKNQSLVASFASFVDFYR PKYGVLENVSGIV
Cgrayi QLGEVEFISAGSPCPGFSTSNHNKANDQGLFDAS MVASVVSFVDFYRPRYALMENVKGMA
Lsp QLDDVQVI SAGSPCQGFSNANQWRENDRALLNVSMVASVVA FVDFYRPKYALLENVTGMA
Dbaeomy RPGEVDVVIAGSPCQGFSTLNNNKSSEKGLRNQSMVASVAAYVDFYRPKYAIL ENVPGMA
Astriga QVGEVDFI SAGSPCQGYSVANMARGSDNSLKQCSMVA AVAAFVDFYRPKYALLENVT SMA
Epallid -LHDIELI SAGSPCKGWSRANPYRGDDNGMRNCSLVASTMSYI ETFLPSYAVLENVGAMG
Arubroc RVGEVDFI SAGSPCQGYSVANMARGSDNSLKQCSMVA AVAAFVDFYRPKYALLENVTAMA
Pcy lind EPGEVDFI SAGSPCQGFSNANQKKSNEKSLRNSSLVASVA AAFVDFYRPKYALLENV VAMA

Pstella -----
#####

Ncrassa QT-FVNRKQDVLSQLFCALVGMGYQAQLILGDAWAHGAPQSRERVFLYFAAPGLPLPDP
Cgrayi SGQD---TNNVLAQVICALLGMGYQVRTSGLDAWNFGSPQSRSRIFISIAAPGMEPMPEP
Lsp KCNAKDRDNNVFAQVLGALVGMGYQTQSFILDAWSFGSPQSRTRLFISIAAPGLAPLSRP
Dbaeomy SK-NGQSK-SVLAQMNCALVGMQYQVQNFITLDAWSFGSPQSRTRLFVAAAAPGMMPLEVP
Astriga KKGAQDPDDNIFSLMLCTLVGMGYQVRPLHLDAWSFGSPQSRRLFISIAAPGFQLPPHP
Epallid SG----VGNSCNQVIACLVLGLGYQIRKMKLNSRDFGSSQGRDRLFLAAAPNVPLPKEP
Arubroc KKGAQDPDDNIFSLMLCTLVGMGYQVRPLHLDAWSFGSPQSRRLFISIAASGLELPPHP
Pcylind AKGKEKQDQNVFSQLLCSLVGMGYQVQFNLDAWSFGSPQSRRLFVSIAPGLRPLPPP
Pstella -----
#####

Ncrassa LPSHSHYR-VKNRNIGFLCNGESYVQRSF-IPTAFKFVSAGEGTADLPKIGDGKPDACVR
Cgrayi PHIHSHPETVTSASLGKTPNGLHTGSRYT-SLTPFSYISAAEATKDLRSR-TDGR-TSCIP
Lsp PQTHSHPDFIPSRSLGKSANGLTFANRHW-DITPFSYVTIGDATKDLPMNNDGR-VDCIA
Dbaeomy QPSHSHPEGVRNHSGLRVANGLPVSQRQLGVSTPLKYVTIGEATADIPYNHDGR-TTSIC
Astriga APSHSHPLNMTQRSLGKAANGLGFGCRQF-GATPYKYVTASEATS DLPSIGDSRVAVCIT
Epallid RPSHSHKTTGFAKAS-----EVSQGLPPMDNDD-LICIS
Arubroc APSHSHPPRMTQRSLGKAANGLGFGCRQF-GATPYKYVTASEATS DLPSIGDSRILVCVP
Pcylind SLSHSHPPGT-----
Pstella -----
#####

Ncrassa FPDHRLASG-ITPYIRAQYACIPHTPYGMNFIKAWNNGNGVMSKSDRDLFPSEGKTRTSD
Cgrayi HPDHRMSFA-LSTLHRVCISSIPRFPGGGSFISAFSA--KGYMPQAQIDDFSWDNPIRASK
Lsp FPDHRPGRT-MTTVNRVRMSCIPRFPGMNYVQAAL--LGLMPPQMEAHNWNNTTRINP
Dbaeomy FPDHRLTGAGMDTEHRLRISCI PRHPENQSVNTARSFIHEKLIKTCPTIFKHS--MRSRG
Astriga YPDHRPSRI-ESTHTRVLISHIPRFPYGETFMTAYR--QGRMSQPQIDTYPFKSNVRGSD
Epallid HPDHIAGAK-QTPLFRELIRRI PRYPRTMNFLLQSVK--KGYQGKAQLEWLLSRKGLHDVE
Arubroc YPDHRPSRV-ESTHTRVLISHIPRFPYGETFMTAYR--RGCMSQPQVDMYPLKSNVRGSD
Pcylind -----
Pstella -----
#####DCM Domain#####

Ncrassa ASVGWKRLNPKTLFPTVTTTTSNPSDARMGP-GLHWDEDRPYTVQEMRRAQGYLDEEVLV
Cgrayi GSKSWKRVRGSLIPTIMTHPRPWDGLGSG-CLHWDQHRLLTVM EVRRAQGFDPHEVLIG
Lsp GGRAWKRVNQNGLIPTIITTCRPEDDFCGA-WIHWDACRCMTVM EVRRAQGFDPDEEVLIG
Dbaeomy DSRCYRRTSANKLLPTVTGYCNP GDYKSGR-VLHWEASRCLTVM EVRRAQSY PDRDVLIG
Astriga DAKTWRRLRPDGLFPTITTTGIRPQDGRCGA-TLHWDEHRLMTVM EARRAQGFDPDEEVLIG
Epallid DHMAFTRLDPDGLIPTITTSR PACGWGGGRI IHWDEHRTLTLLEARRAQGIPDEEVIIG
Arubroc DAKTWRRLRPDGLFPTITTTGIRPQDGRCGA-TLHWDEHRLVTVM EARRAQGFDPDEEVIIG
Pcylind -----
Pstella -----
#####

Ncrassa RTTDQWKLVGNSVSRHMALAI GLKFREAWLGTLYDES AVVATATATATATAAVGVTVPVM
Cgrayi SPSEQWKI IGNSVDRSVALALGLSLRTAWLANAAKETINARFENTTTTETIKEP-----
Lsp MPSSQWKI IGNSVARPVAFALGMELRS AWLKNSVPLANQSAAGPKSTS ISSKM-----
Dbaeomy RPSQQWKIVGNSVARTVALALGMSLRAAWLTTVAFQSLVSTEDGVTGQQEVL R-----

Astriga SPAQQWKIIGNSVARPVAVALGMSLREAWLANNPDEPPSDEVYRTQTTDTSSHQOTLPRE
Epallid DLPKQWAQVGNVNRQVARALGKVVADSWFASKRASPGPEIVVAIPANKRSGRSQK-----
Arubroc SPAQQWKIIGNSVARPVAVALGMSLREAWLANKPDEGPLDDVYRTRTSNAS-----
Pcylind -----
Pstella -----
 #####

Ncrassa EEPGIGTTESSRPSRSPVHTAVDLDDSKSERSRSTTPATVLSTSSAAGDGSANAAGLEDD
Cgrayi -----PFDHLTQTLKS-----
Lsp -----PQDAVNLTWESDDSDL-----
Dbaeomy -----
Astriga TILSKSSPAPRSKQNPQKEVIDLISDDDDDELGQPQPILMN-----
Epallid ---AKKEARLSSVGDDHVNGGLNLDMEQGEVQLVQDIFRHYG-----
Arubroc -----RTAKQPQQEVIDLTSSGD-ELSQPQQGILSN-----
Pcylind -----
Pstella -----

Ncrassa DNDDMEMMEVTRKRSSPAVDEEGMRPSKVQKVEVTVASPASRRSSRQASRNPTASPSSKA
Cgrayi -----AMIEGEAPSIIWQPTDMLDSSRFATTTSGATTIATSNSAGRVSRESIIIEE
Lsp -----EMLPNPPPQSKPCMQSMLSAN---PSTSVDFNFISLKRVRSHSSTSHLDD
Dbaeomy -----AEFSEPSLDSKSKHQTVQVQVGPKEGKCEERAAECRDNKRARP----
Astriga --VQVKIPNRRISAPQQSPPFSAPGPTRTTLTRQITHSETRITRTITTTTRQYTASPDITP
Epallid --RTSDTKQVSINNNASDLSFRQIQRREISTIERNAAAGDKKSVRMFAERSIEVCIKAPT
Arubroc --VQAKIPNRRNSAPQRPPPLSAPAPTRTTLTTQVTHSETRITRTVTTTTRQYTAGPDTP
Pcylind -----
Pstella -----

Ncrassa SKATTHEAPAPEELESDAESYSETYDKEGFDGDYHSGHEDQYSEEEEEEEYAEPETMTVN
Cgrayi PTSGRYQESSSQSSQGPSHQRPQSQDSSHPTLLPSKVTTRETTISKVTITYYNPS-----
Lsp SHGKRQRQRHHRPVPTVITSSSECDSSTESLPESQSTTQTNRSMANQVVIYIG-----
Dbaeomy -----IDSVLSNSTLRHECLEGKLK-----
Astriga SALQARSRTTQKPRQRSVVQRLADGYRALFDDKSGDNGIPKKFEDGSVLEISD-----
Epallid SPAEIEYGGAEHPIDLEAEVSIKRRREDISDDEEDEDLKNWVGKRVRSRMNN-----
Arubroc TALRARPRTAQKPKQRSVVQRLADSYRALFDGKDEDGNGGPKKFEESLTVETSD-----
Pcylind -----
Pstella -----

Ncrassa GMTIVKL
Cgrayi -----
Lsp -----
Dbaeomy -----
Astriga -----
Epallid -----
Arubroc -----
Pcylind -----
Pstella -----

Appendix H. DNA methyltransferase proteins identified from lichen genomes aligned to *Neurospora crassa* RID-1 DNA methyltransferase.

Ncrassa = *Neurospora crassa* DIM-2; Cgrayi = *Cladonia grayi* homologue; Lsp = *Leptogium* sp. homologue; Gscript = *Graphis scripta* homologue; Dbaeomyces = *Dibaeis baeomyces* homologue; Astriga = *Acarospora strigata* homologue; Epallid = *Endocarpon pallidulum* homologue; Arubroc = *Arthonia cf rubrocincta* homologue; Pcy lind = *Peltula cylindrica* homologue; **** = C5 DNA methyltransferase domain.

| | |
|-----------------|---|
| <i>Cgrayi</i> | MDRFFTHDFQ RSLRTNIYPANGARDENIVTITKDDGTTSDGGEGVRSFRR |
| <i>Lsp</i> | ----- |
| <i>Gscript</i> | ----- |
| <i>Dbaeomy</i> | ----- |
| <i>Astriga</i> | ----- |
| <i>Arubroc</i> | ----- |
| <i>Epallid</i> | ----- |
| <i>Pcy lind</i> | ----- |
| <i>Ncrassa</i> | ----- |

| | |
|-----------------|--|
| <i>Cgrayi</i> | LLAHDPVQISDDEDSKSDEGFLDWVDLSGEGETTTSPAISSDFDIPTPPL |
| <i>Lsp</i> | ----- |
| <i>Gscript</i> | ----- |
| <i>Dbaeomy</i> | -----M |
| <i>Astriga</i> | -----MDDES DSSARMTDVLLEDDSQGQKRSPASSEL |
| <i>Arubroc</i> | -----M |
| <i>Epallid</i> | ----- |
| <i>Pcy lind</i> | ----- |
| <i>Ncrassa</i> | ----- |

| | |
|-----------------|---|
| <i>Cgrayi</i> | SSIILPRNPPIGT PFLSLDDYTYNGTRLHPRVFVELRDGDFMKIVHIVKD |
| <i>Lsp</i> | ----- |
| <i>Gscript</i> | ----- |
| <i>Dbaeomy</i> | VLQLSPEDYMDDDDVPY YLSQLSPKKPSVSSFSSRSTPEPSQRRHTPPRS |
| <i>Astriga</i> | IETMYETLERITSDNRTLKAGKTVEISGGEFLRIQSVIRDITTEIFLQG |
| <i>Arubroc</i> | SAVVIEDDYQE QKRSPVSIPLTSLLEPIYETLERVSSDNRTLKAGKTVEL |
| <i>Epallid</i> | ----- |
| <i>Pcy lind</i> | ----- |
| <i>Ncrassa</i> | ----- |

| | |
|----------------|---|
| <i>Cgrayi</i> | AISSDVTIRGWIFRRTRKTKGIFNAKRNEVCWLICIDDDDP RDHSIQSLE |
| <i>Lsp</i> | ----- |
| <i>Gscript</i> | ----- |
| <i>Dbaeomy</i> | HRVTIQSYSYRKINIRAKATVELDDGIHFMRVVELIQNTDNGEV SIRGWI |
| <i>Astriga</i> | HRLRRAKALEGLLELKL NEMVMILEVDQADPRPVMVKGIEQIPLAQVGRI |
| <i>Arubroc</i> | LDEDFLRIQSVIRDINTTEIFLRGHKLRRAKNLEGLLELKLNEVVMILEV |

| | |
|-----------------|---|
| <i>Epallid</i> | ----- |
| <i>Pcy lind</i> | ----- |
| <i>Ncrassa</i> | ----- |
| <i>Cgrayi</i> | TRRVTEVVKRRQIRLTNQLFPALSYRDDRISESEDETIEDVRVLVCRYKYV |
| <i>Lsp</i> | ----- |
| <i>Gscript</i> | ----- |
| <i>Dbaeomy</i> | FRRAAFMYGLLEKKLNEVCWILDVEENDPRDPKEQGIQEIPVARVIKRR |
| <i>Astriga</i> | RELVITNRPFPELSFRERAHERPWNRDVANYGRLVCRWKYIVSFRTARDK |
| <i>Arubroc</i> | EESDARPV MVQGIEQVPLAQVGRIRELVVTNRPFPELSFRECAHKHFADG |
| <i>Epallid</i> | ----- |
| <i>Pcy lind</i> | ----- |
| <i>Ncrassa</i> | ----- |
| | |
| <i>Cgrayi</i> | VHYPNAKARISSHPWSEKMLRLRADDCDRRPNNDVKDEDIRTAWRGETI |
| <i>Lsp</i> | ----- |
| <i>Gscript</i> | -----MGDLNGQYITIDDRDDE RATTPLYESE |
| <i>Dbaeomy</i> | MRLTNLPFPMLSFRDEPSGDSMDVIKNHGIIVCRWKYITSYPAADAKMRG |
| <i>Astriga</i> | QRNKYPETCLKR FQQSEIDPRYAVSDDLRLQRWRGDTVKGGA CLNMRAD E |
| <i>Arubroc</i> | DIANHGRLVCRWKYIVSFQTLGDKRSDKNSEKCLTRFQQGEIDSGFAVND |
| <i>Epallid</i> | ----- |
| <i>Pcy lind</i> | ----- |
| <i>Ncrassa</i> | ----- |
| | |
| <i>Cgrayi</i> | PGGSQEGWLPGEKEYLRQESVSHRGIKAHQSLKGPAGLEYPAGDPMTRGN |
| <i>Lsp</i> | ----- |
| <i>Gscript</i> | GDVESETYKTACETVCEAPLASEWEVQDGT PFGGIVGDEEYS ENTWNE DR |
| <i>Dbaeomy</i> | SYSEKALLRLNEKESDKRNLGDDAIRRMWRGPTIKGGTSFGIRPEESAHL |
| <i>Astriga</i> | EDFDQKERVASQASRNIEPTNSRRSSNSAVRDQAQSTGSRSSKTTTESGI |
| <i>Arubroc</i> | DLLRQRWRGDTVKGGA CLDMSAE EEEFDEKEGVTTEASETIRPSQATSTA |
| <i>Epallid</i> | -----MMKANAPLSAQEINEEEGVDAED |
| <i>Pcy lind</i> | ----- |
| <i>Ncrassa</i> | ----- |
| | |
| <i>Cgrayi</i> | VGTL LGVHDLNPIRATSSNAPNRTRAVPFEATTRHATIDSDKEDAVMLVS |
| <i>Lsp</i> | ----- |
| <i>Gscript</i> | ISVIHTHEQIHRDQTSRRQTIHGTLMPARARVHENMDTFPERYRKPQGHK |
| <i>Dbaeomy</i> | KHEHSEEEERARSLVQQLSQKPSASPSKGPVSYVVDISTDDDEIQELRY |
| <i>Astriga</i> | FSTSMSSSSTTG NRPHFRVLEHDTETS VSLFVQGNSPDQASGNRRQAAR |
| <i>Arubroc</i> | RKSSKTIT EFGSSSETLSASTSSSFPTGNRTHPYTVLNHETETLPDWA AE |
| <i>Epallid</i> | VVTEDFIVDHLRSLVIDPSTETSSRSVGENFWPGQRVADTVNIKGLNLRR |
| <i>Pcy lind</i> | ----- |
| <i>Ncrassa</i> | ----- |
| | |
| <i>Cgrayi</i> | SDPDKKTDMASAPIQPGVRNKRKRSPTFTFDEDYDVSDQESRKDSTIIVD |
| <i>Lsp</i> | ----- |
| <i>Gscript</i> | NSEGEMDTLPCRSWDEL PFTLFRDEEAQCRRIDLVNHRDVELTGGSFRR |
| <i>Dbaeomy</i> | LPTRVTDFQWIERRCSQGSPPRRRSRSLEPPERQIRSNISPDRNFDKRH |

Astriga HQIVIDITGEEESPKIISLNPLSLGESRGQSSKRHDHSPENETPQSSWHYNTK
Arubroc NQRQAVRHQIVIDITGDESPKISSNPLSLDQSRGQSSKRRYSHEDGTSQT
Epallid GLTVELQNGSFLWIDSLKNNIWGPVTIKGYKLARDGYCGAKLPSGRLNEL
Pcy lind -----
Ncrassa -----

Cgrayi LTQSLRESSIQPRRQERKPSKVNKINAKGESWSNSGIDMERTENMTTARS
Lsp -----MARKLNETCWIQNVFLDGTPEPNDDVPLGAVMR
Gscript ITMIVGNDRTGDVWIKGRKMIRSKSMAPVIGGLKNQLVWIAEVHRSLTEK
Dbaeomy LSLAPSARSPPHFVDSLSEDEIKGARSNSISGSPPKLRTSAYKEDVPLRR
Astriga RQRTGPEGFTRQEPSPDIEIEISAKVTAKSSGVYTTKYIGPLPSSHSSCIH
Arubroc LWHHNNQSQVASAASSSDIEIEISAQVTANYIGALRNSYAPHPQFTKRLKT
Epallid VWINEIDEEGHQVGLSVQQEVQVSDVKRVRDVVYTNCPPWPMFSSHDEIT
Pcy lind -----QCNGNISTWRSLETSP
Ncrassa -----

Cgrayi HPQSGISSERQAVLQSFNTPPGYSNRKREEVRRALMDRWEPKVEVIDLSR
Lsp LRQLRMTNRSFSSFNITD TDGNTGSKAQREAEGLFCRFRFC--SVYESQ
Gscript LSGKNGEIWIRPSRIRKLRKIVFSNLPTENWQRADDFEFNANAVLRCSW
Dbaeomy RLPLDMQPVSLESETQVGRMSFVSHD VDDGKPLIGTHR PSTVSIDDDS
Astriga STKRIKTTHRETGMSSIMDPAPPLSDARLEHLGKRWPVSVTLGQRRAS
Arubroc TGEATPTSPTMGLTNPLSDTGHGHSGLWPAKSESLGQRPASAKSRDQTL
Epallid GRRLTQTATNNAEAESEHGKLYCRWKFIQVKSRLKNDDEACL-----TLLS
Pcy lind KKQRREVIQCPTPPSSCGSSRVEVRGSRGEDPNRRSVYITP---SSNRSP
Ncrassa -----MVCRYSYVEYRHVDKPREWSFIRVEE-NEADEGFRL---SDDVLV

Cgrayi PHTTSFSSPKPGT LAPLQSI--SRFATQFSSATLAPPRNITLTPRRRLPA
Lsp AQKSKQSYANFSEKYITCLDPEDIDPGWSVDLPNASRLH-RDISRSSSAT
Gscript KFFSLYEHDTKHRRVFS E MCVLRMMACEADEGHCEGDSVVQEKFRRLGRQ
Dbaeomy DAMSQRKHSIISINDDEPTGCVATPKTIPPRDIVVESPP-EIDNLEVTTG
Astriga DLTDQIASSGTALKEPGQKTPAESSTISALTTKQSLGQK-ALSKQNPST
Arubroc SFGIALKEPEQKISVQATAFSA LTAKQSHRQRPRRRRV-SSNKPRQKQR
Epallid PEEAIDKAGIEPSLLRKAWR-----GDEDPIPGGSSTMP-AFDVETGKMT
Pcy lind SRKRPFNDRPIPFRLKHA---QKGGKLVTLTPDIEIEI-P-SEQVDLITE-
Ncrassa NGWRGGKVPGGSFLPAGQ---EHGHGHVHNVDDLSTLTP-SLTARGPKQ-

Cgrayi -RPFYRRYTFGDCFCGAGGM-SRGAVNAGLRIAWGDFDNL PACRSYQLN
Lsp LVDDDEIPYTFGDVFCGAGGV-SSGAKQAGLKVVWGLERDIDAARSFNQN
Gscript HIEQPGR-YTFADGFCCSGGA-TTGALAAGAVFDRAFDKDP AACASYERN
Dbaeomy VSGVGQR-YTFADAFCGAGGT-SRGACMAGLRVQWGLDHNEHA CLSWSQN
Astriga KPRHKQR-YTFGDAFCGAGGT-SRGAKSAGFRVEWGFDFDLAAINSYRLN
Arubroc -----YTFGDAFCGAGGT-SRGAKSAGFRVQWGFDFDFAAITSYSLN
Epallid SIPSYSL----GDCFCGAGGV-SRGAIQAGLQVAWGFVDVPEAIK AHAEN
Pcy lind -IRPNST-YSLGDCFCGAGGM-SCAHSAGYDVAWAFDHDYAAITSYGLN
Ncrassa -VPSDQK-YTAGDTFAGAGGA-SRGITDAGVHLEFCVDNWEHAVASLNAN

Cgrayi FLG--TPIYHVWANEF SNCKRDLK----CDICHLSPPCQFFSPL-HTSNG

Lsp FPEAVCEVASVEQFVHFADYD-----VDIMHISPPCPSFSPM-QIIPG
Gscript FGEICFVADTHKYCDLVRELGLP---PVAIAHYSPPCKTWSAA-HTHIG
Dbaeomy FPFANILQRDI-FEVITRQSGLSR---VDILHLSPPCQFFSPA-HTTAG
Astriga FYDTGCIVAWAHEFAEISVRDQN---KVDVLHLSPPCQVFSWA-HTVDG
Arubroc FYNTDCIVAWAHQFAEISVGDKH---KVDILHLSPPCQVFSWA-HTVDG
Epallid FAAYGTKSLELTDSAIIDLIKKNPRKFHVVDIAHYSPPCQPFSSANHNKNV
Pcy lind HGGTQRHLIEADQFVFI EGENLK-----VDVMHLSPPCQFFSPA-HTMQG
Ncrassa FQGQDTTTYDIDMHNFI VNK EIRHR---VDILHLSPPCQVWSPA-HTRPG

Cgrayi KNDELNTASLFAIEALLK KAKPRVATLEETAGLVKISQHIDYFNAVINMF
Lsp KNFEEQQVVILSIEELVKRIRPRILTMEETFGLLFERNMDF FATVVR-S-L
Gscript KDDEANEATLLALSELLMVDRPRIT TLENTSG-LWTHQPDMYGM LGQ-Y
Dbaeomy QNDEANTAASFAVAGMLQFAKPRIVTLENTSG-LLRQHQVWLYAVINQ-F
Astriga KDDDMNSATFFAVEELVKRTKPRVVTLENTSG-LPILYPRWFNSAIHF-F
Arubroc KDDEMNTATFFAVEELVKRIKPRVVTLENTSG-LPVLHPLWLN SAIHI-F
Epallid DRDFLNQKALFSLHDLTQLLKPRIATIEETAG-LMHRHGEWFNALVHI-F
Pcy lind QNDERNTAAQFAVPELIKPCRPRIVTLEQTFG-LFQRHRYWFNMIQF-F
Ncrassa QNDERNLAILFSCTHLIEKIRPRLFTVEQTFGILHPRLDNFFQSLVHG-F

Cgrayi TSQGF SVRWKVLCSADYGV PQQRKRLV I IASC-----PGEVLPFFPA
Lsp NDFGYSVRWKILNLRHYGAPQPRRLII IAARYEHVRNR-PGETLPSFPP
Gscript TSLDFSVRFVKNLQYGV PQKRKRF LII GACDGF TDRSSPGEPLPPFPI
Dbaeomy TSLGFSIRWKLLCFQDYGLGQRRERLII IASC-----PGEPLPTFPP
Astriga TRLGFSVRWKIMNFAEYGLVQARKRLII IASC-----PGEGLPVLPT
Arubroc TRLGFSIRWKIMNFAEYGLVQARKRLII IASI-----PHGFANHDPD
Epallid TNLGYSIRWKIVRCNEHGIPQNRVRL LLM AAA-----PGESLPKFPR
Pcy lind TDLGFSIRWKVVNFAEYGLPQARKRLII FASC-----PGEALPPFPK
Ncrassa TDHGYSVRWKVVNF SHYGLPQPRRLIMIGAG-----PGEKLPFFPS

Cgrayi PTHAKDPAGTGLKPWTTVNKA-INSI-PTDWNHDVDGANRVF-----KE
Lsp PTHLNN-----PNTIADLISDLPP---TSTDPLHNPTAKLFLVPRDPY
Gscript ETHSLSECVTLGQAVEKIELLHNLGFEIPD-----HTPAESTGVY
Dbaeomy RTHSSNPPQTGLKPHTTMSNA-LQNI---PHGWSHHPDPSVTRR--DLPP
Astriga PTHGAPSS--GLAPWVTVDVLSRI---PYGFANHDPDSVRK----RNEP
Arubroc SAGQRNE-----PPYRA-----
Epallid ATHG-EKGS-GLKPYVKIRDV-IYNI----PLTAKNQNLSESKKPYLLT
Pcy lind PTHS-G---DHRQRSKYGLARLTTVNDAIRNIPHDAPNHG PENSPRMSLS
Ncrassa PTHGN----GLKPVTTARQA-LAAI-DGRRRYPLHQPYLQPF--PTRKA

Cgrayi PYS-GDRTANCITTTGNGA--YHPSGTRPFTHRELACLQEFPLGHKFGSV
Lsp NAH-GPLREIIACSPTS G--NYHPSGLRPFTVREFGCLQTFPKDFRFVGT

Gscript DDKIYSEVAKSITTSASE--NWHYNG-RQLTVRELAI IQTFPIKHLFGEG
Dbaeomy TKY-GDSLKCMTTSGAG-- I IHPSGKRALTNREFACLQGFPLVHTFSDR
Astriga PYRADQPLRACITCSGGG--NYHPSGKRDLTLREYACLQGFPLEHKFGLA
Arubroc ----DQVLRACITCSGGG--NYHPSGKRDFTLREYACLQGFPLEHKFGLT
Epallid PFS-DDSFAKCITTSGGQF--NHHPSGERRYNVREMAQLQGFPTNHAFNNH
Pcyclind RYCGDMPLPRCITTSGGG--NAHPDGTRKF TLREFACLQGFPIDYKFSAT
Ncrassa HWD-GDKPLPYTVTCGAAE--NYHWSGLRQFTPQEYALLQGFPMHKKFAGN

Cgrayi EVRK----QIGNAVPPTLATVILEEVKKTLMKADGL-----
Lsp -----
Gscript ISKTAKKSQIGNAFPPYFTWQMLHQVVQSLRRADGLPASNL PFDHNAARH
Dbaeomy EVKK----QIGNAVSPPFAKILLGHIRKELEKVDGIPSRHEEEIIEV---
Astriga RVKK----QIGNAVPPLFAKILFEEIKKTLLKADGF-----
Arubroc RVKK----QIGNAVPPLFARVLFEEIKKSLLKADGF-----
Epallid SITIATR-QVGNVPPTLAKPWLEGI I KSLRETDMKQKMR-----
Pcyclind GVLK----QIGNAVPPSVGAVFLAEVKKALMKADGLL-----
Ncrassa YIKK----QIGNAFPPIFVKLLYKHLVECLDKRDNIIRQAQARTEEAAPF

Cgrayi -----
Lsp -----
Gscript TLKHKWTSEKVDIVHGSPFKIVENKDSPRVRGVPYDHLLPRALGNTTED
Dbaeomy -----
Astriga -----
Arubroc -----
Epallid -----
Pcyclind -----
Ncrassa QTPRKLVGSVGRNEVGSVGRNEEDDGEVTFLSSRKRRRQFAAVEVIDNNN

Cgrayi -----
Lsp -----
Gscript EGEVAGRAPSKDANNEVFILD-----
Dbaeomy -----
Astriga -----
Arubroc -----
Epallid -----
Pcyclind -----
Ncrassa RISTHRSKRARRLVGNTPPCQAAAAQAKQKTIIDLDEDISNLDLDQDRD

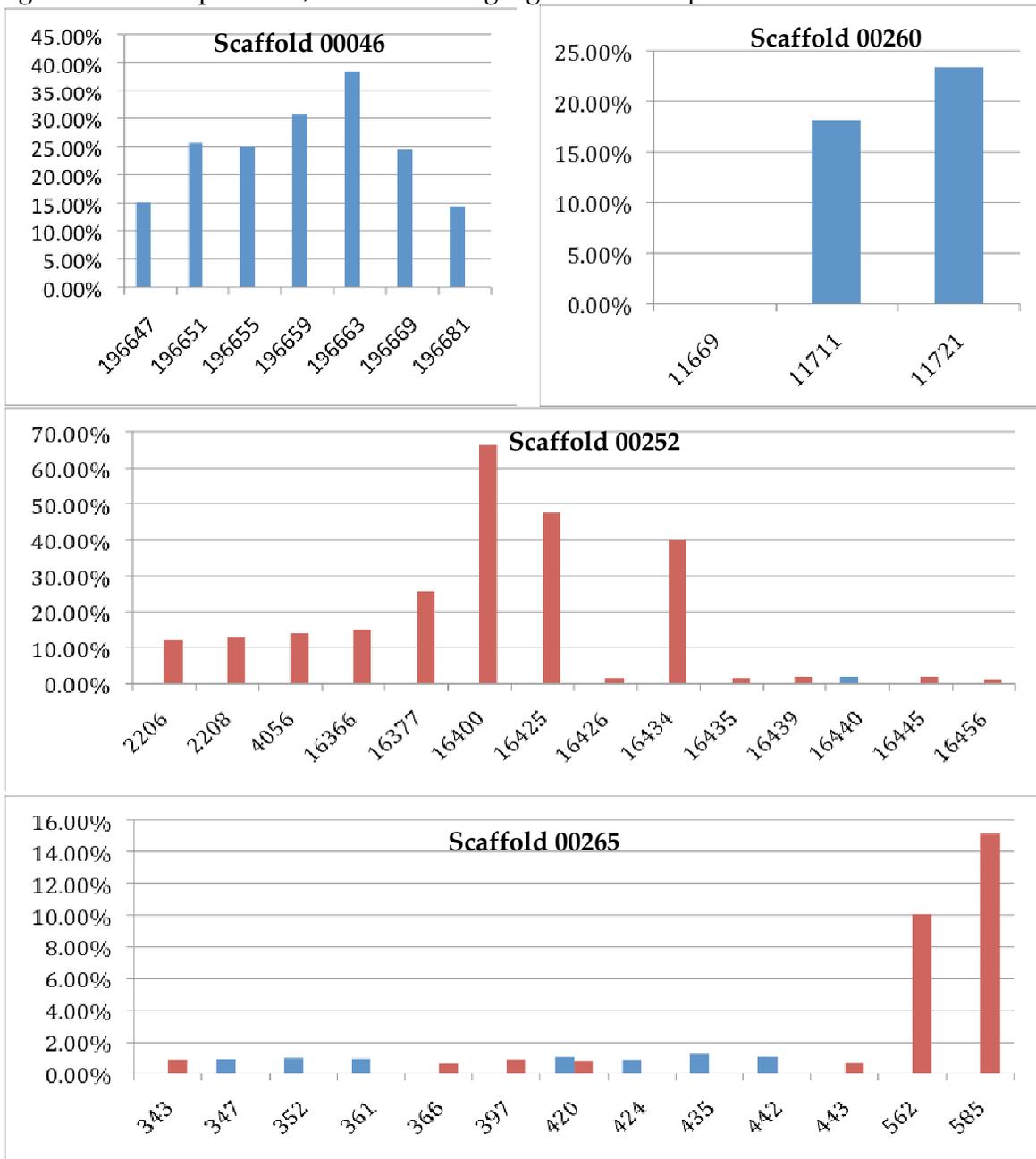
Cgrayi -----
Lsp -----
Gscript -----
Dbaeomy -----
Astriga -----

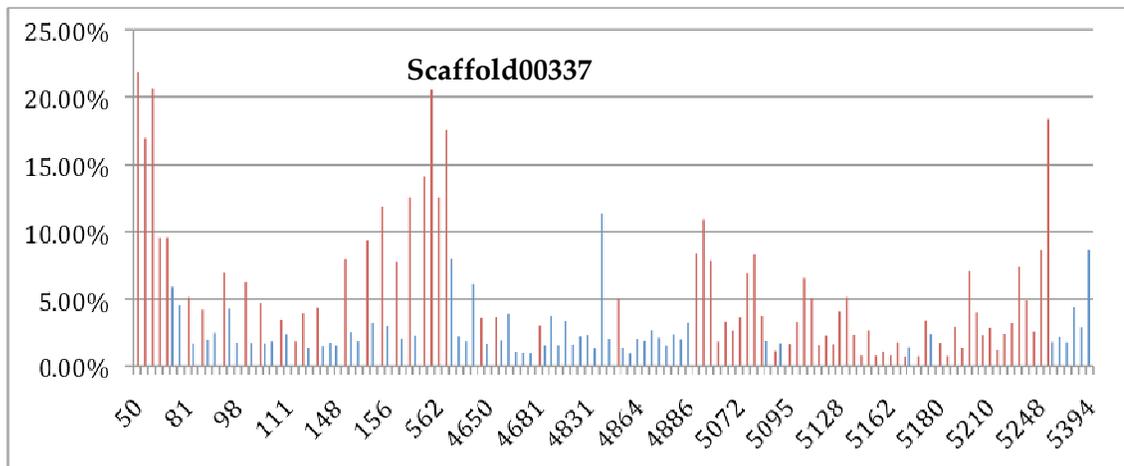
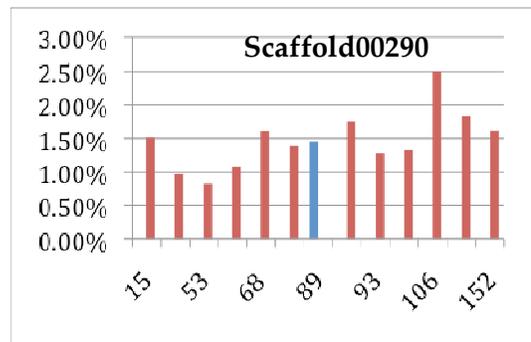
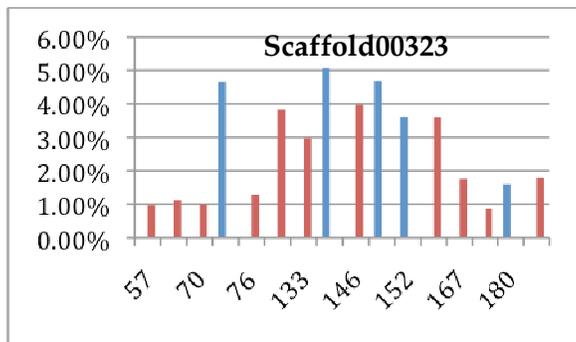
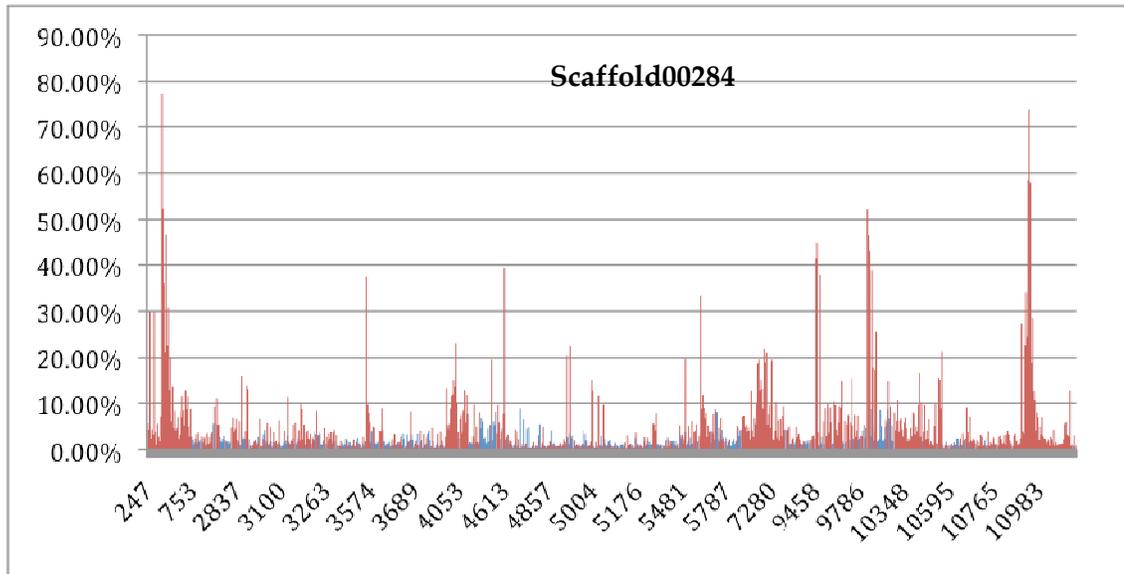
Arubroc -----
Epallid -----
Pcy lind -----
Ncrassa DDRSDTATIRESSVEVEASPRVSPVRRHPAAPPSSHLPLGPAGLTNPIQG

Cgrayi -----
Lsp -----
Gscript -----
Dbaeomy -----
Astriga -----
Arubroc -----
Epallid -----
Pcy lind -----
Ncrassa GPSTSTSTRTGSKASSSSQQQTHHNNMQGETVRRKLFFAVPPSRTEPFSS
Cgrayi -----
Lsp -----
Gscript -----
Dbaeomy -----
Astriga -----
Arubroc -----
Epallid -----
Pcy lind -----
Ncrassa PSSTSSTSSSTTASSSAAGSSNGSNSSSPVVKKENQKGTRKEPMELFDD-

Appendix I. Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*

X axis = coordinates of the methylated position within the contig. Y axis = percentage of reads covering a site indicating methylation at that site. Blue bars = fungal genome from squamules; Red bars = fungal genome from podetia.





Appendix J. Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*. Data is summarized in the graphs shown in Appendix I.

Table 13: Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00046 | 196647 | 15.15% | 0 |
| scaffold00046 | 196651 | 25.64% | 0 |
| scaffold00046 | 196655 | 25.00% | 0 |
| scaffold00046 | 196659 | 30.68% | 0 |
| scaffold00046 | 196663 | 38.54% | 0 |
| scaffold00046 | 196669 | 24.47% | 0 |
| scaffold00046 | 196681 | 14.49% | 0 |
| | | | |
| scaffold00252 | 2206 | 0 | 12.20% |
| scaffold00252 | 2208 | 0 | 12.99% |
| scaffold00252 | 4056 | 0 | 13.92% |
| scaffold00252 | 16366 | 0 | 15.15% |
| scaffold00252 | 16377 | 0 | 25.64% |
| scaffold00252 | 16400 | 0 | 66.67% |
| scaffold00252 | 16425 | 0 | 47.62% |
| scaffold00252 | 16426 | 0 | 1.46% |
| scaffold00252 | 16434 | 0 | 40 |
| scaffold00252 | 16435 | 0 | 1.46% |
| scaffold00252 | 16439 | 0 | 1.89% |
| scaffold00252 | 16440 | 1.98% | 0 |
| scaffold00252 | 16445 | 0 | 2.06% |
| scaffold00252 | 16456 | 0 | 1.08% |
| | | | |
| scaffold00260 | 11669 | 0 | 13.19% |
| scaffold00260 | 11711 | 18.18% | 0 |
| scaffold00260 | 11721 | 23.40% | 11721 |
| | | | |
| scaffold00265 | 343 | 0 | 0.87% |
| scaffold00265 | 347 | 0.96% | 0 |
| scaffold00265 | 352 | 1.01% | 0 |
| scaffold00265 | 361 | 0.98% | 0 |
| scaffold00265 | 366 | 0 | 0.64% |
| scaffold00265 | 397 | 0 | 0.88% |
| scaffold00265 | 420 | 1.07% | 0.79% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00265 | 424 | 0.86% | 0 |
| scaffold00265 | 435 | 1.27% | 0 |
| scaffold00265 | 442 | 1.08% | 0 |
| scaffold00265 | 443 | 0 | 0.67% |
| scaffold00265 | 562 | 0 | 10.07% |
| scaffold00265 | 585 | 0 | 15.12% |
| scaffold00284 | 247 | 0 | 5.67% |
| scaffold00284 | 364 | 0 | 4.27% |
| scaffold00284 | 368 | 0 | 30 |
| scaffold00284 | 375 | 0 | 2.37% |
| scaffold00284 | 376 | 0 | 4.39% |
| scaffold00284 | 382 | 0 | 3.26% |
| scaffold00284 | 384 | 0 | 30 |
| scaffold00284 | 387 | 0 | 3.80% |
| scaffold00284 | 388 | 0 | 1.06% |
| scaffold00284 | 389 | 0 | 5.58% |
| scaffold00284 | 390 | 0 | 2.75% |
| scaffold00284 | 391 | 0 | 1.79% |
| scaffold00284 | 392 | 0 | 7.13% |
| scaffold00284 | 393 | 0 | 77.14% |
| scaffold00284 | 397 | 0 | 52.33% |
| scaffold00284 | 399 | 0 | 36.00% |
| scaffold00284 | 407 | 0 | 21.30% |
| scaffold00284 | 409 | 0 | 46.85% |
| scaffold00284 | 410 | 0 | 22.58% |
| scaffold00284 | 411 | 0 | 30.88% |
| scaffold00284 | 418 | 0 | 12.71% |
| scaffold00284 | 421 | 0 | 20.16% |
| scaffold00284 | 422 | 0 | 6.06% |
| scaffold00284 | 423 | 0 | 13.45% |
| scaffold00284 | 424 | 0 | 4.53% |
| scaffold00284 | 425 | 0 | 8.30% |
| scaffold00284 | 428 | 0 | 4.12% |
| scaffold00284 | 430 | 0 | 4.64% |
| scaffold00284 | 438 | 0 | 6.93% |
| scaffold00284 | 439 | 0 | 2.33% |
| scaffold00284 | 441 | 0 | 3.15% |
| scaffold00284 | 477 | 2.52% | 8.27% |
| scaffold00284 | 506 | 0 | 6.94% |
| scaffold00284 | 523 | 0 | 11.37% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 524 | 0 | 8.37% |
| scaffold00284 | 525 | 0 | 5.13% |
| scaffold00284 | 526 | 0 | 12.70% |
| scaffold00284 | 536 | 0 | 8.70% |
| scaffold00284 | 545 | 0 | 11.36% |
| scaffold00284 | 742 | 0 | 4.88% |
| scaffold00284 | 744 | 0 | 2.71% |
| scaffold00284 | 753 | 0 | 8.72% |
| scaffold00284 | 753 | 2.73% | 0 |
| scaffold00284 | 761 | 2.13% | 0 |
| scaffold00284 | 762 | 1.14% | 0 |
| scaffold00284 | 766 | 1.46% | 2.45% |
| scaffold00284 | 769 | 0 | 3.42% |
| scaffold00284 | 769 | 1.43% | 0 |
| scaffold00284 | 772 | 1.60% | 3.68% |
| scaffold00284 | 772 | 1.60% | 0 |
| scaffold00284 | 777 | 2.17% | 0 |
| scaffold00284 | 779 | 0 | 2.78% |
| scaffold00284 | 780 | 1.60% | 0.53% |
| scaffold00284 | 783 | 1.44% | 2.55% |
| scaffold00284 | 789 | 0 | 2.58% |
| scaffold00284 | 790 | 1.04% | 0.63% |
| scaffold00284 | 793 | 0 | 3.59% |
| scaffold00284 | 796 | 0 | 1.24% |
| scaffold00284 | 799 | 0 | 2.68% |
| scaffold00284 | 806 | 0 | 1.40% |
| scaffold00284 | 810 | 0 | 3.58% |
| scaffold00284 | 818 | 0 | 3.72% |
| scaffold00284 | 1014 | 5.58% | 0 |
| scaffold00284 | 1023 | 0 | 9.26% |
| scaffold00284 | 1023 | 5.71% | 0 |
| scaffold00284 | 1117 | 0 | 11.05% |
| scaffold00284 | 1123 | 0 | 5.36% |
| scaffold00284 | 1153 | 0 | 5.36% |
| scaffold00284 | 2757 | 2.65% | 0 |
| scaffold00284 | 2771 | 1.78% | 0 |
| scaffold00284 | 2773 | 1.55% | 0 |
| scaffold00284 | 2789 | 2.30% | 0 |
| scaffold00284 | 2790 | 2.93% | 0 |
| scaffold00284 | 2796 | 0 | 1.96% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 2801 | 1.75% | 0 |
| scaffold00284 | 2806 | 1.49% | 0 |
| scaffold00284 | 2818 | 1.81% | 0 |
| scaffold00284 | 2820 | 1.30% | 0 |
| scaffold00284 | 2822 | 1.66% | 0 |
| scaffold00284 | 2828 | 0 | 4.71% |
| scaffold00284 | 2829 | 0 | 6.86% |
| scaffold00284 | 2830 | 0 | 3.81% |
| scaffold00284 | 2834 | 0 | 4.42% |
| scaffold00284 | 2836 | 0 | 6.55% |
| scaffold00284 | 2837 | 0 | 2.08% |
| scaffold00284 | 2837 | 1.49% | 0 |
| scaffold00284 | 2841 | 1.71% | 6.15% |
| scaffold00284 | 2847 | 1.00% | 0 |
| scaffold00284 | 2854 | 0 | 16.09% |
| scaffold00284 | 2858 | 2.17% | 0 |
| scaffold00284 | 2865 | 2.54% | 0 |
| scaffold00284 | 2870 | 2.34% | 0 |
| scaffold00284 | 2884 | 2.32% | 3.96% |
| scaffold00284 | 2924 | 0 | 13.70% |
| scaffold00284 | 2925 | 0 | 12.99% |
| scaffold00284 | 2958 | 0 | 2.49% |
| scaffold00284 | 2972 | 2.25% | 0 |
| scaffold00284 | 2979 | 0 | 0.92% |
| scaffold00284 | 3010 | 0.78% | 0.73% |
| scaffold00284 | 3013 | 0.54% | 0.34% |
| scaffold00284 | 3018 | 1.02% | 0.80% |
| scaffold00284 | 3019 | 0.56% | 1.71% |
| scaffold00284 | 3029 | 1.69% | 2.14% |
| scaffold00284 | 3032 | 0.70% | 0.84% |
| scaffold00284 | 3051 | 1.29% | 1.48% |
| scaffold00284 | 3054 | 2.69% | 0 |
| scaffold00284 | 3058 | 0 | 6.57% |
| scaffold00284 | 3064 | 0.63% | 0.76% |
| scaffold00284 | 3066 | 0.66% | 0 |
| scaffold00284 | 3067 | 0.80% | 2.30% |
| scaffold00284 | 3068 | 3.30% | 0 |
| scaffold00284 | 3070 | 4.23% | 1.90% |
| scaffold00284 | 3073 | 1.30% | 0 |
| scaffold00284 | 3075 | 1.64% | 5.68% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 3076 | 1.03% | 0 |
| scaffold00284 | 3079 | 1.24% | 0.92% |
| scaffold00284 | 3080 | 1.69% | 4.81% |
| scaffold00284 | 3080 | 1.69% | 0 |
| scaffold00284 | 3082 | 0.66% | 0 |
| scaffold00284 | 3084 | 1.57% | 3.79% |
| scaffold00284 | 3087 | 0.99% | 1.45% |
| scaffold00284 | 3090 | 1.53% | 1.85% |
| scaffold00284 | 3091 | 0 | 1.21% |
| scaffold00284 | 3092 | 0.98% | 0 |
| scaffold00284 | 3093 | 1.58% | 6.29% |
| scaffold00284 | 3098 | 1.32% | 0 |
| scaffold00284 | 3100 | 0.66% | 0 |
| scaffold00284 | 3103 | 0.81% | 0 |
| scaffold00284 | 3104 | 0.94% | 0 |
| scaffold00284 | 3106 | 1.51% | 4.02% |
| scaffold00284 | 3107 | 1.20% | 0 |
| scaffold00284 | 3109 | 1.95% | 0 |
| scaffold00284 | 3110 | 3.03% | 11.31% |
| scaffold00284 | 3111 | 0 | 1.50% |
| scaffold00284 | 3114 | 1.22% | 0 |
| scaffold00284 | 3124 | 1.17% | 1.92% |
| scaffold00284 | 3128 | 0 | 1.83% |
| scaffold00284 | 3129 | 2.42% | 5.36% |
| scaffold00284 | 3140 | 1.24% | 0 |
| scaffold00284 | 3145 | 0 | 5.73% |
| scaffold00284 | 3147 | 1.15% | 0 |
| scaffold00284 | 3152 | 1.53% | 0 |
| scaffold00284 | 3158 | 2.45% | 4.16% |
| scaffold00284 | 3160 | 0 | 1.77% |
| scaffold00284 | 3165 | 1.99% | 10.09% |
| scaffold00284 | 3167 | 2.66% | 8.69% |
| scaffold00284 | 3170 | 0 | 2.17% |
| scaffold00284 | 3175 | 1.50% | 5.41% |
| scaffold00284 | 3177 | 1.14% | 0 |
| scaffold00284 | 3180 | 1.03% | 3.71% |
| scaffold00284 | 3182 | 1.64% | 4.08% |
| scaffold00284 | 3188 | 0 | 2.42% |
| scaffold00284 | 3191 | 0 | 2.17% |
| scaffold00284 | 3200 | 0 | 3.54% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 3206 | 1.22% | 0 |
| scaffold00284 | 3208 | 0 | 2.51% |
| scaffold00284 | 3211 | 0 | 2.02% |
| scaffold00284 | 3217 | 0 | 3.92% |
| scaffold00284 | 3219 | 1.95% | 0 |
| scaffold00284 | 3230 | 3.35% | 0 |
| scaffold00284 | 3235 | 0 | 8.30% |
| scaffold00284 | 3236 | 3.09% | 3.09% |
| scaffold00284 | 3246 | 3.09% | 0 |
| scaffold00284 | 3251 | 3.61% | 0 |
| scaffold00284 | 3256 | 1.09% | 0 |
| scaffold00284 | 3258 | 1.07% | 0 |
| scaffold00284 | 3261 | 0 | 1.92% |
| scaffold00284 | 3262 | 1.92% | 4.59% |
| scaffold00284 | 3263 | 1.23% | 0 |
| scaffold00284 | 3268 | 0 | 2.67% |
| scaffold00284 | 3269 | 1.17% | 0 |
| scaffold00284 | 3271 | 1.94% | 3.56% |
| scaffold00284 | 3274 | 0 | 2.37% |
| scaffold00284 | 3283 | 0 | 3.75% |
| scaffold00284 | 3321 | 1.32% | 0 |
| scaffold00284 | 3322 | 2.19% | 3.70% |
| scaffold00284 | 3324 | 0 | 4.30% |
| scaffold00284 | 3327 | 1.80% | 0 |
| scaffold00284 | 3340 | 0 | 3.14% |
| scaffold00284 | 3351 | 0.92% | 0 |
| scaffold00284 | 3353 | 0.73% | 0 |
| scaffold00284 | 3354 | 0 | 0.81% |
| scaffold00284 | 3355 | 0 | 1.89% |
| scaffold00284 | 3358 | 0.69% | 0.75% |
| scaffold00284 | 3359 | 1.42% | 0 |
| scaffold00284 | 3361 | 0.87% | 0 |
| scaffold00284 | 3367 | 1.10% | 0 |
| scaffold00284 | 3369 | 2.38% | 1.73% |
| scaffold00284 | 3378 | 2.01% | 0.65% |
| scaffold00284 | 3383 | 0.78% | 0.54% |
| scaffold00284 | 3384 | 1.54% | 1.06% |
| scaffold00284 | 3384 | 1.54% | 0 |
| scaffold00284 | 3386 | 0.77% | 0 |
| scaffold00284 | 3402 | 0.88% | 0 |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 3403 | 1.12% | 2.15% |
| scaffold00284 | 3407 | 0.64% | 0.62% |
| scaffold00284 | 3418 | 0.73% | 0.46% |
| scaffold00284 | 3426 | 1.71% | 1.33% |
| scaffold00284 | 3429 | 2.47% | 0.93% |
| scaffold00284 | 3437 | 1.37% | 0.59% |
| scaffold00284 | 3438 | 0.46% | 1.18% |
| scaffold00284 | 3456 | 0.48% | 0 |
| scaffold00284 | 3465 | 0.66% | 0 |
| scaffold00284 | 3475 | 1.37% | 2.70% |
| scaffold00284 | 3477 | 0.71% | 0 |
| scaffold00284 | 3483 | 1.77% | 1.47% |
| scaffold00284 | 3542 | 0 | 37.50% |
| scaffold00284 | 3564 | 0 | 9.63% |
| scaffold00284 | 3572 | 1.98% | 7.82% |
| scaffold00284 | 3573 | 1.62% | 5.88% |
| scaffold00284 | 3574 | 1.46% | 0 |
| scaffold00284 | 3576 | 3.94% | 0 |
| scaffold00284 | 3582 | 0 | 4.93% |
| scaffold00284 | 3584 | 0 | 4.66% |
| scaffold00284 | 3586 | 1.43% | 0 |
| scaffold00284 | 3595 | 1.19% | 0 |
| scaffold00284 | 3597 | 1.41% | 0 |
| scaffold00284 | 3601 | 1.55% | 0 |
| scaffold00284 | 3603 | 0 | 3.98% |
| scaffold00284 | 3606 | 1.69% | 3.97% |
| scaffold00284 | 3611 | 3.06% | 8.87% |
| scaffold00284 | 3614 | 3.40% | 5.25% |
| scaffold00284 | 3615 | 0 | 4.21% |
| scaffold00284 | 3621 | 1.66% | 0 |
| scaffold00284 | 3623 | 1.77% | 0 |
| scaffold00284 | 3626 | 0.88% | 0 |
| scaffold00284 | 3627 | 1.14% | 1.27% |
| scaffold00284 | 3628 | 1.05% | 0 |
| scaffold00284 | 3632 | 1.46% | 0 |
| scaffold00284 | 3637 | 1.62% | 0 |
| scaffold00284 | 3644 | 0.66% | 0.77% |
| scaffold00284 | 3645 | 0.80% | 2.27% |
| scaffold00284 | 3647 | 0.93% | 0 |
| scaffold00284 | 3648 | 1.91% | 3.42% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 3650 | 0.99% | 0 |
| scaffold00284 | 3652 | 1.12% | 0 |
| scaffold00284 | 3653 | 1.44% | 1.55% |
| scaffold00284 | 3657 | 1.50% | 0.53% |
| scaffold00284 | 3658 | 1.10% | 1.63% |
| scaffold00284 | 3659 | 0.80% | 2.74% |
| scaffold00284 | 3662 | 0.60% | 0.98% |
| scaffold00284 | 3663 | 3.53% | 0 |
| scaffold00284 | 3665 | 2.32% | 0 |
| scaffold00284 | 3666 | 0 | 0.58% |
| scaffold00284 | 3670 | 0.74% | 0.66% |
| scaffold00284 | 3673 | 3.28% | 0 |
| scaffold00284 | 3674 | 1.34% | 1.64% |
| scaffold00284 | 3675 | 1.85% | 0 |
| scaffold00284 | 3679 | 2.12% | 8.20% |
| scaffold00284 | 3680 | 3.36% | 0 |
| scaffold00284 | 3686 | 2.06% | 0 |
| scaffold00284 | 3688 | 0.62% | 0 |
| scaffold00284 | 3689 | 0 | 1.28% |
| scaffold00284 | 3690 | 0.67% | 1.91% |
| scaffold00284 | 3693 | 0.93% | 1.00% |
| scaffold00284 | 3697 | 3.48% | 0 |
| scaffold00284 | 3702 | 0 | 0.90% |
| scaffold00284 | 3703 | 1.50% | 4.04% |
| scaffold00284 | 3704 | 2.01% | 0 |
| scaffold00284 | 3706 | 0.48% | 0 |
| scaffold00284 | 3707 | 0.65% | 3.39% |
| scaffold00284 | 3709 | 1.15% | 1.00% |
| scaffold00284 | 3711 | 0.72% | 0.54% |
| scaffold00284 | 3714 | 1.29% | 2.56% |
| scaffold00284 | 3718 | 3.51% | 0 |
| scaffold00284 | 3720 | 4.02% | 0 |
| scaffold00284 | 3724 | 2.01% | 0 |
| scaffold00284 | 3731 | 0.62% | 0.69% |
| scaffold00284 | 3732 | 0 | 4.64% |
| scaffold00284 | 3732 | 1.85% | 0 |
| scaffold00284 | 3738 | 0.49% | 0.53% |
| scaffold00284 | 3738 | 0.49% | 0 |
| scaffold00284 | 3747 | 1.04% | 0.71% |
| scaffold00284 | 3748 | 1.48% | 3.44% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 3754 | 0 | 0.50% |
| scaffold00284 | 3758 | 0.74% | 0 |
| scaffold00284 | 3759 | 1.02% | 3.85% |
| scaffold00284 | 3762 | 1.14% | 1.89% |
| scaffold00284 | 3767 | 1.04% | 0 |
| scaffold00284 | 3770 | 0 | 0.63% |
| scaffold00284 | 3810 | 0 | 2.72% |
| scaffold00284 | 3815 | 0 | 13.18% |
| scaffold00284 | 3831 | 0 | 5.32% |
| scaffold00284 | 3832 | 0 | 4.44% |
| scaffold00284 | 3833 | 0 | 5.53% |
| scaffold00284 | 3836 | 0 | 10.26% |
| scaffold00284 | 3847 | 0 | 8.85% |
| scaffold00284 | 3871 | 0 | 11.48% |
| scaffold00284 | 3886 | 0 | 11.81% |
| scaffold00284 | 3889 | 0 | 15.20% |
| scaffold00284 | 3892 | 0 | 13.51% |
| scaffold00284 | 3947 | 0 | 23.17% |
| scaffold00284 | 3973 | 0 | 10.16% |
| scaffold00284 | 4030 | 0 | 3.74% |
| scaffold00284 | 4053 | 2.36% | 0 |
| scaffold00284 | 4054 | 0 | 6.51% |
| scaffold00284 | 4063 | 0 | 7.46% |
| scaffold00284 | 4075 | 0 | 3.59% |
| scaffold00284 | 4085 | 3.97% | 8.33% |
| scaffold00284 | 4088 | 0 | 12.75% |
| scaffold00284 | 4090 | 0 | 5.81% |
| scaffold00284 | 4111 | 3.72% | 11.69% |
| scaffold00284 | 4131 | 1.42% | 7.66% |
| scaffold00284 | 4140 | 1.25% | 0 |
| scaffold00284 | 4145 | 1.59% | 0 |
| scaffold00284 | 4149 | 1.38% | 0 |
| scaffold00284 | 4154 | 1.48% | 0 |
| scaffold00284 | 4167 | 1.16% | 2.82% |
| scaffold00284 | 4168 | 2.51% | 9.65% |
| scaffold00284 | 4181 | 1.16% | 0 |
| scaffold00284 | 4182 | 1.55% | 4.87% |
| scaffold00284 | 4194 | 1.19% | 1.56% |
| scaffold00284 | 4210 | 1.19% | 0 |
| scaffold00284 | 4218 | 1.57% | 8.00% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 4241 | 4.56% | 0 |
| scaffold00284 | 4292 | 6.71% | 0 |
| scaffold00284 | 4295 | 6.06% | 0 |
| scaffold00284 | 4314 | 2.39% | 0 |
| scaffold00284 | 4321 | 2.55% | 0 |
| scaffold00284 | 4332 | 1.46% | 0 |
| scaffold00284 | 4335 | 1.40% | 0 |
| scaffold00284 | 4338 | 4.47% | 0 |
| scaffold00284 | 4345 | 1.65% | 0 |
| scaffold00284 | 4359 | 5.31% | 0 |
| scaffold00284 | 4363 | 0 | 19.61% |
| scaffold00284 | 4364 | 2.00% | 0 |
| scaffold00284 | 4369 | 6.09% | 0 |
| scaffold00284 | 4377 | 5.25% | 0 |
| scaffold00284 | 4392 | 8.09% | 0 |
| scaffold00284 | 4503 | 0 | 3.57% |
| scaffold00284 | 4515 | 0 | 9.59% |
| scaffold00284 | 4528 | 0 | 5.84% |
| scaffold00284 | 4530 | 2.35% | 0 |
| scaffold00284 | 4533 | 0 | 3.81% |
| scaffold00284 | 4536 | 4.49% | 0 |
| scaffold00284 | 4545 | 0 | 7.69% |
| scaffold00284 | 4613 | 0 | 39.29% |
| scaffold00284 | 4661 | 2.17% | 0 |
| scaffold00284 | 4668 | 0 | 2.72% |
| scaffold00284 | 4673 | 0 | 3.25% |
| scaffold00284 | 4673 | 2.18% | 0 |
| scaffold00284 | 4686 | 1.33% | 1.88% |
| scaffold00284 | 4689 | 1.62% | 1.70% |
| scaffold00284 | 4691 | 1.02% | 0 |
| scaffold00284 | 4703 | 0.59% | 0 |
| scaffold00284 | 4704 | 0.94% | 0 |
| scaffold00284 | 4705 | 1.14% | 4.36% |
| scaffold00284 | 4714 | 1.77% | 0.73% |
| scaffold00284 | 4715 | 2.59% | 3.90% |
| scaffold00284 | 4730 | 0 | 0.59% |
| scaffold00284 | 4731 | 0.97% | 2.21% |
| scaffold00284 | 4736 | 0.41% | 0 |
| scaffold00284 | 4739 | 8.85% | 0 |
| scaffold00284 | 4751 | 0.54% | 0.34% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 4752 | 0.72% | 1.26% |
| scaffold00284 | 4756 | 6.43% | 0 |
| scaffold00284 | 4761 | 0.66% | 0 |
| scaffold00284 | 4766 | 1.15% | 1.27% |
| scaffold00284 | 4769 | 4.48% | 0 |
| scaffold00284 | 4770 | 0.43% | 0.33% |
| scaffold00284 | 4781 | 4.74% | 0 |
| scaffold00284 | 4786 | 0.46% | 0 |
| scaffold00284 | 4789 | 0.50% | 0.34% |
| scaffold00284 | 4796 | 0.76% | 0.49% |
| scaffold00284 | 4797 | 0.92% | 1.64% |
| scaffold00284 | 4801 | 0 | 0.34% |
| scaffold00284 | 4802 | 0.46% | 0.91% |
| scaffold00284 | 4805 | 0.73% | 0.29% |
| scaffold00284 | 4808 | 0.68% | 0.42% |
| scaffold00284 | 4811 | 3.09% | 2.07% |
| scaffold00284 | 4814 | 5.43% | 0 |
| scaffold00284 | 4821 | 0.50% | 0 |
| scaffold00284 | 4822 | 0.71% | 1.41% |
| scaffold00284 | 4823 | 0.85% | 4.84% |
| scaffold00284 | 4835 | 0.81% | 0.52% |
| scaffold00284 | 4850 | 0.76% | 0.58% |
| scaffold00284 | 4852 | 0.51% | 0.44% |
| scaffold00284 | 4856 | 0.83% | 0.52% |
| scaffold00284 | 4857 | 0.56% | 0.74% |
| scaffold00284 | 4858 | 0.94% | 1.72% |
| scaffold00284 | 4862 | 1.65% | 1.60% |
| scaffold00284 | 4867 | 4.12% | 0 |
| scaffold00284 | 4877 | 0.97% | 0.79% |
| scaffold00284 | 4882 | 1.35% | 0.58% |
| scaffold00284 | 4885 | 0.89% | 0.78% |
| scaffold00284 | 4889 | 0.60% | 0.55% |
| scaffold00284 | 4897 | 0.77% | 0.57% |
| scaffold00284 | 4898 | 0.73% | 0.81% |
| scaffold00284 | 4902 | 0 | 0.73% |
| scaffold00284 | 4904 | 1.39% | 1.24% |
| scaffold00284 | 4911 | 0.70% | 0 |
| scaffold00284 | 4914 | 0.55% | 0.90% |
| scaffold00284 | 4915 | 1.30% | 2.30% |
| scaffold00284 | 4916 | 0 | 1.62% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 4919 | 1.22% | 0 |
| scaffold00284 | 4927 | 2.82% | 20.37% |
| scaffold00284 | 4928 | 2.59% | 0 |
| scaffold00284 | 4934 | 1.16% | 0 |
| scaffold00284 | 4935 | 3.57% | 22.45% |
| scaffold00284 | 4936 | 2.49% | 0 |
| scaffold00284 | 4939 | 3.00% | 0 |
| scaffold00284 | 4953 | 0 | 0.53% |
| scaffold00284 | 4954 | 1.69% | 2.06% |
| scaffold00284 | 4958 | 0.74% | 0 |
| scaffold00284 | 4959 | 0.75% | 1.05% |
| scaffold00284 | 4965 | 0.75% | 0 |
| scaffold00284 | 4966 | 0.81% | 1.18% |
| scaffold00284 | 4970 | 0.60% | 0 |
| scaffold00284 | 4971 | 0.72% | 0.55% |
| scaffold00284 | 4972 | 1.11% | 0.70% |
| scaffold00284 | 4975 | 1.30% | 0.93% |
| scaffold00284 | 4976 | 4.15% | 4.13% |
| scaffold00284 | 4979 | 0.90% | 0 |
| scaffold00284 | 4980 | 3.43% | 0 |
| scaffold00284 | 4984 | 2.00% | 1.92% |
| scaffold00284 | 4989 | 0.71% | 0 |
| scaffold00284 | 4991 | 0.78% | 0.39% |
| scaffold00284 | 4992 | 0.85% | 0.72% |
| scaffold00284 | 4993 | 0.96% | 1.68% |
| scaffold00284 | 5000 | 0 | 15.07% |
| scaffold00284 | 5004 | 0 | 12.66% |
| scaffold00284 | 5008 | 0.76% | 0 |
| scaffold00284 | 5012 | 0.47% | 0.44% |
| scaffold00284 | 5013 | 1.40% | 1.60% |
| scaffold00284 | 5015 | 0.67% | 0.43% |
| scaffold00284 | 5016 | 0 | 11.54% |
| scaffold00284 | 5028 | 0.95% | 0 |
| scaffold00284 | 5029 | 0.77% | 0.88% |
| scaffold00284 | 5030 | 3.01% | 0 |
| scaffold00284 | 5039 | 0.88% | 0 |
| scaffold00284 | 5041 | 0 | 9.64% |
| scaffold00284 | 5042 | 2.33% | 0 |
| scaffold00284 | 5047 | 1.26% | 0 |
| scaffold00284 | 5052 | 0.72% | 0.52% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 5055 | 0.94% | 0.81% |
| scaffold00284 | 5058 | 1.73% | 0.68% |
| scaffold00284 | 5059 | 2.02% | 1.29% |
| scaffold00284 | 5063 | 0.97% | 0 |
| scaffold00284 | 5070 | 0.77% | 0 |
| scaffold00284 | 5073 | 0.60% | 0 |
| scaffold00284 | 5075 | 0.78% | 0.47% |
| scaffold00284 | 5077 | 1.40% | 1.40% |
| scaffold00284 | 5087 | 0.59% | 0 |
| scaffold00284 | 5092 | 0.89% | 0.51% |
| scaffold00284 | 5093 | 0.67% | 0 |
| scaffold00284 | 5095 | 0 | 0.39% |
| scaffold00284 | 5098 | 1.02% | 0.72% |
| scaffold00284 | 5102 | 1.00% | 0 |
| scaffold00284 | 5106 | 1.14% | 0.41% |
| scaffold00284 | 5108 | 0.66% | 0 |
| scaffold00284 | 5109 | 1.29% | 1.50% |
| scaffold00284 | 5113 | 0.50% | 0 |
| scaffold00284 | 5114 | 0 | 3.06% |
| scaffold00284 | 5118 | 0.77% | 0.79% |
| scaffold00284 | 5119 | 0.98% | 1.07% |
| scaffold00284 | 5122 | 1.10% | 1.24% |
| scaffold00284 | 5126 | 0.64% | 0.28% |
| scaffold00284 | 5128 | 1.05% | 0.52% |
| scaffold00284 | 5150 | 0.56% | 0.33% |
| scaffold00284 | 5150 | 0.56% | 0 |
| scaffold00284 | 5159 | 2.49% | 3.69% |
| scaffold00284 | 5174 | 0.90% | 0.44% |
| scaffold00284 | 5176 | 1.23% | 0.72% |
| scaffold00284 | 5190 | 0.62% | 0.43% |
| scaffold00284 | 5203 | 1.13% | 0.84% |
| scaffold00284 | 5205 | 0 | 0.73% |
| scaffold00284 | 5207 | 0.71% | 0.44% |
| scaffold00284 | 5213 | 0.53% | 0 |
| scaffold00284 | 5214 | 1.79% | 2.66% |
| scaffold00284 | 5220 | 0.81% | 0.58% |
| scaffold00284 | 5221 | 1.13% | 1.11% |
| scaffold00284 | 5225 | 1.85% | 2.59% |
| scaffold00284 | 5238 | 0.96% | 0 |
| scaffold00284 | 5241 | 1.59% | 0 |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 5253 | 7.04% | 0 |
| scaffold00284 | 5260 | 1.12% | 0 |
| scaffold00284 | 5265 | 1.12% | 0 |
| scaffold00284 | 5294 | 1.06% | 5.63% |
| scaffold00284 | 5296 | 0 | 5.46% |
| scaffold00284 | 5306 | 1.26% | 4.10% |
| scaffold00284 | 5307 | 1.06% | 7.78% |
| scaffold00284 | 5311 | 1.10% | 0 |
| scaffold00284 | 5346 | 1.15% | 0 |
| scaffold00284 | 5354 | 0.75% | 0 |
| scaffold00284 | 5363 | 1.12% | 0.86% |
| scaffold00284 | 5376 | 1.06% | 0 |
| scaffold00284 | 5412 | 0 | 0.66% |
| scaffold00284 | 5418 | 0.87% | 0.42% |
| scaffold00284 | 5422 | 0.86% | 0 |
| scaffold00284 | 5423 | 0 | 1.97% |
| scaffold00284 | 5428 | 0 | 0.64% |
| scaffold00284 | 5440 | 2.60% | 2.12% |
| scaffold00284 | 5442 | 0.63% | 0 |
| scaffold00284 | 5445 | 0.79% | 1.87% |
| scaffold00284 | 5448 | 1.74% | 1.23% |
| scaffold00284 | 5449 | 1.80% | 0 |
| scaffold00284 | 5455 | 1.85% | 0 |
| scaffold00284 | 5460 | 1.15% | 0.60% |
| scaffold00284 | 5461 | 1.82% | 0 |
| scaffold00284 | 5463 | 0.86% | 0.69% |
| scaffold00284 | 5469 | 0.54% | 1.27% |
| scaffold00284 | 5471 | 0.43% | 0 |
| scaffold00284 | 5472 | 3.53% | 5.23% |
| scaffold00284 | 5477 | 2.51% | 0 |
| scaffold00284 | 5481 | 3.21% | 3.12% |
| scaffold00284 | 5484 | 1.26% | 1.13% |
| scaffold00284 | 5485 | 2.04% | 2.57% |
| scaffold00284 | 5497 | 3.76% | 19.79% |
| scaffold00284 | 5499 | 0.85% | 0 |
| scaffold00284 | 5506 | 1.30% | 0 |
| scaffold00284 | 5511 | 1.84% | 5.05% |
| scaffold00284 | 5517 | 0.92% | 0 |
| scaffold00284 | 5520 | 2.56% | 0 |
| scaffold00284 | 5525 | 1.25% | 6.16% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 5526 | 1.28% | 0 |
| scaffold00284 | 5530 | 0 | 3.81% |
| scaffold00284 | 5531 | 1.74% | 3.99% |
| scaffold00284 | 5535 | 0 | 5.45% |
| scaffold00284 | 5536 | 1.81% | 0 |
| scaffold00284 | 5538 | 1.69% | 0 |
| scaffold00284 | 5548 | 0 | 2.86% |
| scaffold00284 | 5575 | 0 | 33.33% |
| scaffold00284 | 5580 | 0 | 8.78% |
| scaffold00284 | 5581 | 0 | 11.66% |
| scaffold00284 | 5584 | 0 | 8.92% |
| scaffold00284 | 5594 | 0 | 6.70% |
| scaffold00284 | 5670 | 0 | 7.84% |
| scaffold00284 | 5681 | 0 | 2.69% |
| scaffold00284 | 5682 | 0 | 5.12% |
| scaffold00284 | 5683 | 0 | 3.77% |
| scaffold00284 | 5699 | 0 | 3.94% |
| scaffold00284 | 5712 | 1.48% | 0 |
| scaffold00284 | 5715 | 3.90% | 3.61% |
| scaffold00284 | 5717 | 3.39% | 7.64% |
| scaffold00284 | 5719 | 2.18% | 0 |
| scaffold00284 | 5737 | 8.67% | 8.10% |
| scaffold00284 | 5742 | 8.05% | 0 |
| scaffold00284 | 5743 | 8.16% | 0 |
| scaffold00284 | 5756 | 1.31% | 0 |
| scaffold00284 | 5759 | 2.24% | 4.63% |
| scaffold00284 | 5760 | 0 | 1.77% |
| scaffold00284 | 5761 | 3.17% | 6.78% |
| scaffold00284 | 5768 | 4.26% | 1.73% |
| scaffold00284 | 5770 | 2.14% | 2.55% |
| scaffold00284 | 5779 | 1.50% | 0 |
| scaffold00284 | 5784 | 0 | 1.79% |
| scaffold00284 | 5787 | 2.51% | 0 |
| scaffold00284 | 5795 | 0.87% | 0 |
| scaffold00284 | 5802 | 1.57% | 0 |
| scaffold00284 | 5831 | 2.22% | 2.10% |
| scaffold00284 | 5833 | 0.85% | 0 |
| scaffold00284 | 5843 | 0.83% | 0 |
| scaffold00284 | 5846 | 1.70% | 0 |
| scaffold00284 | 5850 | 2.29% | 0 |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 5857 | 1.20% | 0 |
| scaffold00284 | 5859 | 1.75% | 0 |
| scaffold00284 | 5867 | 2.13% | 5.08% |
| scaffold00284 | 5870 | 4.62% | 0 |
| scaffold00284 | 5874 | 2.93% | 0 |
| scaffold00284 | 5879 | 4.15% | 0 |
| scaffold00284 | 5929 | 0 | 4.26% |
| scaffold00284 | 6343 | 0 | 7.07% |
| scaffold00284 | 6351 | 0 | 7.33% |
| scaffold00284 | 6363 | 0 | 4.77% |
| scaffold00284 | 6366 | 0 | 5.32% |
| scaffold00284 | 6626 | 0 | 3.84% |
| scaffold00284 | 6632 | 0 | 4.76% |
| scaffold00284 | 6638 | 0 | 4.03% |
| scaffold00284 | 6639 | 0 | 2.19% |
| scaffold00284 | 6642 | 0 | 12.63% |
| scaffold00284 | 6644 | 0 | 4.07% |
| scaffold00284 | 6646 | 0 | 2.65% |
| scaffold00284 | 6668 | 0 | 6.64% |
| scaffold00284 | 6670 | 0 | 5.50% |
| scaffold00284 | 6676 | 0 | 10 |
| scaffold00284 | 6697 | 0 | 18.75% |
| scaffold00284 | 6718 | 0 | 19.57% |
| scaffold00284 | 6721 | 0 | 12.63% |
| scaffold00284 | 6826 | 0 | 15.12% |
| scaffold00284 | 6847 | 0 | 13.08% |
| scaffold00284 | 6849 | 0 | 8.85% |
| scaffold00284 | 6868 | 0 | 21.95% |
| scaffold00284 | 6869 | 0 | 18.99% |
| scaffold00284 | 6874 | 0 | 21.13% |
| scaffold00284 | 7248 | 0 | 5.39% |
| scaffold00284 | 7274 | 0 | 7.55% |
| scaffold00284 | 7275 | 0 | 9.58% |
| scaffold00284 | 7277 | 0 | 4.62% |
| scaffold00284 | 7280 | 0 | 19.52% |
| scaffold00284 | 7302 | 0 | 2.19% |
| scaffold00284 | 7308 | 0 | 2.10% |
| scaffold00284 | 7310 | 0 | 4.08% |
| scaffold00284 | 7321 | 0 | 10.26% |
| scaffold00284 | 7332 | 0 | 2.52% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 7342 | 0 | 6.34% |
| scaffold00284 | 7351 | 0 | 1.97% |
| scaffold00284 | 7360 | 0 | 6.49% |
| scaffold00284 | 7362 | 0 | 2.83% |
| scaffold00284 | 7401 | 0 | 7.26% |
| scaffold00284 | 7420 | 0 | 9.32% |
| scaffold00284 | 7438 | 0 | 4.20% |
| scaffold00284 | 7584 | 0 | 8.00% |
| scaffold00284 | 7584 | 4.32% | 0 |
| scaffold00284 | 7611 | 0 | 4.14% |
| scaffold00284 | 7623 | 0 | 4.41% |
| scaffold00284 | 7625 | 0 | 4.92% |
| scaffold00284 | 9223 | 1.20% | 0 |
| scaffold00284 | 9264 | 1.23% | 0 |
| scaffold00284 | 9269 | 1.41% | 2.23% |
| scaffold00284 | 9277 | 0 | 2.50% |
| scaffold00284 | 9283 | 1.01% | 0 |
| scaffold00284 | 9288 | 1.46% | 0 |
| scaffold00284 | 9289 | 2.19% | 4.86% |
| scaffold00284 | 9296 | 0 | 2.65% |
| scaffold00284 | 9311 | 0.98% | 3.46% |
| scaffold00284 | 9317 | 0 | 2.18% |
| scaffold00284 | 9318 | 0 | 1.26% |
| scaffold00284 | 9321 | 1.15% | 0 |
| scaffold00284 | 9352 | 1.08% | 0 |
| scaffold00284 | 9365 | 1.37% | 1.85% |
| scaffold00284 | 9394 | 1.59% | 0 |
| scaffold00284 | 9411 | 1.18% | 0 |
| scaffold00284 | 9422 | 1.96% | 0 |
| scaffold00284 | 9426 | 1.38% | 0 |
| scaffold00284 | 9427 | 0 | 2.09% |
| scaffold00284 | 9429 | 1.47% | 0 |
| scaffold00284 | 9433 | 0 | 2.16% |
| scaffold00284 | 9439 | 0 | 2.89% |
| scaffold00284 | 9440 | 1.68% | 3.01% |
| scaffold00284 | 9448 | 2.35% | 0 |
| scaffold00284 | 9458 | 0 | 41.38% |
| scaffold00284 | 9459 | 0 | 44.83% |
| scaffold00284 | 9465 | 0 | 3.72% |
| scaffold00284 | 9475 | 0.91% | 0 |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 9479 | 0 | 37.93% |
| scaffold00284 | 9485 | 1.35% | 0 |
| scaffold00284 | 9490 | 2.75% | 0 |
| scaffold00284 | 9501 | 0 | 6.07% |
| scaffold00284 | 9536 | 1.18% | 0 |
| scaffold00284 | 9537 | 0 | 8.92% |
| scaffold00284 | 9540 | 0 | 2.86% |
| scaffold00284 | 9541 | 0 | 9.97% |
| scaffold00284 | 9544 | 0 | 3.63% |
| scaffold00284 | 9545 | 0 | 9.05% |
| scaffold00284 | 9546 | 0 | 5.88% |
| scaffold00284 | 9558 | 1.05% | 0 |
| scaffold00284 | 9560 | 1.33% | 0 |
| scaffold00284 | 9561 | 0 | 10.39% |
| scaffold00284 | 9562 | 0 | 9.52% |
| scaffold00284 | 9565 | 0 | 4.27% |
| scaffold00284 | 9567 | 0 | 3.08% |
| scaffold00284 | 9568 | 1.67% | 5.83% |
| scaffold00284 | 9572 | 0 | 9.32% |
| scaffold00284 | 9573 | 0 | 9.02% |
| scaffold00284 | 9576 | 0 | 14.78% |
| scaffold00284 | 9593 | 1.85% | 0 |
| scaffold00284 | 9598 | 1.32% | 0 |
| scaffold00284 | 9639 | 0 | 6.19% |
| scaffold00284 | 9643 | 1.56% | 0 |
| scaffold00284 | 9648 | 0 | 5.48% |
| scaffold00284 | 9649 | 0 | 7.47% |
| scaffold00284 | 9668 | 2.02% | 0 |
| scaffold00284 | 9694 | 0 | 5.86% |
| scaffold00284 | 9718 | 2.15% | 0 |
| scaffold00284 | 9749 | 0 | 15.45% |
| scaffold00284 | 9757 | 3.17% | 5.24% |
| scaffold00284 | 9770 | 2.33% | 0 |
| scaffold00284 | 9775 | 2.14% | 7.47% |
| scaffold00284 | 9778 | 0 | 2.71% |
| scaffold00284 | 9779 | 0 | 4.16% |
| scaffold00284 | 9780 | 2.61% | 7.36% |
| scaffold00284 | 9782 | 2.11% | 0 |
| scaffold00284 | 9786 | 2.50% | 0 |
| scaffold00284 | 9803 | 0 | 2.96% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 9807 | 0 | 2.80% |
| scaffold00284 | 9811 | 4.11% | 0 |
| scaffold00284 | 9814 | 0 | 5.32% |
| scaffold00284 | 9815 | 2.53% | 0 |
| scaffold00284 | 9829 | 4.63% | 0 |
| scaffold00284 | 9921 | 0 | 52.17% |
| scaffold00284 | 9938 | 0 | 46.67% |
| scaffold00284 | 9947 | 0 | 42.86% |
| scaffold00284 | 9952 | 2.89% | 0 |
| scaffold00284 | 9957 | 8.72% | 38.89% |
| scaffold00284 | 9968 | 4.78% | 17.86% |
| scaffold00284 | 9973 | 1.91% | 0 |
| scaffold00284 | 9980 | 3.72% | 17.50% |
| scaffold00284 | 9985 | 0 | 25.58% |
| scaffold00284 | 9996 | 1.70% | 0 |
| scaffold00284 | 10000 | 1.87% | 0 |
| scaffold00284 | 10044 | 2.25% | 0 |
| scaffold00284 | 10211 | 8.47% | 0 |
| scaffold00284 | 10213 | 5.05% | 0 |
| scaffold00284 | 10258 | 1.93% | 0 |
| scaffold00284 | 10262 | 1.91% | 0 |
| scaffold00284 | 10269 | 2.32% | 0 |
| scaffold00284 | 10272 | 4.87% | 6.37% |
| scaffold00284 | 10274 | 5.57% | 0 |
| scaffold00284 | 10276 | 8.64% | 14.72% |
| scaffold00284 | 10278 | 6.65% | 0 |
| scaffold00284 | 10285 | 0 | 9.30% |
| scaffold00284 | 10286 | 1.96% | 0 |
| scaffold00284 | 10288 | 5.15% | 0 |
| scaffold00284 | 10293 | 1.86% | 0 |
| scaffold00284 | 10310 | 0 | 7.92% |
| scaffold00284 | 10311 | 0 | 4.16% |
| scaffold00284 | 10314 | 0 | 6.21% |
| scaffold00284 | 10317 | 0 | 10.58% |
| scaffold00284 | 10318 | 0 | 3.84% |
| scaffold00284 | 10319 | 0 | 6.29% |
| scaffold00284 | 10338 | 0 | 6.75% |
| scaffold00284 | 10340 | 0 | 4.34% |
| scaffold00284 | 10344 | 0 | 2.79% |
| scaffold00284 | 10347 | 0 | 5.82% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 10348 | 0 | 6.88% |
| scaffold00284 | 10351 | 0 | 4.50% |
| scaffold00284 | 10394 | 0 | 1.77% |
| scaffold00284 | 10400 | 0 | 2.40% |
| scaffold00284 | 10401 | 0 | 1.94% |
| scaffold00284 | 10407 | 0 | 4.62% |
| scaffold00284 | 10408 | 0 | 2.85% |
| scaffold00284 | 10418 | 0 | 7.99% |
| scaffold00284 | 10421 | 0 | 7.69% |
| scaffold00284 | 10439 | 0 | 3.31% |
| scaffold00284 | 10445 | 0 | 4.95% |
| scaffold00284 | 10449 | 0 | 3.91% |
| scaffold00284 | 10460 | 0 | 9.55% |
| scaffold00284 | 10461 | 0 | 2.29% |
| scaffold00284 | 10466 | 0 | 16.67% |
| scaffold00284 | 10468 | 0 | 2.73% |
| scaffold00284 | 10470 | 0 | 10.20% |
| scaffold00284 | 10473 | 1.78% | 0 |
| scaffold00284 | 10481 | 1.86% | 3.85% |
| scaffold00284 | 10482 | 1.81% | 9.53% |
| scaffold00284 | 10483 | 0 | 2.23% |
| scaffold00284 | 10487 | 0 | 4.12% |
| scaffold00284 | 10488 | 0 | 2.36% |
| scaffold00284 | 10494 | 0 | 6.49% |
| scaffold00284 | 10507 | 0.93% | 0.97% |
| scaffold00284 | 10508 | 1.30% | 1.76% |
| scaffold00284 | 10510 | 0 | 1.15% |
| scaffold00284 | 10527 | 0 | 0.85% |
| scaffold00284 | 10533 | 0 | 5.04% |
| scaffold00284 | 10534 | 0 | 9.96% |
| scaffold00284 | 10538 | 1.33% | 0 |
| scaffold00284 | 10541 | 0 | 1.31% |
| scaffold00284 | 10544 | 0 | 15.54% |
| scaffold00284 | 10551 | 0 | 15.24% |
| scaffold00284 | 10552 | 0 | 8.88% |
| scaffold00284 | 10553 | 0 | 21.38% |
| scaffold00284 | 10558 | 1.11% | 0 |
| scaffold00284 | 10561 | 0 | 0.51% |
| scaffold00284 | 10562 | 0.86% | 0.75% |
| scaffold00284 | 10565 | 1.67% | 1.07% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 10589 | 0 | 0.77% |
| scaffold00284 | 10591 | 0.70% | 0 |
| scaffold00284 | 10595 | 0 | 1.00% |
| scaffold00284 | 10601 | 0 | 0.64% |
| scaffold00284 | 10602 | 1.24% | 0.56% |
| scaffold00284 | 10607 | 0.78% | 0 |
| scaffold00284 | 10608 | 1.05% | 1.57% |
| scaffold00284 | 10608 | 1.05% | 0 |
| scaffold00284 | 10611 | 1.06% | 0.86% |
| scaffold00284 | 10616 | 2.43% | 0 |
| scaffold00284 | 10617 | 2.38% | 0 |
| scaffold00284 | 10620 | 0 | 0.80% |
| scaffold00284 | 10622 | 1.65% | 2.42% |
| scaffold00284 | 10628 | 0.94% | 0 |
| scaffold00284 | 10629 | 0.96% | 3.57% |
| scaffold00284 | 10635 | 1.83% | 0 |
| scaffold00284 | 10644 | 2.17% | 0 |
| scaffold00284 | 10646 | 0 | 2.42% |
| scaffold00284 | 10647 | 0 | 9.09% |
| scaffold00284 | 10651 | 0 | 3.75% |
| scaffold00284 | 10652 | 1.48% | 0 |
| scaffold00284 | 10655 | 1.53% | 7.06% |
| scaffold00284 | 10658 | 0 | 1.97% |
| scaffold00284 | 10664 | 0 | 1.71% |
| scaffold00284 | 10686 | 1.21% | 2.34% |
| scaffold00284 | 10702 | 0 | 2.00% |
| scaffold00284 | 10703 | 1.22% | 0 |
| scaffold00284 | 10705 | 0 | 1.98% |
| scaffold00284 | 10706 | 0 | 0.69% |
| scaffold00284 | 10711 | 0 | 1.62% |
| scaffold00284 | 10714 | 0 | 0.81% |
| scaffold00284 | 10716 | 0 | 3.19% |
| scaffold00284 | 10721 | 1.52% | 2.78% |
| scaffold00284 | 10725 | 0 | 0.78% |
| scaffold00284 | 10729 | 0.93% | 0.93% |
| scaffold00284 | 10734 | 1.95% | 0 |
| scaffold00284 | 10739 | 0 | 2.54% |
| scaffold00284 | 10740 | 0 | 0.66% |
| scaffold00284 | 10744 | 0.93% | 0 |
| scaffold00284 | 10748 | 0 | 3.88% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 10755 | 0 | 0.72% |
| scaffold00284 | 10758 | 0.90% | 1.59% |
| scaffold00284 | 10761 | 0 | 1.34% |
| scaffold00284 | 10764 | 1.00% | 0 |
| scaffold00284 | 10765 | 1.45% | 2.56% |
| scaffold00284 | 10766 | 0 | 1.08% |
| scaffold00284 | 10767 | 0 | 1.15% |
| scaffold00284 | 10769 | 1.20% | 0 |
| scaffold00284 | 10771 | 0 | 0.89% |
| scaffold00284 | 10775 | 0 | 2.75% |
| scaffold00284 | 10778 | 0 | 2.96% |
| scaffold00284 | 10779 | 0 | 2.29% |
| scaffold00284 | 10784 | 0 | 1.25% |
| scaffold00284 | 10787 | 0 | 1.39% |
| scaffold00284 | 10790 | 0 | 3.12% |
| scaffold00284 | 10791 | 0 | 1.11% |
| scaffold00284 | 10796 | 0 | 3.30% |
| scaffold00284 | 10801 | 0 | 4.02% |
| scaffold00284 | 10810 | 0 | 3.28% |
| scaffold00284 | 10812 | 0 | 1.95% |
| scaffold00284 | 10817 | 0 | 0.92% |
| scaffold00284 | 10831 | 1.40% | 0 |
| scaffold00284 | 10832 | 0 | 0.75% |
| scaffold00284 | 10833 | 0 | 2.98% |
| scaffold00284 | 10834 | 0 | 3.54% |
| scaffold00284 | 10839 | 0 | 1.89% |
| scaffold00284 | 10842 | 0 | 0.94% |
| scaffold00284 | 10843 | 0 | 1.36% |
| scaffold00284 | 10854 | 0 | 1.17% |
| scaffold00284 | 10883 | 0 | 2.43% |
| scaffold00284 | 10886 | 0 | 27.27% |
| scaffold00284 | 10895 | 0 | 3.88% |
| scaffold00284 | 10897 | 0 | 3.58% |
| scaffold00284 | 10905 | 0 | 22.64% |
| scaffold00284 | 10916 | 0 | 34.04% |
| scaffold00284 | 10940 | 0 | 24.49% |
| scaffold00284 | 10942 | 0 | 58.49% |
| scaffold00284 | 10945 | 0 | 73.68% |
| scaffold00284 | 10946 | 0 | 57.89% |
| scaffold00284 | 10953 | 0 | 18.89% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 10956 | 0 | 28.57% |
| scaffold00284 | 10963 | 0 | 12.56% |
| scaffold00284 | 10964 | 0 | 10.57% |
| scaffold00284 | 10965 | 0 | 4.12% |
| scaffold00284 | 10966 | 0 | 7.97% |
| scaffold00284 | 10976 | 0 | 6.99% |
| scaffold00284 | 10983 | 0 | 2.11% |
| scaffold00284 | 10985 | 0 | 4.92% |
| scaffold00284 | 10986 | 0 | 3.25% |
| scaffold00284 | 10987 | 0 | 6.99% |
| scaffold00284 | 11005 | 0 | 2.78% |
| scaffold00284 | 11008 | 0 | 2.38% |
| scaffold00284 | 11013 | 0 | 2.30% |
| scaffold00284 | 11014 | 0 | 1.42% |
| scaffold00284 | 11019 | 0 | 1.68% |
| scaffold00284 | 11033 | 0 | 2.19% |
| scaffold00284 | 11044 | 0 | 0.82% |
| scaffold00284 | 11045 | 0 | 2.69% |
| scaffold00284 | 11055 | 0 | 2.27% |
| scaffold00284 | 11057 | 0 | 2.07% |
| scaffold00284 | 11065 | 1.60% | 1.04% |
| scaffold00284 | 11081 | 0 | 4.24% |
| scaffold00284 | 11086 | 0 | 0.74% |
| scaffold00284 | 11089 | 1.36% | 1.52% |
| scaffold00284 | 11091 | 0 | 0.91% |
| scaffold00284 | 11097 | 0 | 1.13% |
| scaffold00284 | 11115 | 0 | 1.20% |
| scaffold00284 | 11134 | 0 | 1.08% |
| scaffold00284 | 11140 | 0 | 1.18% |
| scaffold00284 | 11153 | 0 | 1.30% |
| scaffold00284 | 11154 | 0 | 2.25% |
| scaffold00284 | 11185 | 0 | 5.56% |
| scaffold00284 | 11187 | 0 | 5.90% |
| scaffold00284 | 11198 | 0 | 3.19% |
| scaffold00284 | 11200 | 0 | 3.08% |
| scaffold00284 | 11201 | 0 | 2.78% |
| scaffold00284 | 11203 | 0 | 12.66% |
| scaffold00284 | 11224 | 0 | 0.77% |
| scaffold00284 | 11243 | 1.14% | 0 |
| scaffold00284 | 11258 | 0 | 0.73% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00284 | 11259 | 0 | 3.03% |
| scaffold00284 | 11287 | 0 | 0.62% |
| scaffold00284 | 11288 | 0 | 0.94% |
| scaffold00290 | 15 | 0 | 1.51% |
| scaffold00290 | 40 | 0 | 0.97% |
| scaffold00290 | 53 | 0 | 0.83% |
| scaffold00290 | 59 | 0 | 1.08% |
| scaffold00290 | 68 | 0 | 1.62% |
| scaffold00290 | 88 | 0 | 1.39% |
| scaffold00290 | 89 | 1.45% | 0 |
| scaffold00290 | 89 | 0 | 1.76% |
| scaffold00290 | 93 | 0 | 1.29% |
| scaffold00290 | 102 | 0 | 1.33% |
| scaffold00290 | 106 | 0 | 2.50% |
| scaffold00290 | 127 | 0 | 1.84% |
| scaffold00290 | 152 | 0 | 1.62% |
| | | | |
| scaffold00323 | 57 | 0 | 0.97% |
| scaffold00323 | 66 | 0 | 1.12% |
| scaffold00323 | 70 | 0 | 0.99% |
| scaffold00323 | 76 | 4.65% | 0 |
| scaffold00323 | 76 | 0 | 1.29% |
| scaffold00323 | 91 | 0 | 3.82% |
| scaffold00323 | 133 | 0 | 2.98% |
| scaffold00323 | 142 | 5.07% | 0 |
| scaffold00323 | 146 | 0 | 3.97% |
| scaffold00323 | 147 | 4.68% | 0 |
| scaffold00323 | 152 | 3.61% | 0 |
| scaffold00323 | 152 | 0 | 3.60% |
| scaffold00323 | 167 | 0 | 1.76% |
| scaffold00323 | 171 | 0 | 0.87% |
| scaffold00323 | 180 | 1.60% | 0 |
| scaffold00323 | 193 | 0 | 1.79% |
| | | | |
| scaffold00337 | 50 | 0 | 21.84% |
| scaffold00337 | 51 | 0 | 17.00% |
| scaffold00337 | 57 | 0 | 20.63% |
| scaffold00337 | 60 | 0 | 9.58% |
| scaffold00337 | 65 | 0 | 9.59% |
| scaffold00337 | 72 | 5.91% | 0 |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00337 | 76 | 4.53% | 0 |
| scaffold00337 | 81 | 0 | 5.13% |
| scaffold00337 | 81 | 1.69% | 0 |
| scaffold00337 | 87 | 0 | 4.21% |
| scaffold00337 | 87 | 1.96% | 0 |
| scaffold00337 | 89 | 2.48% | 0 |
| scaffold00337 | 96 | 0 | 6.96% |
| scaffold00337 | 96 | 4.29% | 0 |
| scaffold00337 | 98 | 1.76% | 0 |
| scaffold00337 | 99 | 0 | 6.25% |
| scaffold00337 | 99 | 1.74% | 0 |
| scaffold00337 | 101 | 0 | 4.72% |
| scaffold00337 | 101 | 1.70% | 0 |
| scaffold00337 | 103 | 1.84% | 0 |
| scaffold00337 | 111 | 0 | 3.48% |
| scaffold00337 | 111 | 2.39% | 0 |
| scaffold00337 | 114 | 0 | 1.84% |
| scaffold00337 | 126 | 0 | 3.93% |
| scaffold00337 | 126 | 1.37% | 0 |
| scaffold00337 | 130 | 0 | 4.36% |
| scaffold00337 | 138 | 1.46% | 0 |
| scaffold00337 | 144 | 1.75% | 0 |
| scaffold00337 | 148 | 1.51% | 0 |
| scaffold00337 | 151 | 0 | 7.98% |
| scaffold00337 | 151 | 2.51% | 0 |
| scaffold00337 | 154 | 1.85% | 0 |
| scaffold00337 | 155 | 0 | 9.38% |
| scaffold00337 | 155 | 3.19% | 0 |
| scaffold00337 | 156 | 0 | 11.92% |
| scaffold00337 | 156 | 3.00% | 0 |
| scaffold00337 | 159 | 0 | 7.80% |
| scaffold00337 | 159 | 2.05% | 0 |
| scaffold00337 | 169 | 0 | 12.61% |
| scaffold00337 | 169 | 2.27% | 0 |
| scaffold00337 | 172 | 0 | 14.15% |
| scaffold00337 | 199 | 0 | 20.55% |
| scaffold00337 | 562 | 0 | 12.61% |
| scaffold00337 | 581 | 0 | 17.65% |
| scaffold00337 | 1087 | 8.00% | 0 |
| scaffold00337 | 4624 | 2.20% | 0 |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00337 | 4629 | 1.84% | 0 |
| scaffold00337 | 4649 | 6.15% | 0 |
| scaffold00337 | 4650 | 0 | 3.59% |
| scaffold00337 | 4650 | 1.67% | 0 |
| scaffold00337 | 4656 | 0 | 3.65% |
| scaffold00337 | 4656 | 1.90% | 0 |
| scaffold00337 | 4659 | 3.88% | 0 |
| scaffold00337 | 4661 | 1.09% | 0 |
| scaffold00337 | 4669 | 1.01% | 0 |
| scaffold00337 | 4670 | 0.98% | 0 |
| scaffold00337 | 4681 | 0 | 3.04% |
| scaffold00337 | 4681 | 1.53% | 0 |
| scaffold00337 | 4686 | 3.76% | 0 |
| scaffold00337 | 4687 | 1.53% | 0 |
| scaffold00337 | 4689 | 3.38% | 0 |
| scaffold00337 | 4691 | 1.58% | 0 |
| scaffold00337 | 4695 | 2.20% | 0 |
| scaffold00337 | 4831 | 2.32% | 0 |
| scaffold00337 | 4852 | 1.35% | 0 |
| scaffold00337 | 4854 | 11.36% | 0 |
| scaffold00337 | 4855 | 2.03% | 0 |
| scaffold00337 | 4856 | 0 | 5.03% |
| scaffold00337 | 4856 | 1.38% | 0 |
| scaffold00337 | 4857 | 0.97% | 0 |
| scaffold00337 | 4864 | 2.03% | 0 |
| scaffold00337 | 4866 | 1.87% | 0 |
| scaffold00337 | 4868 | 2.68% | 0 |
| scaffold00337 | 4869 | 2.10% | 0 |
| scaffold00337 | 4872 | 1.50% | 0 |
| scaffold00337 | 4873 | 2.38% | 0 |
| scaffold00337 | 4876 | 2.00% | 0 |
| scaffold00337 | 4886 | 3.22% | 0 |
| scaffold00337 | 5021 | 0 | 8.40% |
| scaffold00337 | 5025 | 0 | 10.92% |
| scaffold00337 | 5031 | 0 | 7.87% |
| scaffold00337 | 5064 | 0 | 1.83% |
| scaffold00337 | 5070 | 0 | 3.30% |
| scaffold00337 | 5071 | 0 | 2.66% |
| scaffold00337 | 5072 | 0 | 3.64% |
| scaffold00337 | 5080 | 0 | 6.92% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00337 | 5081 | 0 | 8.33% |
| scaffold00337 | 5091 | 0 | 3.76% |
| scaffold00337 | 5091 | 1.86% | 0 |
| scaffold00337 | 5092 | 0 | 1.14% |
| scaffold00337 | 5092 | 1.71% | 0 |
| scaffold00337 | 5095 | 0 | 1.64% |
| scaffold00337 | 5108 | 0 | 3.28% |
| scaffold00337 | 5116 | 0 | 6.57% |
| scaffold00337 | 5118 | 0 | 5.07% |
| scaffold00337 | 5121 | 0 | 1.56% |
| scaffold00337 | 5126 | 0 | 2.28% |
| scaffold00337 | 5127 | 0 | 1.62% |
| scaffold00337 | 5128 | 0 | 4.11% |
| scaffold00337 | 5129 | 0 | 5.13% |
| scaffold00337 | 5133 | 0 | 2.34% |
| scaffold00337 | 5134 | 0 | 0.79% |
| scaffold00337 | 5136 | 0 | 2.67% |
| scaffold00337 | 5139 | 0 | 0.78% |
| scaffold00337 | 5140 | 0 | 1.10% |
| scaffold00337 | 5162 | 0 | 0.80% |
| scaffold00337 | 5165 | 0 | 1.78% |
| scaffold00337 | 5167 | 0 | 0.72% |
| scaffold00337 | 5167 | 1.41% | 0 |
| scaffold00337 | 5178 | 0 | 0.75% |
| scaffold00337 | 5179 | 0 | 3.41% |
| scaffold00337 | 5179 | 2.41% | 0 |
| scaffold00337 | 5180 | 0 | 1.74% |
| scaffold00337 | 5181 | 0 | 0.76% |
| scaffold00337 | 5188 | 0 | 2.91% |
| scaffold00337 | 5190 | 0 | 1.38% |
| scaffold00337 | 5191 | 0 | 7.13% |
| scaffold00337 | 5206 | 0 | 4.01% |
| scaffold00337 | 5207 | 0 | 2.31% |
| scaffold00337 | 5210 | 0 | 2.82% |
| scaffold00337 | 5216 | 0 | 1.19% |
| scaffold00337 | 5219 | 0 | 2.43% |
| scaffold00337 | 5227 | 0 | 3.19% |
| scaffold00337 | 5236 | 0 | 7.43% |
| scaffold00337 | 5239 | 0 | 4.89% |
| scaffold00337 | 5242 | 0 | 2.56% |

Table 13 (continued): Distribution of methylated sites in the fungal genome from squamules and podetia of *Cladonia grayi*.

| Scaffold | Position | % methylated Squamules | % methylated Podetia |
|---------------|----------|------------------------|----------------------|
| scaffold00337 | 5248 | 0 | 8.65% |
| scaffold00337 | 5269 | 0 | 18.33% |
| scaffold00337 | 5362 | 1.80% | 0 |
| scaffold00337 | 5365 | 2.14% | 0 |
| scaffold00337 | 5370 | 1.78% | 0 |
| scaffold00337 | 5371 | 4.42% | 0 |
| scaffold00337 | 5374 | 2.87% | 0 |
| scaffold00337 | 5394 | 8.65% | 0 |

References

- Adams, G. M. & Blumenthal, R. M. (1997) The PvuII DNA (cytosine-N4)-methyltransferase comprises two trypsin-defined domains, each of which binds a molecule of S-adenosyl-L-methionine. *Biochemistry* **36**: 8284-8292.
- Agatep R, Kirkpatrick RD, Parchaliuk DL, Wood RA, Gietz, RD. 1998. Transformation of *Saccharomyces cerevisiae* by the lithium acetate/single-stranded carrier DNA/polyethylene glycol (LiAc/ss-DNA/PEG) protocol. Technical Tips Online (<http://tto.trends.com>).
- Ahmadjian, V. 1966. Artificial Reestablishment of the Lichen *Cladonia cristatella*. *Science* **151**:199-201.
- Ahmadjian V. 1993. The lichen symbiosis. New York: John Wiley.
- Ahmadjian, V. & Heikkilä, H. (1970) The culture and synthesis of *Endocarpon pusillum* and *Staurothele clopima*. *Lichenologist* **4**: 259-267.
- Andrade SL, Dickmanns A, Ficner R, Einsle O. 2005. Crystal structure of the archaeal ammonium transporter Amt-1 from *Archaeoglobus fulgidus*. *Proc Natl Acad Sci USA* **102**:14994-14999.
- Antequera F, Tamame M., Villanueva JR, Santos T. 1984. DNA methylation in the fungi. *J Biol Chem* **259**:8033-8036.
- Armaleo D, May S. 2009. Sizing the fungal and algal genomes of the lichen *Cladonia grayi* through quantitative PCR. *Symbiosis* **49**:43-51.

- Armaleo D, Miao V. 1999. Symbiosis and DNA methylation in the *Cladonia* lichen fungus. 26:143-163.
- Armaelo D, Sun X, Culberson C. 2011. Insights from the first putative biosynthetic gene cluster for a lichen depside and depsidone. *Mycologia* 103: doi:10.3852/10-335.
- Arnold AE, Miadlikowska J, Higgins KL, Sarvate SD, Gugger P, Way A, Hofstetter V, Kauff F, Lutzoni F. 2009. A phylogenetic estimation of trophic transition networks for ascomycetous fungi: are lichens cradles of symbiotrophic fungal diversification? *Syst Biol* 58:283-297.
- Barry C, Faugeron G, Rossignol JL. 1993. Methylation induced premeiotically in *Ascobolus*: coextension with DNA repeat lengths and effect on transcript elongation. *Proc Natl Acad Sci U S A* 90:4557-4561.
- Bhattacharya D, Yoon HS, Hedges SB, Hackett JD. 2009. Eukaryotes (Eukaryota). In: Hedges SB, Kumar S, editors. *The Timetree of Life*. New York, NY: Oxford University Press. p. 116-120.
- Blair JE. 2009. Fungi. In: Hedges SB, Kumar S, editors. *The Timetree of Life*. New York, NY: Oxford University Press. p. 215-219.
- Brock TD, Brock KM, Belly RT, Weiss RL. 1972. *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch Mikrobiol.* 84:54-68.
- Brodo IM. 1973. The Lichen Genus *Coccotrema* in North America. *The Bryologist* 76:260-270.

- Butkus V, Petrauskiene L, Maneliene Z, Klimasauskas S, Laucys V, Janulaitis A. 1987. Cleavage of methylated CCCGGG sequences containing either N4-methylcytosine or 5-methylcytosine with MspI, HpaII, SmaI, XmaI and Cfr9I restriction endonucleases. *Nucleic Acids Res* 15:7091-7102.
- Clark DA, Norris PR. 1996. *Acidimicrobium ferrooxidans* gen. nov., sp. nov.: Mixed-culture ferrous iron oxidation with *Sulfobacillus* species. *Microbiology-UK* 142:785-790.
- Cokus SJ, Feng S, Zhang X, Chen Z., Merriman B, Haudenschild CD, Pradhan S, Nelson, SF, Pellegrini M, Jacobsen SE. 2008. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* 452:215-219.
- Couturier J, Montanini B, Martin F, Brun A, Blaudez D, Chalot M. 2007. The expanded family of ammonium transporters in the perennial poplar plant. *New Phytol.* 174:137-150.
- Dopson M, Baker-Austin C, Hind A, Bowman JP, Bond PL. 2004. Characterization of *Ferroplasma* isolates and *Ferroplasma acidarmanus* sp. nov., extreme acidophiles from acid mine drainage and industrial bioleaching environments. *Appl Environ Microbiol.* 70:2079-2088.
- Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH, Halpern ME, Ukomadu C, Sadler KC, Pradhan S, Pellegrini M, Jacobsen SE. 2010. Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci U S A* 107:8689-8694.
- Futterer O, Angelov A, Liesegang H, Gottschalk G, Schleper C, Schepers B, Dock C, Antranikian G, Liebl W. 2004. Genome sequence of *Picrophilus torridus* and its implications for life around pH 0. *Proc Natl Acad Sci USA* 101:9091-9096.

- Gaya E, Holguin A, Fernandez-Brime S, Högnabba F, Stenroos, S, Molnar K, Lutzoni, F. 2011. Resolving deep internodes in the Teloschistales with a multi-gene approach. *Molecular Phylogenetics and Evolution* **In review**.
- Gimeno CJ, Ljungdahl P O, Styles C A, Fink GR. 1992. Unipolar cell divisions in the yeast *S. cerevisiae* lead to filamentous growth: regulation by starvation and RAS. *Cell* 68:1077-1090.
- Gogarten, J. P., W. F. Doolittle, and J. G. Lawrence. 2002. Prokaryotic evolution in light of gene transfer. *Mol Biol Evol.* 19:2226-2238.
- Goltsman DS, Denev VJ, Singer SW, et al. (17 co-authors). 2009. Community genomic and proteomic analyses of chemoautotrophic iron-oxidizing "*Leptospirillum rubarum*" (Group II) and "*Leptospirillum ferrodiazotrophum*" (Group III) bacteria in acid mine drainage biofilms. *Appl Environ Microbiol.* 75:4599-4615.
- Gruswitz F, Chaudhary S, Ho JD, Schlessinger A, Pezeshki B, Ho CM, Sali A, Westhoff CM, Stroud RM. 2010. Function of human Rh based on structure of RhCG at 2.1 A. *Proc. Natl Acad Sci U S A.* 107:9638-43.
- Gueidan C, Roux C, Lutzoni F. 2007. Using a multigene phylogenetic analysis to assess generic delineation and character evolution in *Verrucariaceae* (Verrucariales, Ascomycota). *Mycol Res.* 111:1145-1168.
- Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol.* 52: 696-704.
- Hibbett DS, Binder M, Bischoff JF, et al. (67 co-authors). 2007. A higher-level phylogenetic classification of the Fungi. *Mycol Res.* 111:509-547.

- Hickman, MA, Froyd, CA, Rusche, LN. 2011. Reinventing heterochromatin in budding yeasts: Sir2 and the origin recognition complex take center stage. *Eukaryot. Cell* 10(9):1183-9.
- Hoffman CS, Winston F. 1987. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene* 57:267-272.
- Huang CH, Peng J. 2005. Evolutionary conservation and diversification of Rh family genes and proteins. *Proc Natl Acad Sci USA* 102:15512-15517.
- Itoh T, Suzuki K, Sanchez PC, Nakase T. 1999. *Caldivirga maquilingensis* gen. nov., sp. nov., a new genus of rod-shaped crenarchaeote isolated from a hot spring in the Philippines. *Int J Syst Bacteriol.* 49 Pt 3:1157-1163.
- Iyer LM, Abhiman S, Aravind L. 2011. Natural history of eukaryotic DNA methylation systems. *Prog Mol Biol Transl Sci* 101:25-104.
- James TY, Kauff F, Schoch CL, et al. (70 co-authors). 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* 443:818-822.
- Javelle A, Andre B, Marini AM, Chalot M. 2003. High-affinity ammonium transporters and nitrogen sensing in mycorrhizas. *Trends Microbiol.* 11:53-55.
- Javelle A, Lupo D, Zheng L, Li XD, Winkler FK, Merrick M. 2006. An unusual twin-his arrangement in the pore of ammonia channels is essential for substrate conductance. *J Biol Chem.* 281:39492-39498.
- Javelle A, Morel M, Rodríguez-Pastrana BR, Botton B, André B, Marini AM, Brun A, Chalot M. 2003. Molecular characterization, function and regulation of ammonium transporters (Amt) and ammonia-metabolizing enzymes (GS,

NADP-GDH) in the ectomycorrhizal fungus *Hebeloma cylindrosporum*. *Mol. Microbiol.* 47(2):411-30

Keeling P, Leander BS, Simpson A. 2009. Eukaryotes. Eukaryota, Organisms with nucleated cells. Version 28 October 2009.
<http://tolweb.org/Eukaryotes/3/2009.10.28> in The Tree of Life Web Project,
<http://tolweb.org>.

Kelly DP, Wood AP. 2000. Reclassification of some species of *Thiobacillus* to the newly designated genera *Acidithiobacillus* gen. nov., *Halothiobacillus* gen. nov. and *Thermithiobacillus* gen. nov. *Int J Syst Evol Microbiol.* 50 Pt 2:511-516.

Khademi S, O'Connell J 3rd, Remis J, Y. Robles-Colmenares Y, Miercke LJ, Stroud RM. 2004. Mechanism of ammonia transport by Amt/MEP/Rh: structure of AmtB at 1.35 Å. *Science* 305:1587-1594.

Kustu S, Inwood W. 2006. Biological gas channels for NH₃ and CO₂: evidence that Rh (Rhesus) proteins are CO₂ channels. *Transfus Clin Biol.* 13:103-110.

Lamoureux G, Javelle A, Baday S, Wang S, Berneche S. 2010. Transport mechanisms in the ammonium transporter family. *Transfus Clin Biol.* 17:168-175.

Lamoureux G, Klein ML, Berneche S. 2007. A stable water chain in the hydrophobic pore of the AmtB ammonium transporter. *Biophys J.* 92:L82-84.

Lee DW, Freitag M, Selker EU, Aramayo R. 2008. A cytosine methyltransferase homologue is essential for sexual development in *Aspergillus nidulans*. *PLoS One* 3: e2531.

- Li LC, Dahiya R. 2002. MethPrimer: designing primers for methylation PCRs. *Bioinformatics* 18:1427-1431.
- Li X, Lupo D, Zheng L, Winkler F. 2006. Structural and functional insights into the AmtB/Mep/Rh protein family. *Transfus Clin Biol.* 13:165-69.
- Li X, Jayachandran S, Nguyen HH, Chan MK. 2007. Structure of the *Nitrosomonas europaea* Rh protein. *Proc Natl Acad Sci USA* 104:19279-19284.
- Lopez-Pedrosa A, Gonzalez-Guerrero M, Valderas A, Azcon-Aguilar C, Ferrol N. 2006. *GintAMT1* encodes a functional high-affinity ammonium transporter that is expressed in the extraradical mycelium of *Glomus intraradices*. *Fungal Genet Biol.* 43:102-110.
- Lorenz, M. C., and J. Heitman. 1998. The MEP2 ammonium permease regulates pseudohyphal differentiation in *Saccharomyces cerevisiae*. *Embo J.* 17:1236-1247.
- Ludewig, U., N. von Wiren, and W. B. Frommer. 2002. Uniport of NH₄⁺ by the root hair plasma membrane ammonium transporter LeAMT1;1. *J Biol Chem.* 277:13548-13555.
- Lutzoni F, Kauff F, Cox JC, et al. (44 co-authors). 2004. Assembling the fungal tree of life: Progress, classification, and evolution of subcellular traits. *Am J Bot.* 91:1446-1480.
- Lutzoni, F., M. Pagel, and V. Reeb. 2001. Major fungal lineages are derived from lichen symbiotic ancestors. *Nature* 411:937-940.
- Maddison DR, Maddison WP. 2005. MacClade 4. Sunderland, MA: Sinauer Associates.

- Marcet-Houben M, Gabaldon T. 2010. Acquisition of prokaryotic genes by fungal genomes. *Trends Genet.* 26:5-8.
- Marini AM, Vissers S, Urrestarazu S, Andre B. 1994. Cloning and expression of the *MEP1* gene encoding an ammonium transporter in *Saccharomyces cerevisiae*. *EMBO J.* 13 (15): 3456-63.
- Marini AM, Soussi-Boudekou S, Vissers S, Andre B. 1997a. A family of ammonium transporters in *Saccharomyces cerevisiae*. *Mol Cell Biol.* 17:4282-4293.
- Marini AM, Urrestarazu A, Beauwens R, Andre B. 1997b. The Rh (rhesus) blood group polypeptides are related to NH₄⁺ transporters. *Trends Biochem Sci.* 22:460-461.
- Mayer, M., M. Dynowski, and U. Ludewig. 2006. Ammonium ion transport by the AMT/Rh homologue LeAMT1;1. *Biochem J.* 396:431-437.
- Mayer M, Ludewig U. 2006. Role of AMT1;1 in NH₄⁺ acquisition in *Arabidopsis thaliana*. *Plant Biol. (Stuttg)* 8:522-528.
- McDonald SM, Plant JN, Worden AZ. The mixed lineage nature of nitrogen transport and assimilation in marine eukaryotic phytoplankton: a case study of *Micromonas*. *Mol Biol Evol.* 27:2268-2283.
- McDonald, T. R., Dietrich, F. S. & Lutzoni, F. (2011) Multiple Horizontal Gene Transfers of Ammonium Transporters/Ammonia Permeases from Prokaryotes to Eukaryotes: Toward a New Functional and Evolutionary Classification. *Mol Biol Evol.* Aug 27 (Epub ahead of print).
- Miadlikowska J, Lutzoni F. 2004. Phylogenetic classification of peltigeralean fungi (Peltigerales, Ascomycota) based on ribosomal RNA small and large subunits. *Am J Bot* 91:449-464.

- Miao VP, Freitag M, Selker EU. 2000. Short TpA-rich segments of the zeta-eta region induce DNA methylation in *Neurospora crassa*. *J Mol Biol* 300: 249-273.
- Mishra PK, Baum M, Carbon J. 2011. DNA methylation regulates phenotype-dependent transcriptional activity in *Candida albicans*. *Proc Natl Acad Sci U S A* 108:11965-11970.
- Monahan BJ, Fraser JA, Hynes MJ, Davis MA. 2002. Isolation and characterization of two ammonium permease genes, *meaA* and *mepA*, from *Aspergillus nidulans*. *Eukaryot Cell* 1:85-94.
- Montanini B, Moretto N, Soragni E, Percudani R, Ottonello S. 2002. A high-affinity ammonium transporter from the mycorrhizal ascomycete *Tuber borchii*. *Fungal Genet Biol.* 36:22-34.
- Mumberg D, Muller R, Funk M. 1995. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 156:119-122.
- Ninnemann O, Jauniaux JC, Frommer WB. 1994. Identification of a high affinity NH₄⁺ transporter from plants. *EMBO J.* 13 (15): 3464-71.
- Nylander JAA. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden.
- Oudet P, Gross-Bellard M, Chambon P. 1975. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* 4:281-300.
- Posada D. 2008. jModelTest: phylogenetic model averaging. *Mol Biol Evol.* 25:1253-1256.

- Raymond CK, Pownder T A, Sexson SL. 1999. General method for plasmid construction using homologous recombination. *Biotechniques* 26:134-138, 140-131.
- Richards TA, Soanes DM, Foster PG, Leonard G, Thornton CR, Talbot NJ. 2009. Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. *Plant Cell* 21:1897-1911.
- Rippka R, Herdman M. 1992. *Pasteur Culture Collection of Cyanobacterial Strains in Axenic Culture, Catalogue and Taxonomic Handbook*. Paris: Institut Pasteur.
- Rusche LN, Rine J. 2001. Conversion of a gene-specific repressor to a regional silencer. *Genes Dev* 15:955-967.
- Rutherford JC, Lin X, Nielsen K, Heitman J. 2008. Amt2 permease is required to induce ammonium-responsive invasive growth and mating in *Cryptococcus neoformans*. *Eukaryot Cell* 7:237-246.
- Savic S, Tibell L, Gueidan C, Lutzoni F. 2008. Molecular phylogeny and systematics of *Polyblastia* (Verrucariaceae, Eurotiomycetes) and allied genera. *Mycol Res* 112:1307-1318.
- Schleper C, Puehler G, Holz, I, Gambacorta A, Janekovic D, Santarius U, Klenk HP, Zillig W. 1995. *Picrophilus* gen. nov., fam. nov.: a novel aerobic, heterotrophic, thermoacidophilic genus and family comprising archaea capable of growth around pH 0. *J Bacteriol.* 177:7050-7059.
- Schleper C, Puhler G, Klenk HP, Zillig W. 1996. *Picrophilus oshimae* and *Picrophilus torridus* fam. nov., gen. nov., sp. nov., two species of hyperacidophilic, thermophilic, heterotrophic, aerobic archaea. *Int. J. Syst. Bacteriology* 46:814-816.

- Schoch CL, Sung GH, Lopez-Giraldez F, et al. (64 co-authors). 2009. The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. *Syst Biol*. 58:224-239.
- Selker EU, Tountas NA, Cross SH, Margolin BS, Murphy JG, Bird AP, Freitag M. 2003. The methylated component of the *Neurospora crassa* genome. *Nature* 422:893-897.
- She Q, Singh RK, Confalonieri F, et al. (31 co-authors). 2001. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc Natl Acad Sci USA* 98:7835-7840.
- Simmons EG. 1992. *Alternaria* taxonomy: Current status, Viewpoint, Challenge. Pp. 1-36 in J. C. a. A. Visconti, ed. *Alternaria: Biology, Plant Diseases and Metabolites*. Elsevier, New York.
- Simon-Rosin U, Wood, C, Udvardi MK. 2003. Molecular and cellular characterisation of LjAMT2;1, an ammonium transporter from the model legume *Lotus japonicus*. *Plant Mol Biol* 51:99-108.
- Slot JC, Hibbett DS. 2007. Horizontal transfer of a nitrate assimilation gene cluster and ecological transitions in fungi: a phylogenetic study. *PLoS One* 2 (10): e1097.
- Smith DG, Garcia-Pedrajas MD, Gold SE, Perlin MH. 2003. Isolation and characterization from pathogenic fungi of genes encoding ammonium permeases and their roles in dimorphism. *Mol Microbiol*. 50(1):259-75.
- Sohlenkamp C, Shelden M., Howitt S, Udvardi, M. 2000. Characterization of *Arabidopsis* AtAMT2, a novel ammonium transporter in plants. *FEBS Lett* 467:273-278.

- Sohlenkamp C, Wood CC, Roeb GW, Udvardi MK. 2002. Characterization of Arabidopsis AtAMT2, a high-affinity ammonium transporter of the plasma membrane. *Plant Physiol.* 130:1788-1796.
- Stamatakis A, Ludwig T, Meier H. 2005. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* 21:456-463.
- Strahl BD, Allis CD. 2000. The language of covalent histone modifications. *Nature* 403: 41-45.
- Suzuki T, Iwasaki T, Uzawa T, Hara K, Nemoto N, Kon T, Ueki T, Yamagishi A, Oshima T. 2002. *Sulfolobus tokodaii* sp. nov. (f. *Sulfolobus* sp. strain 7), a new member of the genus *Sulfolobus* isolated from Beppu Hot Springs, Japan. *Extremophiles* 6:39-44.
- Swofford DL. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sunderland, MA: Sinauer Associates.
- Trembley ML, Ringli C, and Honegger R. 2002. Morphological and molecular analysis of early stages in the resynthesis of the lichen *Baeomyces rufus*. *Mycol. Res* 106 (7):768-776.
- Valdes J, Quatrini R, Hallberg K, Dopson M, Valenzuela PD, and Holmes DS. 2009. Draft genome sequence of the extremely acidophilic bacterium *Acidithiobacillus caldus* ATCC 51756 reveals metabolic versatility in the genus *Acidithiobacillus*. *J Bacteriol.* 191:5877-5878.
- Vos P, Hogers R, Bleeker M., Reijans M., van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, et al. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 23: 4407-4414.

- Weber M, Davies J J, Wittig D, Oakeley E J, Haase M, Lam W L, Schubeler D. 2005. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet* 37:853-862.
- Winkler FK. 2006. Amt/MEP/Rh proteins conduct ammonia. *Pflugers Arch.* 451:701-707.
- Yang H, Xu Y, Zhu W, Chen K, Jiang H. 2007. Detailed mechanism for AmtB conducting NH₄⁺/NH₃: molecular dynamics simulations. *Biophys J.* 92:877-885.
- Zemach A, McDaniel IE, Silva P, Zilberman D. 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* 328:916-919.
- Zerbino, DR, Birney, E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821-829.
- Zheng L, Kostrewa D, Berneche S, Winkler FK, Li XD. 2004. The mechanism of ammonia transport based on the crystal structure of AmtB of *Escherichia coli*. *Proc Natl Acad Sci USA* 101:17090-17095.
- Zolan ME, Pukkila PJ. 1986. Inheritance of DNA methylation in *Coprinus cinereus*. *Mol Cell Biol.* 6:195-200.

Biography

Tami Renae McDonald was born January 14, 1974 in St. Paul, Minnesota. She attended Grinnell College in Grinnell, Iowa and received a Bachelor of the Arts degree in May 1996, with majors in Biology and English, graduating with Honors in English. She received a Post-Baccalauareate certificate in Environmental Education from the University of Minnesota at Duluth in conjunction with Wolf Ridge Environmental Learning Center (Finland, Minnesota) in May, 1997 and graduated with High Distinction. She received her Master of Science degree in December 2000 from the University of Minnesota Twin Cities for her thesis entitled "The lichen genus *Sticta* in the Great Smoky Mountains." She has published five articles on which she is the first author, including: *Multiple horizontal gene transfers of ammonium transporters/ammonia permeases from prokaryotes to eukaryotes: Toward a new functional and evolutionary classification* (Molecular Biology and Evolution, 2011); *RNA silencing of mycotoxin production in Aspergillus and Fusarium species* (Molecular Plant-Microbe Interactions 18, 539-545); *The ST cluster revisited: Lessons from a genetic model* (In: Aflatoxin and Food Safety, H. Abbas, ed. (Boca Raton, FL: CRC Press), pp. 117-136, 2005); *Signaling events connecting mycotoxin biosynthesis and sporulation in Aspergillus and Fusarium spp.*, (In: New Horizon of Mycotoxicology for Assuring Food Safety, T. Yoshizawa, ed. (Takamatsu, Kagawa, Japan), pp. 139-147, 2005), and *The lichen genus Sticta in the Great Smoky*

Mountains: A phylogenetic study of morphological, chemical, and molecular data (The Bryologist Vol. 106, pp. 61-79, 2003).

Since obtaining the bachelor's degree, she has received several awards including: The Anne T. and Robert M. Bass Fellowship for Undergraduate Instruction (2010 – 2011) for her original course entitled "Scientific reading and writing: Epigenetics" which she then taught at Duke University in the Spring of 2011 (Biol 295S); FESIN/Nordforsk travel grant to Edinburgh, Scotland, 2010; Sigma Xi, Grant-in-Aid of Research, 2008; Mycological Society of America, Graduate Research Award, 2008; Mycological Society of America, Luttrell Mentor Travel Award, 2008; Mycological Society of America, Best Poster Award, 2007; Duke University, Department of Biology Grant-in-aid, 2006; Mycological Society of America, Alexopolous Mentor Travel Award, 2006; Duke University, James B. Duke Fellowship (4 years), 2004; University of Minnesota Dayton-Wilke Natural History Fund grant, 1999; and a University of Minnesota Graduate Fellowship for first-year students, 1998.