## **EVOLUTIONARY HISTORY OF THE SEPTIN CYTOSKELETON PROTEINS IN**

#### **OPISTHOKONTS**

by

**Benjamin Auxier** 

B.Sc. (Hons. Plant Biology), UBC, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

## THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Botany)

## THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

November 2017

© Benjamin Auxier, 2017

#### Abstract:

Septins are cytoskeletal proteins important in morphogenesis, cell division and establishing and maintaining cell polarity. Over the course of more than a billion years as Animals and Fungi originated and diversified, their septin genes duplicated and diverged, giving rise to paralogs that encode modular proteins. The septin monomers assemble into heteropolymeric higher order structures that affect cell form by creating physical barriers to diffusion or serving as scaffolds organizing groups of diverse proteins. Here we take advantage of newly sequenced genomes to track the history of septin gene expansions and losses within the phylogeny of Animals and Fungi, including their close protist relatives. By sampling broadly across Opisthokonts, we identified the likely presence of early-diverging animal lineages within Groups 4 and 2A and discovered a novel group of fungal septins not found in Ascomycetes or Basidiomycetes. We hypothesized that previously identified sequence conservation is linked to interface interactions. Using protein homology folding, we mapped interacting residues across Opisthokonts and found that all previously identified motifs were involved in interface interactions, and contained almost all interacting residues. As septin subunit interactions are likely driven by residue identity, we categorized the interacting residues and found specific interface residues associated with each septin Group. We suggest that these residues may explain patterns of septin subunit binding affinity. Notably, we found that Group 3 septins show little conservation within the polybasic region that forms the first alpha helix, found in the NC interface of other septin Groups. This may explain the capping role of Group 3 septins in the yeast septin octamer. With increased sampling, this work identified increased diversity of Opisthokont septins. These proteins show patterns of sequence conservation that are largely driven by conserved interface interactions, in addition to binding of GTP. This work highlights the likely duplications that predate the Opisthokont ancestor, and the structural constraints that shaped the evolution of these multi-purpose septins.

Additionally, I attempted to validate and optimize an *Agrobacterium*-mediated transformation protocol for the chytrid fungus *Blastocladiella*. While I was unable to conclusively repeat previously published experiments, my work highlighted the difficulties in transformation of these distinct fungi.

ii

## Lay Summary:

The molds and mushrooms that we commonly encounter share an ancestor with animals roughly one billion years ago. This ancient ancestor was likely an aquatic organism that we could call a protist if we saw it today. Fungi and Animals took separate evolutionary paths over this time, resulting in distinct morphologies, yet share many genes, some of which determine cell shape. Initially I attempted to directly assess these genes in a chytrid fungus with a distinct morphology, but was unsuccessful. Instead, I investigated the evolutionary history of a shared morphological gene family. I surveyed these genes from diverse Animal and Fungal lineages, and found more shared genes than previously thought. I also found that the evolution of these genes was constrained by required interactions between genes. This work can be used to further our understanding of gene evolution in Animals and Fungi, and how their interactions evolve.

#### Preface:

The work presented in Chapter 1 was the result of a collaboration with Dr. Michelle Momany at the University of Georgia. The structural analysis of the septin genes was performed with the assistance of Dr. Momany at the University of Georgia. The initial idea to focus on septins as a morphological driver was formulated by Ms. Jaclyn Dee in the Berbee Lab, but I performed all of the analyses presented here. The results of this chapter are being prepared for publication as a journal article co-authored with Dr. Berbee, Jaclyn Dee and Dr. Momany. I wrote the manuscript with editing assistance from Drs. Berbee and Momany.

The work presented in Chapter 2 on *Agrobacterium*-mediated transformation was performed myself at UBC with the assistance of Dr. Berbee and Ms. Dee.

An outline of the work on transformation of *Blastocladiella* was presented in 2016 to the Mycological Society of America under the heading of:

Progress towards genetic transformation in a Blastocladiomycete fungus. Auxier B, Dee J, M Berbee, Mycological Society of America, August 2016

The results of the septin phylogenetic and structural analysis was presented in 2017 to the Mycological Society of America under the title of:

Evolutionary expansion of the cytoskeletal septin proteins predates diversification in fungal morphology. Auxier B, Dee J, M Momany, M Berbee. Mycological Society of America, July 2017

### **Table of Contents**

Abstra	act	
Lay S	umma	ryiii
Prefac	e	iv
Table	of Cor	ntentsv
List of	f Table	svi
List of	f Figur	esvii
Gloss	ary	viii
List of	f Abbre	eviationsix
Ackno	wledg	ementsx
Introd	uction	1
1	Analys	sis of Opisthokont Septins
	1.1 1.2 1.3 1.4 1.5	Introduction
2	Agrob 2.1 2.2 2.3 2.4 2.5	acterium-mediated Transformation of a Chytrid Fungus Introduction
Concl	usion	
Biblio	graphy	
Apper	ndix A	Supplementary Material for Chapter 141
Apper	ndix B	Agrobacterium-Mediated Transformation Protocol

## List of Tables:

Table A.1 : Sources of proteomes for this study	49
Table A.2 : Proteins used in homology modeling	50

Table B.1 : First Round F	PCR Fragment Amplification	51
Table B.2 : Second Rour	nd Overlap PCR Extension	51
Table B.3 : Final Product	Amplification with Restriction	Sites on Primers51
Table B.4 : Blastocladiell	a Primers and Genomic Reg	ons55

## List of Figures:

Figure 1.1:	Organization of a Septin Protein		
Figure 1.2:	Presence of Septin Groups Throughout Opisthokonts	11	
Figure 1.3:	Conservation of Sequence Identity is Related to Interface Inter	actions14	
Figure 1.4: Identity of Interface Residues Varies Between Septin Groups			
Figure 2.1:	Plasmids used in Agrobacterium-mediated transformation	26	
Figure 2.2:	Lifecycle and Hygromycin Sensitivity of Blastocladiella	27	
Figure 2.3:	First Southern blot of Genomic DNA from Transformants	28	
Figure 2.4:	Second Southern blot of Genomic DNA from Transformants	29	

Figure A.1:	RAxML Septin Phylogeny	.41
Figure A.2:	MrBayes Septin Phylogeny	.43
Figure A.3:	jPRIME Septin Phylogeny	.45
Figure A.4:	Septin Group Ancestral State Reconstructions	.47
Figure A.5:	Septin Alignment with Structural Annotation	.48

#### Glossary

Animal – The use of capitalized Animal indicates the monophyletic group encompassing Metazoans as well as related lineages, resulting is a sister group to Fungi

BNI5 – Bud Neck Involved gene 5. Involved in cell division in *Saccharomyces*, and directly interacts with septins as shown by Bertin et al., 2008.

CDC3/10/11/12 – Cell Division Cycle genes originally identified in *Saccharomyces* mutants. CDC3/10/11/12 were found to be the four core septins in *Saccharomyces*.

Fungi – The use of capitalized Fungi indicates the monophyletic group that is sister to Animals, encompassing terrestrial Fungi as well as parasitic lineages.

G Interface – The interface between septin subunits composed of residues near the GTP-binding domain

G1/3/4 - GTP-binding boxes. G1/3/4 are unique to septins.

Group 1/2/3/4/5 – Septin phylogenetic clades identified by Pan et al., 2007. Group 3,4,5 were originally found to be restricted to Fungi.

NC Interface – The interface between septin subunits composed of residues from the amino and carboxy termini of the protein.

PB Domain – A region unique to septin proteins that resides near the N terminus and binds phosphoinositol in plasma membranes using the multiple basic residues (commonly referred to as either Poly Basic or Phosphoinositol Binding).

S1/2/3/4 – Motifs of unknown role identified by Pan et al., 2007.

Sept2/6/7 – Septins from Homo sapiens, used to obtain crystal structures.

#### **List of Abbreviations**

AIC – Aikake Information Criterion. A criterion used for model selection based on model likelihood and number of parameters.

CDD – Conserved Domain Database. An online database used to identify conserved regions in protein sequences. Provided by the National Center for Biotechnology Institute.

CTE – Carboxy Terminal Element. A septin region found terminal to the SUE, and highly variable between septin Groups. This regions forms coiled-coil domains in many septins.

GTP – Guanosine Triphosphate. Septins bind and hydrolyze GTP during oligomer assembly.

HOS – Higher Order Structures. Formed by septin heterooligomers, found as rings or bars inside cells.

LG – Le and Gascuel protein substitution matrix. Used to infer probability of protein residue change.

NTE – Amino Terminal Element. A septin region found proximal to the PB domain, and highly variable between septins Groups. This region binds membranes in many septins.

SUE – Septin Unique Element. A region found next to the CTE, unique to septins. Residues in this region bind both the NC and G interface.

#### Acknowledgements

I would like to thank Dr. Mary Berbee for both allowing me to spend considerable time and resources on genetic transformation of a chytrid, and then providing the support for the subsequent phylogenetic analysis. Her curiosity for the strange and wonderful world of Fungi is infectious. Thanks also to Dr. Michelle Momany, and all of her lab members for hosting me in Georgia, and showing me the mysterious world of septin biology. All the members of the Berbee laboratory were invaluable for me to formulate my ideas, especially Jaclyn Dee in the lab, and Ludo Le Renard and Anna Bazzicalupo for their endless ability to discuss evolution from the perspective of a protein.

A piece of work such as this is not possible without family support, and I would like to point out my wife Jenny Auxier for her unwavering support throughout this period, including several extended absences.

This work was funded by the National Sciences and Engineering Research Council of Canada, both through and the Alexander Graham Bell CGS-M and through a Michael Smith Foreign Study Supplement to allow me to travel to the University of Georgia. My degree was also funded through a Botany Department TAship, which provided an opportunity to consider some of the more fundamental parts of fungal biology and biology in general, in addition to providing financial support.

#### Introduction

I am interested in the processes that underlie the evolution of morphological complexity in Fungi. The ancestor of this diverse group was a unicellular, aquatic, protist, yet we now see many morphologies represented including flagellated, non-flagellated, unicellular, and filamentous. Across Fungi we see a diversity of cellular organizations, and repeated evolution of filamentous growth with multiple reversions to a unicellular state. These repeated changes between cellular morphologies on an evolutionary timescale seems to be a special feature of Fungi, as transitions in other lineages like animals, plants, or heterokonts are much less common. Presumably this distinctive number of morphological transitions in Fungi is due to differences in the genetic machinery of Fungi, which may be more pliable than in other lineages. To investigate this Fungal morphogenesis machinery, I used two approaches.

The first approach I present is a detailed analysis of a family of cytoskeleton proteins, septins, that in Fungi are involved with many aspects of cellular morphology as well as other cellular processes. I wanted to investigate the origin of these proteins, and compare patterns of gene evolution to organismal evolution, especially morphological complexity. I used a combination of bioinformatics and phylogenetics to understand the evolution of these septin proteins and the constraints on these structural proteins. Previous analyses lacked the diversity of genomes from early-diverging lineages now available, and I believed these especially would yield new data.

A second, more direct, approach to understand early evolution of Fungi was to refine a transformation system in an early-diverging zoosporic fungus to assess the role of morphogenesis genes. Using such a system, it would be possible to localize proteins and collect mutant phenotypes, which could be compared to model systems like *Saccharomyces*. The well-known hyphal form did not evolve until after the zoosporic fungi diverged, yet they share a surprising number of morphogenesis-related genes. This indicates that the machinery for morphological transitions was present prior to hyphal growth, but the role of this machinery is these non-hyphal lineages is unknown. A reliable transformation system would enable me to investigate these questions.

#### Chapter 1: Analysis of Opisthokont Septins

#### 1.1 Introduction:

From their common unicellular ancestor ~1.3 billion years ago, Opisthokonts, the clade uniting animals and fungi, inherited a core set of genes which through duplications, deletions and other modifications gave rise to an astounding range of morphological and physiological diversity (Ruiz-Trillo et al., 2007). Septin genes are among those that appear to have expanded during opisthokont diversification. Given their important roles in morphogenesis, septin proteins may have contributed to the evolution of opisthokont complexity and diversity. Alongside the better known proteins that form actin filaments, intermediate filaments, and microtubules, septins assemble into filaments or rings that form part of the cytoskeleton (Mostowy & Cossart, 2012). In both Animals and Fungi, septins form physical barriers to diffusion and also anchor proteins to substrates such as the plasma membrane and endoplasmic reticulum (Bridges et al., 2014; Takizawa, DeRisi, Wilhelm, & Vale, 2000). In animals, septins are involved in cell division, and localize to the plasma membrane during cytokinesis (Spiliotis & Gladfelter, 2012). Perturbation of septins is associated with many diseases (Angelis & Spiliotis, 2016; Mostowy & Cossart, 2012). In Fungi, septins are involved in cell division and in determining morphology (Gladfelter, 2010; Oh & Bi, 2011). Septins are required for virulence in many fungal pathogens of plants and animals (Bridges et al., 2014; Kozubowski & Heitman, 2010; Momany & Talbot, 2017).



Figure 1.1 – Organization of a septin protein. A) Three-dimensional crystal structure from PDB (2QA5) in three orientations, connected with dashed arrows, each rotated by 90°. Position of GTP binding is located with arrow with solid line. Structures are coloured according to regions in B). Note that coiled coil regions are not resolved in the crystal structure. B) Linear representation of a septin gene, indicating the arrangements of the major septin elements. B) reproduced with permission from Pan et al., 2007.

Structurally, septin proteins (Fig 1.1) consist of an amino-terminal extension (NTE) that is highly variable, which borders the poly-basic (PB) domain that is thought to be involved in membrane binding, and the PB abbreviation also alludes to phosphoinositol binding (Bertin et al., 2010; Sirajuddin et al., 2007). The center of the protein is a highly conserved Ras-type GTPase, which is thought to be solely for conformational shifting due to extremely slow hydrolysis of GTP (Sirajuddin, Farkasovsky, Zent, & Wittinghofer, 2009; Weirich, Erzberger, & Barral, 2008). Carboxy-terminal to the GTPase there is a septin unique element (SUE) that has a largely unknown role, although some studies have found involvement in interface binding (Sirajuddin et al., 2007). The Carboxy-terminus extension (CTE) of the protein is also highly variable and is predicted to form coiled-coils in many Opisthokont septins, contributing to forming septin higher order structures (HOS) (Pan, Malmberg, &

Momany, 2007). To form HOS, individual septins first form heterooligomers by interacting through their G- and NC-interfaces, and these heterooligomers interact with each other through the coiled-coil domains and the septins that cap the ends of the heterooligomers. The order of septins is non-random and apparently consistent among all the septin heterooligomers within a cell (Sellin, Sandblad, Stenmark, & Gullberg, 2011). Human septin monomers and their heterooligomers have been analyzed by crystallography and transmission electron microscopy, leading to an increasingly sophisticated understanding of the basis for monomer assembly at the level of interactions of interface amino acid residues (Sirajuddin et al., 2007). With available crystal structures of the human septins as spatial constraints, it becomes possible to model the three-dimensional structure of other orthologous Opisthokont sequences. This opens the door to recognizing interactions between conserved amino acid residues in aligned Opisthokont septins, which direct septin heterooligomer self-assembly.

New whole genome sequences from unicellular relatives of animals and from non-filamentous fungi that diverged from over a billion years ago now provide opportunities to relate patterns of gene evolution and morphological complexity with increasingly sophisticated organismal phylogenies. This improved sampling of Opisthokont diversity will allow us to infer where genes were lost or duplicated and to evaluate the origin of previously identified groups (Pan et al., 2007).

In this chapter, we take advantage of the new wealth of genome sequence data and reconcile septin gene copy number with the Opisthokont phylogeny. This allows us to evaluate time of origin of septin Groups as well as gains and losses. We combine phylogenetic information with protein modeling to test the hypothesis that interacting interface regions between septin monomers would correspond to previously unexplained conserved motifs in septin alignments. By comparing the identity of interface residues between septin groups, we hoped to suggest informed targets for further mutagenesis studies, the analysis of which may reveal rules for septin partner assembly.

#### **1.2 Materials and Methods:**

#### Septin searches and coiled-coil domain prediction

To identify Opisthokont septins, we downloaded the predicted proteomes of 22 taxa (Table A.1). We searched these predicted proteomes using PSI-BLAST using Saccharomyces cerevisiae CDC3 (NP 013418.2) as the initial guery and an e-value cutoff of 0.001. After three iterations of PSI-BLAST all the known septins from model organisms Aspergillus, Drosophila, and Saccharomyces were recovered. As an alternative search strategy we also used HMMER (Eddy, 2008), using a previous alignment as a search profile (Pan et al., 2007). To remove proteins that only shared sequence similarity due to the GTPase domain but lacked other septin domains, we used the Conserved Domain Database (CDD) (Marchler-Bauer et al., 2015) and a custom python script to retain those with recognized domains "P-Loop NTPase", "CDC3", or "CDC septin", or if they had at least two of three G boxes (G1, G3, G4) in the GTPase domain but no other conflicting domains such as "LvsM", or "DNA binding". These criteria allowed us to include S. cerevisiae CDC3, as not all three of its G boxes are recognized by CDD, and A. nidulans AspE for which CDD does not recognize a CDC3 or CDC\_Septin domain. Filtering retained 120 septins. As the criteria were based on previously known septins, we cannot exclude that we may have missed novel diversity. To assess potential septins in the Opisthokont ancestor, we searched the genomes of *Thecamonas trahens*, a member of the Opisthokont sister group Apusozoa, and the closely related breviate Lenisia limosa (Hamann et al., 2016), but no septin sequences were recovered (data not shown). These filtered proteins were aligned with COBALT (Papadopoulos & Agarwala, 2007) with default settings, and the alignment was visualized with Mesquite (Maddison & Maddison, 2016). To predict coiled-coil domains known to be involved in septin interactions, we used the MarCoil webserver with the 120 proteins (Delorenzi, 2016).

#### Phylogenetic analysis

We processed the septin alignment using Aliscore and Alicut (window size, 12) to remove areas where positional homology was doubtful (Kück et al., 2010; Misof & Misof, 2009), after pruning 344 sites remained. Using ProtTest3 with AIC (Darriba, Taboada, Doallo, & Posada, 2011), we chose LG + gamma + invariant sites as an appropriate model of evolution for subsequent analyses. With RAxML v8.0 using the

CIPRES portal (Miller, Pfeiffer, & Schwartz, 2010; Stamatakis, 2014), we performed 2000 maximum likelihood searches of the 344 site alignment and then 456 bootstrap replicates (a sufficient number based on -autoMRE option (Pattengale, Alipour, Bininda-Emonds, Moret, & Stamatakis, 2009). We also performed a Bayesian analysis with MrBayes v3.2.4 with the LG +gamma + invariant sites model (Huelsenbeck & Ronquist, 2001). We used two independent runs of 8 chains each, modifying the heating parameter to 0.06 to increase the swap frequency and ran the analysis for 200 million generations. We used TRACER to select a burn-in of 50%, by which point the average split frequency was below 0.01, and all parameter estimated sample sizes were over 200 after removing the burnin (Rambaut, Suchard, Xie, & Drummond, 2014).

We also used a current Opisthokont topology to inform a jPRIME analysis of the septin gene phylogeny. *jPRIME* increases phylogenetic resolution by using an ultrametric organismal phylogeny as an informative prior to parameterize gene duplication and loss rates (Åkerborg, Sennblad, Arvestad, & Lagergren, 2009; Sjöstrand, Sennblad, Arvestad, & Lagergren, 2012). For the required organismal phylogeny we started with an accepted topology (Torruella et al., 2015), which is shown in Figure 1B. Because our analysis involved only 22 out of the 83 species in Torruella et al.'s analysis, we re-estimated branch lengths starting with genomic data. With OrthoFinder we identified orthologous gene groups for the 22 taxa (Emms & Kelly, 2015), aligning the orthologs with MAFFT with the –auto setting and concatenating them with FasconCAT (Kück & Longo, 2014). With ProtTest3, we selected the LG model of evolution for the concatenated alignment. Using this alignment, we inferred branch lengths for Torruella's tree with RAxML v8.0. We transformed the tree to be ultrametric with r8s (Sanderson, 2003), applying the Turner-Nash method, and using crossvalidation to select a smoothing value of 14. We constrained the age of the basal node of the Opisthokonts to 1350 million years, within the range estimated by Parfrey et al. (Parfrey, Lahr, Knoll, & Katz, 2011).

We then ran two independent jPRIME analyses of the aligned septin genes using the ultrametric version of Torruella's organismal tree for 50 million generations each, using RWTY (Warren, Geneva, & Lanfear, 2017) to confirm that the split frequency standard deviation was below 0.01. TRACER showed that after discarding the first 50% of the trees, the estimated sample sizes were over 200 for all parameters except for the

gene duplication rate and the model of molecular evolution, which were both over 100. To summarize the jPRIME results with a Most Clade Credible Tree, we applied SumTrees from the DendroPy distribution (Sukumaran & Holder, 2010). We used the packages ape and ggtree (Paradis, Claude, & Strimmer, 2004; Yu, Smith, Zhu, Guan, & Lam, 2017) in R 1.0.143 to visualize the phylogenies. We rooted the septin phylogenies with the clade of Group 3 and 5 for reasons discussed below.

## <u>Septin genealogy reconciliation and ancestral state reconstruction with animal and</u> <u>fungal organismal phylogeny</u>

To reconstruct septin gene duplications and losses along the Opisthokont phylogeny, we reconciled the *jPRIME* septin phylogeny with the species tree using NOTUNG v2.8.1.7 (Vernot, Stolzer, Goldman, & Durand, 2008). We allowed rearrangements of the septin phylogeny for nodes with a posterior probability below 0.90, resulting in 2 equally parsimonious reconciliations. We used the average number of inferred septins from the two reconciliations as thickenings in the branches of the species tree in Figure 1B. To track the origin of septin groups within the organismal phylogeny, we reconstructed ancestral states with Mesquite v3.2. To designate Group identity, we anchored the largest monophyletic group possible around Aspergillus or Drosophila septins (excluding NP\_724659 as it did not consistently group with other two Drosophila Group 2 septins). The Group designations are shown in Supplemental Fig 1-3, with the Aspergillus and Drosophila anchor septins in bold. We used the Groups as recovered in the *iPRIME* analysis to code each taxa for presence or absence of each Group (see Fig. 1C). Using Mesquite, we traced the ancestral state with a transition matrix with gain and loss rates set to be equal (MK1), as well as independent (MK2), for each of the 7 Groups. In all cases, the MK1 model with equal gain and loss rates could not be rejected based on the Wilks hypothesis test (Fig A.4). The complete ancestral state reconstruction is found in Figure A.4, and nodes of interest are found in Figure 1A.

As an additional test of the statistical support for orthology of Animal and Fungal sequences within septin groups, we first constrained all the animal septins within one clade and all the fungal septins within another. We then found the most likely tree from 2000 independent searches, given the constraint using RAxML. Again, with RAxML we calculated the per-site likelihoods of the data given the most likely tree found initially

and that given the constrained tree. We used the per-site likelihoods as input for an AU test with Consel (Shimodaira & Hasegawa, 2001).

## Prediction of interface interacting residues of group 1 and 2 septins based on crystal structures

To predict septin interface residues, we designed a workflow based on a recent publication on interface residue evolution (Ames, Talavera, Williams, Robertson, & Lovell, 2016). First, for each Opisthokont taxon a single Group 1 and Group 2 septin was chosen, arbitrarily choosing one protein when paralogs were present. As several early-diverging lineages lacked either Group 1 or Group 2 septins, we could not use these taxa for the analysis. To increase the size of the dataset, we added *Homo sapiens* Sept2/Sept6 and *Caenorhabditis elegans* unc59/unc61 as they could be assigned to orthologous septin Groups without requiring phylogenetic analysis (Cao et al., 2007). This resulted in 17 taxa (34 septins) for further analysis, summarized in Table A.2.

Next, we downloaded crystal structures for the Group 1-1 G-interface dimer (2QA5) and Group 2-2 G-interface dimer (3TW4) from PDB (Berman et al., 2000; Serrão et al., 2011; Sirajuddin et al., 2007). We also extracted the Group2-2 NC-interface homodimer, the Group1-2 G-interface heterodimer, and the Group1-2 NC-interface heterodimer from the human septin hexamer (2QAG) using PyMOL (Sirajuddin et al., 2007; The PyMOL Molecular Graphics System).

We then individually aligned the septins to the appropriate crystal structure using SALIGN implemented in Modeller (Sali & Blundell, 1993). Using Modeller, we produced twenty-five independent homology-folded models of each Opisthokont septin dimer, selecting the model with the lowest objective score "molpdf". To this model we added hydrogen atoms with the program reduce (Word, Lovell, Richardson, & Richardson, 1999) with the flags –build and –FLIP to allow for sidechain rotation based on steric hindrance. We assessed the distance between residues in the reduced dimer structures with the program probe (Word, Lovell, LaBean, et al., 1999) using the default probe radius of 0.25Å, with the flags –Unformatted to allow parsing of the raw data, –Oneway to ignore chain interactions within a subunit, and –NOCLASHOUT to ignore clashes

between the peptide backbones resulting from improper modelling. These settings were designed to retain only interactions from either hydrogen bonding or Van der Waals interactions. We eliminated duplicate hits from the multiple atoms in a common residue with a bash script. The locations of the residues were mapped onto *S. cerevisiae* CDC3 (NP\_013418) as a reference using an in-house script based on a MAFFT alignment with default settings of all the Group 1 and Group 2 septins in Table A.2.

We measured the conservation of septin protein sequences using Jensen-Shannon divergence, which compares the distribution of residues in each column of an alignment with a null model based on entropy, using the webserver (Capra & Singh, 2007) and a window size of one. As input, we used all the septins excluding Group 5 septins, 103 sequences. We reasoned that since there is no evidence that Group 5 septins are involved in dimerization (Hernández-Rodríguez et al., 2014; Kozubowski & Heitman, 2010) with other septins, they would add noise to the analysis as they likely lack conserved interface residues. To visualize the conservation of the residues that are consistently found interacting in Opisthokonts, we made WebLogos of the interfaces found in at least 10 of the 17 taxa, using the webserver interface (Crooks, Hon, Chandonia, & Brenner, 2004).

#### 1.3 Results:

#### Septin classification into groups

All sampled Opisthokont genomes contained predicted septins, ranging from 2 in *Sphaeroforma* to 14 in *Phycomyces* (Fig. 1.2). Previous phylogenetic analysis showed that septins fell into 7 clades designated Group 1A, 1B, 2A, 2B, 3, 4, and 5 (Pan et al, 2007). As a criterion to enable us to compare our Groups to those of Pan et al., we recognized clades based on the previously classified *Aspergillus* ortholog in each (indicated in Fig. A.1-3 with bold gene names). Two groups (1B and 2B) lacked *Aspergillus* members, so we recognized these based on *Drosophila* sequences (Fig. A.1-3, bold gene names). By these criteria, we could place most Opisthokont septins from recently sequenced genomes within the previously recognized clades. The 7 clades accommodated 114/120 sequences in RAxML analysis (Fig. A.1), 99/120 in MrBayes analysis (Fig. A.2), and all 120 septins with jPRIME (Fig. A.3). Septin Groups most commonly contained zero or one gene copy per organism, with some taxa such as *Phycomyces*, *Basidiobolus*, and *Drosophila* containing duplications in more than one

Group (Fig. 1.2 C). Group 1A septins were the most consistently present Group, appearing as single copy genes in 17-20 of 22 taxa (depending on the analysis) and as duplicated genes in one to two taxa. While most fungi had a single copy of a Group 2A septin, *Encephalitozoon* and *Paramicrosporidium* had none, and *Basidiobolus* and *Phycomyces* each had two. Animal lineages, when present, had a single Group 2A septin. *Saccharomyces* had three copies of septins classified as Group 3 in RAxML and jPRIME but these were divergent and difficult to place. In illustrating septin gene copy numbers across septin Groups, the heat map in Fig. 1.2 C shows that three analytical methods, ML, Bayesian and jPRIME largely agreed on the phylogenetic classification of septins.

The presence or absence of predicted coiled-coil domains was conserved within well-supported septin Groups but sometimes varied when genes were poorly resolved phylogenetically and from early-diverging protists. Coiled-coil domains were present in almost all members of Groups 1B, 2A, 2B, and 4 (Fig. A.1-3). Among Group 1B septins, only Batrachochytrium lacked the predicted domain. Salpingoeca XP\_004994451, placed in Group 2A by the RAxML and MrBayes analyses, seemed to lack the domain, but close inspection showed that the gene prediction was missing the 3' end where coiled-coil sequences would likely reside if present. Of the Group 3 septins, most (13/19) had a predicted coiled-coil domain. Group 4 septins had predicted coiled-coil domains with the exception of four genes of inconsistent phylogenetic placement (Fig. A.1-3). Group 1A lacked coiled-coil domains, again with the exception of poorly supported members such as Fonticula XP\_009497655 (Figs A.1-3). Capsaspora XP\_011270180, placed in 1A by jPRIME analysis was unique among our sequences in having a predicted N-terminal coiled-coil domain (Fig. A.3). None of the Group 5 septins except Catenaria 1512492 had a strongly predicted coiled-coil domain and its domain was positioned in the center of the sequence rather than the end as in other septins. One of the Basidiobolus septins also had a weakly predicted coiled-coil domain (Fig. A.1-3).



Figure 1.2: Analysis of early-diverging lineages provides evidence of ancestral septin duplications. A) Shaded area of each pie chart indicates the proportion of likelihood that the specified ancestor had a member of a septin group. Group 2 may be monophyletic or Group 2A and 2B may have arisen independently; reconstructions resulting from these two alternatives are illustrated. B) Gene duplications reconstructed within a species phylogeny. Branch thickness represents the NOTUNG inferred number of septin gene copies using the jPRIME septin gene phylogeny. Stars indicate a change in morphology of a lineage. This species phylogeny was used to guide jPRIME and ancestral state reconstruction. C) Cell shading indicates copy numbers of septin genes representing each septin group as classified by each of three phylogenetic methods. Classification of septins from an organism sometimes differed depending on the analysis method, reflecting uncertainty in phylogenetic placement of divergent sequences.

## Increased taxon sampling captures ancient gene duplications and losses throughout the evolutionary history of septin evolution

To estimate the timing of the origin of septin Groups, we performed ancestral state reconstructions. The reconstructions suggested that the ancestor of animals, fungi, and related protists had septin Groups 1A, 4 and possibly 1B and 5 (Fig. 1.2A). To explain the origin of these Groups, 1-4 gene duplications must have preceded the divergence of fungi from animals. This was consistent with the NOTUNG reconciliation, which indicated the presence of four septins (resulting from three gene duplications) in the Opisthokont ancestor prior to the divergence of fungal and animal lineages (Fig. 1.2B). Among fungi, the ancestor of the walled osmotrophic fungi (*Batrachochytrium* through Schizophyllum) gained a Group 3 septin, and we see the appearance of Group 2A (Fig. 1.2A). Morphologically simple yeasts (Saccharomyces and Cryptococcus) had as many septins as the morphologically complex mushroom forming fungus Schizophyllum. Among the newly analyzed septins from early diverging animal lineages (Sphaeroforma to Salpingoeca Fig. 1.2B and C) and fungal lineages (Fonticula to *Paramicrosporidium*, Fig. 1.2B and C), the representation of septins across the seven Groups is patchy, consistent with repeated gene losses over an immense period of geological time.

As expected, including additional early-diverging lineages not only increased the diversity of septins recovered, but also the challenges involved in phylogenetic resolution (Figs. 1.2C, A.1-3). The point of origin of Group 2 was difficult to assess because it was unclear whether 2A and 2B (Pan et al., 2007) together formed a monophyletic group (Fig. A.1-3). When we assumed that they were not monophyletic, the Opisthokont ancestor was reconstructed as having neither. When we assumed 2A and 2B were sister clades as suggested, although without statistical support by the jPRIME analysis (Fig. A.3), then Group 2 was reconstructed as present in the Opisthokont ancestor. The absence of recognizable septins from protist relatives of opisthokonts made outgroup rooting of the septin gene tree impossible. We rooted the septin gene tree along a branch present in all inferences (Figs. A.1-3) that separated Groups 3 and 5 from Groups 1A, 1B, 2 and 4. This allowed for consistent presentation (Figs. A.1-3) but a possible artifact of the rooting was the inference that the Opisthokont ancestor had a ~60% likelihood of having had a Group 5 septin (Fig. 1A).

#### Highly conserved residues are in GTP binding and monomer interaction surfaces

Previous analysis of Opisthokont septins revealed three highly conserved motifs involved in GTP binding (G1, G3 and G4) along with five motifs of unknown function (S1, S2, S3, S4 and SUE) (Pan et al., 2007). To determine if these conserved residues were also present in septins from recently sequenced early-diverging Opisthokonts, we analyzed the phylogenetic dataset with Jensen-Shannon divergence (Fig 2A) (Capra & Singh, 2007). As expected, all three GTP binding motifs were highly conserved. The S2 motif showed similarly high levels of conservation, while S1, S3, S4 and the SUE motifs were conserved at somewhat lower levels.

All of the septins that we analyzed shared a small number of highly conserved sites (Fig. 1.3), without which we could not have recognized or included them. However, outside of the few conserved sites, septins from the early-diverging organisms showed little phylogenetic signal due to minimal amino acid sequence similarity even among sequences that appeared to be orthologs. To understand why residues not required for GTP binding might be conserved, we used known structures of Group 1 and 2 heteropolymeric septin filaments from *Homo sapiens* to model the structures of orthologous pairs from diverse Opisthokonts. Because septin monomers that interact with each other must be compatible, we focused on interface regions.

The conservation of interface residues is shown in Figure 2. Of the 520 residues in the *Saccharomyces cerevisiae* CDC3 reference septin, 147 (28%) had a Shannon-Jensen conservation score greater than 0.50 when calculated using all septins except for Group 5, and all residues with greater than 0.5 conservation were found within the central region of the protein. To identify residues consistently involved in septin-septin interactions, we used a criterion of interaction in the modelled dimers in at least 10 of the 17 taxa. With this criterion, there were 29 interacting residues across the G interface, and 20 across the NC interface. In the G interface, 22 of the 29 residues were in previously identified regions/motifs, and 14 of the 20 residues in the NC interface were in previously identified conserved regions.



Figure 1.3: Septin interface residues may account for sequence conservation outside of the GTP-binding site. A) Our predicted septin interface interacting regions correspond to conserved regions of the septin proteins. The solid line is a plot of Shannon-Jensen conservation with values above 0.5 shaded. Red columns indicate the proportion of the 17 Opisthokont taxa with a conserved NC interacting group (ig) residue. Blue indicates the proportion of the 17 taxa with a conserved G interface ig residue. Residues involved with the GTP-binding active site are indicated with downward-facing black arrows. The septin gene diagram at the bottom of 2A was adapted from Pan et al., 2007 and it shows the congruence between previously identified conserved regions with predicted septin interacting regions as well as the GTP-binding site. B) Diagram of a septin monomer showing the ig residues at the NC and G interfaces. Coiled-coil is shown as a curved line at top of the monomer. C) Model showing how monomers from three different septin groups in humans (pink, Sept2, Group 2B; yellow Sept6, Group 1B; green Sept7, Group 2B) may form a symmetric septin hexamer (Sirajuddin et al., 2007). Crystal structures from these septins were the basis for homology folding. The ig residues, colored as in (B), indicate the putative binding sites between septin partners.

All of the previously defined conserved regions included interface residues, often with additional interface residues found in close proximity. Looking from the N- to C-termini, the 11 residue PB region had 4 NC interface residues (Fig. 1.4., among sites 110-118), and the sequence just beyond but still within a 3-residue window of the PB region contributed two additional NC interface residues (Fig. 1.4., among sites 108-118). Within G1, G3 and G4 boxes, we predicted 1, 1, and 3 G interface residues respectively. Considering the neighboring 3-residue windows increased the number of included residues to 1, 3 and 5, respectively. The 6 residue S1 motif had 5 NC interface residues. Within the S2, S3 and S4 motifs were 1, 5 and 2 G interface residues and including sites within a 3-residue window increased total numbers of interface residues to 3, 7 and 2 G. The SUE region had 9 G interface residues and 7 NC interface residues, and was the same if looking at a 3-residue window. Looking at the domains and motifs along with their neighboring 3-residue windows, a total of 25 of the 29 G interface and 18 of the 20 NC interface residues were associated with previously identified conserved regions.

#### Conservation of interface residues varies among septin groups

Having identified a set of residues consistently involved in the Opisthokont septin interfaces, we wanted to check the conservation of the interface residue identities. Septins interact in a predictable manner, and patterns in the interface residues may explain this phenomenon. We separated the septins by Group identity and looked for patterns in the conserved interface residues. We split the septins into Groups 1A/1B/2/3/4/5 based on the jPRIME MCCT, and made WebLogos from the 29 residues of the G interface and the 20 residues from the NC interface (Figure 1.4), using numbering from the reference *Saccharomyces* Group 2 septin CDC3. Overall, the Group 5 septins showed lower conservation of interface residues although residues such as D289, R360, W364, and H374 were conserved. The other septin Groups showed higher overall interface residue conservation. Position 129 in the G1 Box was strongly conserved with G for all Groups. Near the G3 box, G209 and D210 were highly conserved, and position 211, 213, and 214 showed conserved differences of amino acids of differing classes between Groups. In the S2-S3 region, the only difference of



Figure 1.4: Weblogo showing patterns of amino acid conservation and diversity across aligned motifs and domains of the septin groups.

note was an acidic E in Group 3 at site 266 that replaced the proline in Groups 1A/2/4. The G4 box did not have any interface residues that differed notably between Groups. The G interface residues in the SUE were all the same residue class, except for position 361, which was variable both between and within Groups.

The 20 residues of the NC interface showed similar patterns to the G interface (Fig. 1.4), with Group 5 septins having the lowest overall conservation. The PB region had many differences between Groups, and notably Group 3 had overall low sequence conservation similar to Group 5. Position 192 was found as an interacting residue, but was not near any previously defined conserved regions, and was mainly found as E, but with low conservation. In the S1 region the E at 237 and the R at 242 were highly conserved in all Groups, but 239 was variable between Groups, 240 and 241 were variable both between and within Groups. The Septin Unique Element contained conserved residues such as YE at 401-2. Position 386 was R in some Groups, but not conserved in other Groups. Position 393 was a mix of basic [K/R] in Groups 2/3/5 and hydrophobic I in Groups 1A/1B/4. Position 404 varied between [Y/F] which are both bulky aromatic amino residues.

Group 3 septins may lack the  $\alpha$ 0 helix that is unique to septins. This is suggested first by the lack of conservation of Group 3 NC interface residues in the poly-basic region, which fall inside the first alpha helix. Secondly, the distance between the start of the N-terminus and the G1 box is unusually short in Group 3 proteins (Fig A.5) compared with other septins.

#### 1.4 Discussion:

#### Determination of septin group origins

Our analysis of septins from genomes of early-diverging fungi and animals points to multiple ancient septin duplications. Cao et al. (2007) concluded that septin Groups diversified before vertebrate and invertebrate animals diverged. Our results show evidence of Holozoan septin diversification even before animals became multicellular. Like Pan et al. (Pan et al., 2007), we found strong evidence for orthologous Group 1A septins among animals and fungi. We tracked Group 1A septin sequences in genomes that were unavailable to Pan et al., including diverse early diverging fungal phyla and

animal genomes including *Trichoplax*, a placozoan multicellular animal, and *Monosiga*, a collar flagellate from among protist relatives of animals. Sampling from this broad phylogenetic distribution further supported Pan et al's suggestion that a Group 1A septin was present in the Opisthokont ancestor.

With additional genomes available to us, we found a more complex vision of septin evolution, with early diverging species suggesting broader representation of septin Groups among animals and fungi than previously detected. Pan et al. had been unable to differentiate whether Group 5 either arose early and was lost in Ascomycete yeasts, or was gained relatively recently in fungal evolution. Improved sampling of septins from early-diverging lineages brought clarity, showing that Group 5 was ancestral in the clade of fungi from Chytridiomycota to Basidiomycota (see Fig. 1.2). In addition, our analysis suggested that three septins in addition to 1A may already have been present in their shared Opisthokont ancestor. This was because genomes of some early-diverging protistan relatives of both animals and fungi had septins classified in Groups 1B, 4 and perhaps 2. An important qualification here is that the phylogeny of early splits among septins was poorly resolved, and uncertainty in classification of highly divergent sequences limits resolution of phylogenetic age of origin.

The polymerization of septins into hetero-oligomeric rods and higher order filaments interacting with membranes and other proteins may be unique to Opisthokonts. Septins may have arisen early in the evolution of eukaryotes, having been detected in green algae, brown algae and ciliates as well as Opisthokonts (Nishihama, Onishi, & Pringle, 2011). Most non-Opisthokont septins lack a coiled-coil domain, seeming to use hydrophobic interactions to interact with membranes rather interacting with membranes through higher order filaments (Nishihama et al., 2011). Supporting the idea that the common ancestor to the Opisthokonts had at least two paralogs, one with a coiled-coil domain, and at least one lacking the domain, each taxon in our analysis except for *Paramicrosporidium* had one or more of each. Septin monomers with coiled-coil domains in *Saccharomyces* Group 4 CDC12 and Group 2A, CDC3 are essential for polymerization of septins into normal, stable octamers (Bertin et al., 2008; Versele et al., 2004). Coiled-coil domains are further essential in interacting to organize octamers into filaments (Bertin et al., 2008) and in morphological development, with deletion of coiled-coil domains producing aberrant morphology in the ascomycete

yeasts *S. cerevisiae* and *Ashbya gossypii* (Meseroll, Howard, & Gladfelter, 2012; Versele et al., 2004). The role of septin monomers without coiled-coil domains is less clear, but a pair of Group 1A monomers designated as Cdc10 in *Saccharomyces* form the central doublet in octameric septin rods and in *Homo* hexamers a pair of Group 2A Sept2 septins occupy the same position (Sirajuddin et al., 2007; Versele et al., 2004). In *Aspergillus* the ortholog AspD is non-essential even though it is inferred to form the central dimer of the *Aspergillus* octamer (Hernández-Rodríguez, Hastings, & Momany, 2012), but recent work has shown that it may have a role in nuclear maintenance in these syncytial organisms (<u>Ilkay</u> Dörter, pers. comm.). Predicted coiled-coil domains were not randomly distributed through the phylogenies of septin groups containing animal and fungal lineages. A coiled-coil domain was absent from almost all Group 1A members while coiled-coils were generally present in Groups 2, 1B and 4. This level of conservation over the roughly 1.3 billion years of evolution represented by each of these septin Groups is consistent with selection for different but important conserved functions.

#### Rooting of septin phylogeny

Previous analyses have used myosin or septins from non-Opisthokonts lineages as outgroups (Nishihama et al., 2011; Pan et al., 2007) but their limited sequence similarity with Opisthokont septins limited their phylogenetic usefulness. A well-resolved septin phylogeny including non-Opisthokonts outgroups might allow us to evaluate whether septins from Opisthokonts are indeed homologous with septins from green algae and protists (Nishihama et al., 2011). Unfortunately the limited length of septin proteins combined with long evolutionary timescales seems to reduce the possibility of a resolved phylogeny. Previously, Group 5 septins have been clustered as sister to ciliate septins (Nishihama et al., 2011). This topology would indicate lack of monophyly of Opisthokont septins, with Group 5 shared with other Eukaryotes and Groups 1-4 restricted to Opisthokonts, except that it did not receive statistical support and it may be an artifact of overall poor resolution. If rooting opisthokont septins with Group 5 and non-opisthokont septins as the outgroups, then heterooligomer-forming septins with carboxy-terminal coiled-coil domains would be a derived character of Opisthokonts. But it is equally possible that Group 5 are more recently derived from other Opisthokont septins and lost the ability to form hetero-oligomers. This change in function may have

facilitated sequence divergence due to a change in selective pressures, to the point where their evolutionary origin has become difficult to track.

#### Sequence comparison shows the highly conserved septin core

We found septin proteins contain a highly conserved core with diverged C- and N-termini, consistent with previous findings (Versele et al., 2004). Many residues in this core are involved in the septin-septin interactions where the evolution of one septin would be constrained by the required interaction with its partner. In contrast, the N- and C-termini are not highly conserved, and their functional roles are less clear, although they are known to bind other proteins, such as BNI5 in *Saccharomyces* (Finnigan, Takagi, Cho, & Thorner, 2015). Neither subunit binding nor filament formation were affected by removal of the C-terminal extension containing the coiled-coil domain from *Saccharomyces* CDC11, a Group 4 septin (Bertin et al., 2008; Finnigan et al., 2015). Possibly, sequence changes in these divergent termini allowed recruitment of proteins that are utilized in the different morphologies found in Opisthokonts. As these divergent coiled-coil domains are not found outside of Opisthokont cytoskeletal diversity. As these regions have been difficult to crystallize, the conformation adopted by the different coiled-coil ends of the different septin Groups is unknown.

# Integrating evolutionary timescales with experimental crystal structures led to classification of septin-septin interface residues

Our analysis of the residues involved in the septin-septin interactions suggested roles for previously discovered S1-4 conserved regions (Pan et al., 2007). By using homology folding across the breadth of evolutionary diversity in Opisthokonts, we showed that the S regions are all involved in interface interactions. We recovered evidence of interface interactions in the alpha helix found in the polybasic region, consistent with what Agelis & Spiliotis found in an earlier study (2016). While almost all of the putative interface interactions that we detected were in regions and motifs previously identified as highly conserved, the glutamic acid flagged as an NC interface residue at position 192 had not previously been noted, possibly due to difficulties in aligning this region during previous motif surveys. The general correlation of interface residues with conserved gene sequences (Pan et al., 2007) is consistent with evolutionary expectations. The requirement for interactions between residues constrains

the sequence evolution in these regions, as both partners would require compensatory mutations for the heterooligomer to form properly. Our finding that Group 5 septins show little sequence conservation in interface regions provides additional evidence that these do not form part of the core septin oligomer.

Previous inferences of interacting regions from Sirajuddin et al., (Sirajuddin et al., 2007 Supplemental) were inconsistent with expected regions of conservation, possibly because the inferences were based solely on human septins. Angelis & Spiliotis (2016) categorized the mutations found in human septins in cancer cases, but the effect of these mutations, either *in vivo* or *in vitro*, is largely unknown. A wider systematic review of experimental mutagenesis of septins would be valuable. Combining a survey of the effects of a wide variety of mutations, as would be as possible in fungal systems, with the results of conserved interacting residues shown here may reveal interesting correlations.

Septin-septin interactions are thought to be governed largely by the residues in the G- and NC-interfaces, possibly with input from coiled-coil domain interactions. The G interface was generally more conserved than the NC interface, which is not surprising since its role in binding GTP further limits changes in sequence. Exemplifying this importance is the completely conserved tryptophan in the G interface (position 364 in Figure 3). Mutagenizing it to a much smaller alanine eliminated dimerization of Group 3 septins in yeast (Brausemann et al., 2016). The differences among groups in residues in G- and NC-interfaces could explain the preferential affinity of the individual Groups. Some septin Groups differed at one or two interface residues. Asparagine was for example conserved at site 211 within Group 3 but tyrosine was conserved at the same site in Group 4. The substitution of an amino acid with different chemical properties may have affected interface binding. One extension of this interface residue specificity could be the assignment of septin Groups based on these residues alone. While we see differences between Groups, these differences are only apparent when looking across many sequences, and it would be difficult to assign any single sequence due to sequence diversity, although this may be possible with greatly increased sample size.

#### Group 3 septins potentially lack one of the conserved characteristic alpha helices

Lack of conservation of the N-terminus in the Group 3 septins, which are only found in Fungi may be related to roles as capping proteins. Among Group 3 septins the

first alpha-helix,  $\alpha 0$  appear highly divergent. Recently, Brausemann et al., (2016) published a crystal structure of CDC11, the Group 3 septin from Saccharomyces, but unfortunately the first 20 residues of the protein had to be truncated to improve crystallization. Thus, in the CDC11 crystal structure it is not possible to assess whether the first helix is actually present (Brausemann et al., 2016). As the capping septin, the G interface of Saccharomyces CDC11 associates with other septins but the NC interface does not (Bertin et al., 2008). While the first 18-20 residues of CDC11 have been designated as  $\alpha 0$ , it is possible that these residues do not form a helical secondary structure. Regardless of whether or not a helix is formed, the deletion of this segment resulted in the abolition of filament formation in purified septins (Bertin et al., 2008), and mutation of 5 basic residues to uncharged glutamine disrupted septin function (Casamayor & Snyder, 2003). A conserved role of Group 3 septin as capping proteins combined with the limited sequence conservation we have recovered would be an interesting area of research. If the lack of this helix truly prevents further septin monomer stacking, then it raises the question of what prevents the further monomeric stacking in animal lineages, which lack Group 3 septins. It would be interesting to identify the capping proteins in multiple Animal lineages, and see if they are also predicted to lack the  $\alpha 0$  helix mechanism.

#### 1.5 Summary:

By including a diverse set of early-diverging Opisthokont lineages, we recovered more diverse sequences than were analyzed previously, allowing us to show that septin duplications were ancient with up to four septin paralogs in the Opisthokont ancestor. With alignments of these diverse sequences and homology folding, we found that interface residue conservation overlapped with evolutionarily conserved residues, indicating the tight relationship between septin partners over time. Septin groups including septins with coiled-coil domains were ancient in the Opisthokonts, suggesting that not only septin heterooligomers but also higher order filaments were part of the ancestral cellular tool kit of both animals and fungi.

#### Chapter 2: Agrobacterium-mediated Transformation of a Chytrid Fungus

#### 2.1 Introduction:

While the commonly encountered Fungi, molds and mushrooms, are sister groups, there are many extant Fungi from other, earlier, divergences. By comparing lineages resulting from separate early divergences, we can better understand the ancestor of Fungi, which frames how we view Fungal evolution. The earliest Fungal divergences are represented by parasitic lineages, such as Microsporidians, which are difficult to manipulate in the laboratory (Spatafora et al., 2017). Of tractable organisms, the earliest divergence is a zoosporic fungus, either Chytridiomycota or Blastocladiomycota as the phylogeny is not fully resolved yet (Spatafora et al., 2016). These aquatic organisms are unlike more commonly encountered Fungi, and were not truly recognized as members of Fungi until the advent of molecular phylogenetics (Walker & Doolittle, 1982) nevertheless their shared lysine biosynthesis pathway indicated a shared ancestry with terrestrial fungi (Vogel, 1961). Detailed study of chytrid fungi can help us differentiate homologous from derived processes and phenomena.

The species I am interested in is *Blastocladiella emersonii*, a member of the Blastocladiomycota. With a 48-hour asexual cycle, this monocentric fungus, having a sporangium positioned on top of a basal cell, provides a tractable system to investigate zoosporic fungal genetics. It was originally isolated from the grounds of the University of Pennsylvania, and has never been recovered from a wild habitat. While many papers have been published on biochemical studies of this organism (238 Results from Web of Science; September 25, 2017), all laboratory strains originate from this single isolate (Cantino, 1951). While the phylogenetic placement of zoosporic fungal DNA studies confirmed them as true Fungi (Bowman et al., 1992; Walker & Doolittle, 1982). But recent work has failed to determine the specific branching order between zoosporic Chytridiomycota and Blastocladiomycota, and terrestrial fungi (Chang et al., 2015; James et al., 2006).

Genetic transformation is a powerful tool to investigate biology, allowing for the direct interrogation of fundamental cell processes. Although transformation protocols are available for many fungal systems, few of these apply outside of Ascomycetes and

Basidiomycetes, and none has been developed for zoosporic fungi. In addition to a plethora of protocols for Ascomycetes and Basidiomycetes, protocols have been published for several species of Mucoromycota including *Absidia glauca* (Wöstemeyer, Burmester, & Weigel, 1987), *Mucor circinelloides* (Garre, Barredo, & Iturriaga, 2015; Gutiérrez, López-García, & Garre, 2011), (Revuelta & Jayaram, 1986). As well, there have been reports of *Agrobacterium* transformation protocols for two Mucoromycota species (Michielse et al., 2004; Nyilasi, Papp, Csernetics, & Vágvölgyi, 2008).

Previously, a proof of principle transformation of *Blastocladiella* was reported (Vieira & Camilo, 2011), the first report of transformation in a zoosporic fungus to my knowledge. Vieira and Camilo used *Agrobacterium*-mediated transformation, and introduced the hygromycin resistance gene, hph, or GFP into *Blastocladiella* in two separate experiments. The key results from their paper were zoospore fluorescence in the GFP-transformed sporangia, and PCR amplification of the hph gene after >10 generations. GFP-expression was obtained in zoospores in mature sporangia, but in the transformation vector the GFP gene was not paired with a selectable marker, and thus the transformation plates needed to be screened manually under a fluorescence microscope. As well, the hph insertion was not confirmed with Southern blotting.

As part of my Master's Thesis, I attempted to obtain a more reliable transformation system with zoosporic fungi, starting by replicating previous results (Vieira & Camilo, 2011). To do this I generated transformation vectors, and attempted to modify the transformation protocols to provide a more usable method. As well, I wanted to confirm the insertion with Southern blotting to ensure stable genomic integration of transferred genetic material. While I was, in the end, unsuccessful, my work may provide the initial procedure for someone else who wishes to embark on this task.

#### 2.2 Materials and Methods:

I obtained our isolate of *B. emersonii* from the laboratory of Dr. Suely Gomes in the Universidade de São Paulo, Brazil. The isolate was cultured on PYG media, transferred approximately every month, and stored at 4 °C, as cultures older than one month were often not viable. I obtained the *Agrobacterium* strain GV3101, the pCambia0380 vector backbone, and the pCSN44 plasmid containing the TrpC promoter fused to hph from the Breuil Lab, which they used for transformation of the ascomycete *Grosmannia* (Wang, DiGuistini, Wang, Bohlmann, & Breuil, 2010). The pCambia0380

vector is a basic backbone, and designed for expression in plants, although Wang et al. obtained expression in the Ascomycete fungus *Grosmannia* using this vector. Vieira and Camilo used the hygromycin resistance gene driven by the TrpC promoter, which although derived from *Aspergillus nidulans* (Yelton, Hamer, Souza, Mullaney, & Timberlake, 1983) is effective in a wide variety of Fungi.

To produce the transformation vector for *Blastocladiella*, I used the Sall site to insert the TrpC promoter with the hygromycin resistance gene from pCSN44 into pCambia0380, producing pBA011 (Figure 2.1). To obtain long flanking sequences for use in homologous recombination, we contacted the laboratory of Dr. Tom Richards (Exeter University) and he generously provided access to a now-defunct BLAST database for the B. emersonii genome and proteome. For a region for recombination I used cytochrome p450 enzymes, based on the Fungal Cytochrome p450 Database (http://p450.riceblast.snu.ac.kr/species.php?a=intro&spe\_id=1211&ref\_id=58474). Using closely related Allomyces p450 sequences from this database, I used BLAST searches to locate *B. emersonii* homologs. Due to a highly fragmented genome assembly, I was unable to obtain both 3' and 5' flanking regions for any one p450 enzyme. I was able to amplify the 5' region of a p450 gene labelled 1568 from the Blastocladiella genome using the primers Be1568TermF and Be1568TermR (Table B.4), which also introduced Smal restriction sites. I inserted the region upstream of the TrpC promoter in pBA011 to produce the 1568 plasmid (Figure 2.1). For transformation of Blastocladiella with pBA011 and 1568 plasmids I used the protocol found in Appendix B, similar to that of Vieira and Camilo (2011). I selected for transformants with PYG + 150 µg/mL hygromycin. Due to previous false positive PCR results with Agrobacteriummediated transformation of zoosporic fungi I did not perform PCR on putative transformants, instead using Southern blotting of putative resistant isolates. Southern blotting was performed with a DIG probe using the ROCHE PCR-labelling DIG kit (11-636-090-910) with the entire hph sequence as the probe, using primers SBhphBDF and SBhphBDR (Table B.4). Southern blots were prepared by digesting 1µg DNA with a restriction enzyme, and run on 1 % agarose gels. Blotting was performed according to the DIG probe manual.

Figure 2.1: Plasmids used in *Agrobacterium*-mediated transformation. Plasmid regions are shown proportional to actual size. Resistance gene region in shows in red arrows, and homologous region from 1568 p450 enzyme described in text is shown in green.



#### 2.3 Results:

*B. emersonii* only responded to selection with 150  $\mu$ g/mL of hygromycin when flooded with free liquid. Without liquid above the media, *B. emersonii* zoospores still developed into sporangia at rates up to 500  $\mu$ g/mL of hygromycin (Data not shown).

Only some developmental stages of *B. emersonii* were responsive to selection with hygromycin, diagrammed in Figure 2.2. At a certain point after the sporangia started developing, they became resistant to hygromycin. While this point of resistance was not precisely determined, it appeared to occur prior to the sporangia reaching their maximum size. Any small sporangia that were included in the zoospores co-cultivated with *Agrobacterium* would be of a sufficient size after the co-cultivation so that they would be resistant to hygromycin.

Figure 2.2: Hygromycin Sensitivity and Lifecycle of Blastocladiella.



Transformation with *Agrobacterium* strains carrying pBA011 and 1568 plasmids resulted in several apparent positively transformed colonies. Using the pBA011 plasmid, I isolated 3 independent potential *Blastocladiella* hygromycin-resistant (hph) transformants. After culturing them on selective media, I digested genomic DNA with Spel, and probed for the hph gene. As seen in Figure 2.3, the wild-type DNA had no probe signal, but in all three putative transformants, using Southern blotting, DNA hybridized to a band of the same size. The plasmid control was undigested for this blot, but the size of the hybridized band in the *Blastocladiella* samples was consistent with the size of the plasmid itself (9200bp, single cut from Spel). Figure 2.3: Southern blot of 3 *Blastocladiella* Transformants. Rows are wt, 3 independent transformants and uncut hph containing plasmid.



I then produced a second set of transformants using pBA011 and 1568. The results of a Southern blot of these transformants is shown in Figure 2.4.

Figure 2.4: Southern blot of 3 *Blastocladiella* transformants with 2 separate enzymatic digestions. The samples were digested with AvrII on the left, and Smal on the right. Wt lane is untransformed *Blastocladiella*. pBA011-1 is DNA from *Blastocladiella* co-cultivated with pBA011 plasmid. 1568-1 and 1568-2 are *Blastocladiella* co-cultivated with 1568 plasmid. pBA011+ is the plasmid positive control, and 1568+ is the plasmid positive control. Note increase in size of 1568+ compared to pBA011+ in AvrII digestion due to addition of homologous region, but same size with Smal digestion, due to excision of homolog region. Note that region of plasmid to be transferred, T-DNA, is 1.5 kb for pBA011 and 2.5 kb for 1568.



As seen in Figure 2.4, on the left, digested with AvrII, the *Blastocladiella* transformed with pBA011 had a hybridizing band of the same size as the plasmid itself. The two 1568-transformed *Blastocladiella* when digested with AvrII had hybridizing bands slightly higher than the pBA011 sample, consistent with the increased size of the plasmid as seen in the gel after digestion with AvrII. The sample from 1568-2 hybridized at a larger band size than 1568-1. Samples from the same isolates were also digested

with Smal (Fig 2.4), but the Smal digest did not appear to be complete. The bands on the blot were quite smeary, and this was also visible on the ethidium bromide-stained gel (not shown). Assuming the digestion was not complete, the smallest of the hybridizing band for each of the three samples was consistent with the AvrII-digested samples. The DNA from pBA011-1 was the same size as the plasmid control. The 1568 plasmid when digested with Smal was the same size as the pBA011 plasmid due to the *B. emersonii* homologous region, which was inserted with Smal.

The Southern blot of these putative transformants revealed the difficulties involved in zoosporic fungal transformation. As seen in Figure 2.3, three independent putative transformants all had hybridizing bands of the same size, which is not expected with the random insertion of *Agrobacterium*-mediated transformation. Instead, we would expect that any integrated genetic material to be inserted at a different site for each transformation event, resulting in Southern blots with different hybridizing band sizes. Additionally, these bands were sized within the range of the plasmid, which would not be expected as the plasmids were at least 9.2 kb long, but the DNA between the T-DNA borders is only 2.5 kb, or 3.5 kb in the case of 1568.

#### 2.4 Discussion:

The differences found in developmental stage susceptibility of *Blastocladiella* to selection with hygromycin were not reported previously. Figure 3 of Vieira and Camilo, (2011) show that susceptible sporangia have their development arrested at a common developmental stage, the mature sporangia. This hints at a connection between life stage and selection susceptibility, but was not previously made explicit. This interaction between life stage and susceptibility complicates the transformation procedure as mature sporangia can still release zoospores if sufficiently developed prior to selection, which would obscure the growth of any actual transformed zoospores.

The results from the two Southern blots together indicate that the *Agrobacterium* plasmid was still present in the *Blastocladiella* samples. This indicates that whatever antibiotic resistance we have seen was due to contamination by *Agrobacterium* residing on the fungal sporangia. We could not rely on PCR results to verify that transformation occurred, and Southern blots are required for confirmation. Vieira and Camilo were able to obtain several lines of transformants that could survive sequential transfer, but our transformant lines were not stable over successive transfers, although they did grow

substantially for a few generations. Potentially if we had lines that showed stable antibiotic resistance, Southern blots would show evidence of genomic integration.

The main issues with transformation in *Blastocladiella* involved resistance of mature sporangia to hygromycin and the need for liquid above the agar media. To address these, a more robust transformation protocol for zoosporic fungi should be based on either a different selection agent, or a different method of transformation. To deal with issues around timing of developmental stages, electroporation would give much faster transformation, which may allow selection before the sporangia mature. Alternatively, selection agents with different chemical characteristics may be absorbed equally through rhizoids as through the sporangial wall, eliminating the need for liquid above the solid media. As hygromycin is a protein-synthesis inhibitor, the use of antibiotics with alternative modes of action should be pursued, as they may not have the same absorption or developmental stage issues as hygromycin.

#### 2.5 Summary/Future Directions:

While I was unable to replicate the results of Vieira and Camilo's work, I did determine that I could rule out false positives with Southern blots. Because of the GFP evidence presented by Vieira and Camilo it still seems likely that they obtained transformation. Although their report of growth on selective media combined with PCR results may have been due to natural resistance in *Blastocladiella* combined with *Agrobacterium* contamination, the GFP evidence is compelling. Further work on transformation in zoosporic Fungi should focus on identifying selection agents that do not require free water above the media, which would allow for recovery of discrete colonies. As well, the use of electroporation transformation techniques could limit the time before selection was applied, which would help to alleviate some of the trouble with stages of the lifecycle being resistant to selection.

Reliable transformation in early-diverging Fungi would allow the direct investigation of many genes with cryptic functions. For example, many chytrids do not use tip growth, instead developing an isodiametric sporangium, yet they contain all the core genes used by filamentous fungi even though hyphal growth appeared after divergence of zoosporic Fungi (Dee, Mollicone, Longcore, Roberson, & Berbee, 2015; Richards, Leonard, & Wideman, 2017). A recent genomic survey found orthologs of many *Neurospora* genes involved in tip growth in all lineages of Fungi, including

Chytridiomycota and Blastocladiomycota (See Figure 2 in Richards et al., 2017). The role of these genes in Fungi that lack filamentous growth is perplexing, but recent work from our lab (Dee et al., unpublished) has shown evidence for actin accumulation at the tips of rhizoids similar to hyphal tip growth in filamentous fungi. With reliable transformation techniques, it would be possible to localize these core proteins, and see whether they localize to the sporangia, rhizoids, or both during development. This would allow us to understand the ancestral role of the proteins now involved in filamentous growth.

Light perception is another example of where transformation tools would prove especially powerful. While Ascomycetes, Basidiomycetes, and zygomycetes sense light using the conserved White Collar 1 (WC-1) photoreceptor, a genomic comparison did not recover homologs in the chytrid *Batrachochytrium dendrobatidis*, or in the earlydiverging Microsporidian *Encephalitozoon cuniculi* (Salichos & Rokas, 2010). Recently, instead of the White Collar-1 photoreceptor, Ahrendt et al (2017), found an opsin in Chytridiomycota homologous with those found in Animals, although the phylogenetic analysis is still unpublished. While zoosporic Fungi do not have the homologous lightreceptor of other Fungi, they do share homologs of the rest of the light perception system such as FRQ, FWD-1, and FRH (Salichos & Rokas, 2010). Given that the regulatory pathway appears conserved in all Fungi, but the photoreceptor is instead shared with Animals, I would assume that the effect of a FRQ, FWD-1, or FRH knockout mutant would be similar to the perturbation of the circadian rhythm seen in the wellstudied *Neurospora* system (Montenegro-Montero, Canessa, & Larrondo, 2015).

#### **Conclusion:**

To understand the mechanisms of morphological complexity in Fungi I used two approaches, transformation of a zoosporic fungus, as well as an analysis of a protein family involved in morphogenesis in Fungi. While the transformation was in the end not successful, I did provide data that can be used to assist future progress. In investigating the septin protein family I discovered novel diversity, and identified residues that interact across Opisthokonts. These findings can allow testing of hypotheses related to septin binding, and the role of novel septin group in Fungi.

My research in transformation of *Blastocladiella* showed that Southern blot results are needed to confirm successful genomic integration, which was not shown previously (Vieira & Camilo, 2011). Also, I discovered a connection between developmental stage and response to protein synthesis inhibitors, which was previously reported in zoospores (Soll & Sonneborn, 1971), but not in sporangia in relation to hygromycin (Vieira & Camilo, 2011). The aquatic nature of this organism presents particular difficulties, which may be overcome with testing of selection agents with modes of action that do not involve inhibiting protein synthesis. As well, the use of electroporation transformation methods may avoid some of the challenges with the lifecycle of *Blastocladiella* in relation to AMT. As the zoospores of chytrids are not surrounded by a chitinous cell wall, instead what appears to be a glycoprotein matrix, electroporation may be more simple than in other fungi (Dorward & Powell, 1983).

Through a detailed phylogenetic and bioinformatic study of septin proteins in Opisthokonts I revealed a more complicated history than previously inferred (Cao et al., 2007; Nishihama et al., 2011; Pan et al., 2007). The inclusion of early-diverging lineages uncovered hidden diversity, as well as reducing the phylogenetic certainty. My research provides evidence that the Opisthokont ancestor had multiple septin genes, with at least one member likely without a coiled-coil domain, which seems to be a conserved trait among all Opisthokont lineages, even though the septins without coiledcoils are not homologous between taxa. There is limited evidence for the role of septins without coiled-coil domains, but recent results indicate that they may be involved in homopolymers (Ilkay Dörter, pers comm), which has been previously unknown. This role could explain the phylogenetic trait of having septins lacking a coiled-coil domain, without a homologous family. I also identified a set of conserved interacting residues

across Opisthokonts, which can inform future mutagenesis experiments, which until now have been restricted to testing naïve hypotheses by changing interface residues to alternatives almost guaranteed not to associate.

#### **Bibliography:**

- Åkerborg, Ö., Sennblad, B., Arvestad, L., & Lagergren, J. (2009). Simultaneous Bayesian gene tree reconstruction and reconciliation analysis. *Proceedings of the National Academy of Sciences*, *106*(14), 5714–5719. https://doi.org/10.1073/pnas.0806251106
- Ames, R. M., Talavera, D., Williams, S. G., Robertson, D. L., & Lovell, S. C. (2016). Binding interface change and cryptic variation in the evolution of protein-protein interactions. *BMC Evolutionary Biology*, 16, 40. https://doi.org/10.1186/s12862-016-0608-1
- Angelis, D., & Spiliotis, E. T. (2016). Septin mutations in human cancers. *Frontiers in Cell and Developmental Biology*, 4. https://doi.org/10.3389/fcell.2016.00122
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., ... Bourne, P. E. (2000). The protein data bank. *Nucleic Acids Research*, 28(1), 235–242. https://doi.org/10.1093/nar/28.1.235
- Bertin, A., McMurray, M. A., Grob, P., Park, S.-S., Garcia, G., Patanwala, I., ... Nogales, E. (2008). Saccharomyces cerevisiae septins: Supramolecular organization of heterooligomers and the mechanism of filament assembly. *Proceedings of the National Academy of Sciences*, 105(24), 8274–8279. https://doi.org/10.1073/pnas.0803330105
- Bertin, A., McMurray, M. A., Thai, L., Garcia, G., Votin, V., Grob, P., ... Nogales, E. (2010). Phosphatidylinositol-4,5-bisphosphate promotes budding yeast septin filament assembly and organization. *Journal of Molecular Biology*, 404(4), 711–731. https://doi.org/10.1016/j.jmb.2010.10.002
- Bowman, B. H., Taylor, J. W., Brownlee, A. G., Lee, J., Lu, S. D., & White, T. J. (1992). Molecular evolution of the fungi: relationship of the Basidiomycetes, Ascomycetes, and Chytridiomycetes. *Molecular Biology and Evolution*, 9(2), 285–296. https://doi.org/10.1093/oxfordjournals.molbev.a040720
- Brausemann, A., Gerhardt, S., Schott, A.-K., Einsle, O., Große-Berkenbusch, A., Johnsson, N., & Gronemeyer, T. (2016). Crystal structure of Cdc11, a septin subunit from Saccharomyces cerevisiae. *Journal of Structural Biology*, 193(3), 157–161. https://doi.org/10.1016/j.jsb.2016.01.004
- Bridges, A. A., Zhang, H., Mehta, S. B., Occhipinti, P., Tani, T., & Gladfelter, A. S. (2014). Septin assemblies form by diffusion-driven annealing on membranes. *Proceedings of the National Academy of Sciences*, 111(6), 2146–2151. https://doi.org/10.1073/pnas.1314138111
- Cantino, E. C. (1951). Metabolism and morphogenesis in a new Blastocladiella. *Antonie van Leeuwenhoek*, *17*(1), 325–362. https://doi.org/10.1007/BF02062279
- Cao, L., Ding, X., Yu, W., Yang, X., Shen, S., & Yu, L. (2007). Phylogenetic and evolutionary analysis of the septin protein family in metazoan. *FEBS Letters*, *581*(28), 5526–5532. https://doi.org/10.1016/j.febslet.2007.10.032
- Capra, J. A., & Singh, M. (2007). Predicting functionally important residues from sequence conservation. *Bioinformatics*, *23*(15), 1875–1882. https://doi.org/10.1093/bioinformatics/btm270
- Casamayor, A., & Snyder, M. (2003). Molecular dissection of a yeast septin: distinct domains are required for septin interaction, localization, and function. *Molecular and Cellular Biology*, *23*(8), 2762–2777. https://doi.org/10.1128/MCB.23.8.2762-2777.2003
- Cavalier-Smith, T., Chao, E. E., Snell, E. A., Berney, C., Fiore-Donno, A. M., & Lewis, R. (2014). Multigene eukaryote phylogeny reveals the likely protozoan ancestors of opisthokonts

(animals, fungi, choanozoans) and amoebozoa. *Molecular Phylogenetics and Evolution*, *81*, 71–85. https://doi.org/10.1016/j.ympev.2014.08.012

- Chang, Y., Wang, S., Sekimoto, S., Aerts, A. L., Choi, C., Clum, A., ... Berbee, M. L. (2015). Phylogenomic analyses indicate that early fungi evolved digesting cell walls of algal ancestors of land plants. *Genome Biology and Evolution*, 7(6), 1590–1601. https://doi.org/10.1093/gbe/evv090
- Crooks, G. E., Hon, G., Chandonia, J.-M., & Brenner, S. E. (2004). WebLogo: A Sequence logo generator. *Genome Research*, 14(6), 1188–1190. https://doi.org/10.1101/gr.849004
- Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics*, *27*(8), 1164–1165. https://doi.org/10.1093/bioinformatics/btr088
- Dee, J. M., Mollicone, M., Longcore, J. E., Roberson, R. W., & Berbee, M. L. (2015). Cytology and molecular phylogenetics of Monoblepharidomycetes provide evidence for multiple independent origins of the hyphal habit in the Fungi. *Mycologia*, 107(4), 710–728. https://doi.org/10.3852/14-275
- Delorenzi, M. (2016). MARCOIL. Retrieved from bcf.isb-sib.ch/Delorenzi/Marcoil/index.html
- Dorward, D. W., & Powell, M. J. (1983). Cytochemical Detection of Polysaccharides and the Ultrastructure of the Cell Coat of Zoospores of Chytriomyces aureus and Chytriomyces hyalinus. *Mycologia*, *75*(2), 209–220. https://doi.org/10.2307/3792804
- Eddy, S. R. (2008). A probabilistic model of local sequence alignment that simplifies statistical significance estimation. *PLOS Computational Biology*, *4*(5), e1000069. https://doi.org/10.1371/journal.pcbi.1000069
- Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, 16, 157. https://doi.org/10.1186/s13059-015-0721-2
- Finnigan, G. C., Takagi, J., Cho, C., & Thorner, J. (2015). Comprehensive genetic analysis of paralogous terminal septin subunits Shs1 and Cdc11 in Saccharomyces cerevisiae. *Genetics*, 200(3), 821–841. https://doi.org/10.1534/genetics.115.176495
- Garre, V., Barredo, J. L., & Iturriaga, E. A. (2015). Transformation of Mucor circinelloides f. Iusitanicus protoplasts. In M. A. van den Berg & K. Maruthachalam (Eds.), *Genetic Transformation Systems in Fungi, Volume 1* (pp. 49–59). Springer International Publishing. https://doi.org/10.1007/978-3-319-10142-2\_4
- Gladfelter, A. S. (2010). Guides to the final frontier of the cytoskeleton: septins in filamentous fungi. *Current Opinion in Microbiology*, *13*(6), 720–726. https://doi.org/10.1016/j.mib.2010.09.012
- Gutiérrez, A., López-García, S., & Garre, V. (2011). High reliability transformation of the basal fungus Mucor circinelloides by electroporation. *Journal of Microbiological Methods*, 84(3), 442–446. https://doi.org/10.1016/j.mimet.2011.01.002
- Hamann, E., Gruber-Vodicka, H., Kleiner, M., Tegetmeyer, H. E., Riedel, D., Littmann, S., ... Strous, M. (2016). Environmental Breviatea harbour mutualistic Arcobacter epibionts. *Nature*, 534(7606), 254–258. https://doi.org/10.1038/nature18297
- Hernández-Rodríguez, Y., Hastings, S., & Momany, M. (2012). The Septin AspB in Aspergillus nidulans forms bars and filaments and plays roles in growth emergence and conidiation. *Eukaryotic Cell*, *11*(3), 311–323. https://doi.org/10.1128/EC.05164-11
- Hernández-Rodríguez, Y., Masuo, S., Johnson, D., Orlando, R., Smith, A., Couto-Rodriguez, M., & Momany, M. (2014). Distinct septin heteropolymers co-exist during multicellular

development in the filamentous fungus Aspergillus nidulans. *PLOS ONE, 9*(3), e92819. https://doi.org/10.1371/journal.pone.0092819

- Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics (Oxford, England)*, *17*(8), 754–755.
- James, T. Y., Letcher, P. M., Longcore, J. E., Mozley-Standridge, S. E., Porter, D., Powell, M. J., ... Vilgalys, R. (2006). A molecular phylogeny of the flagellated fungi (Chytridiomycota) and description of a new phylum (Blastocladiomycota). *Mycologia*, 98(6), 860–871. https://doi.org/10.3852/mycologia.98.6.860
- Kozubowski, L., & Heitman, J. (2010). Septins enforce morphogenetic events during sexual reproduction and contribute to virulence of Cryptococcus neoformans. *Molecular Microbiology*, *75*(3), 658–675. https://doi.org/10.1111/j.1365-2958.2009.06983.x
- Kück, P., & Longo, G. C. (2014). FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Frontiers in Zoology*, 11. https://doi.org/10.1186/s12983-014-0081-x
- Kück, P., Meusemann, K., Dambach, J., Thormann, B., von Reumont, B. M., Wägele, J. W., & Misof, B. (2010). Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. *Frontiers in Zoology*, 7, 10. https://doi.org/10.1186/1742-9994-7-10
- Maddison, W., & Maddison, D. (2016). *Mesquite: a modular system for evolutionary analysis. Version 3.10. 2016.*
- Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., ... Bryant, S. H. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Research*, 43(D1), D222-226. https://doi.org/10.1093/nar/gku1221
- Meseroll, R. A., Howard, L., & Gladfelter, A. S. (2012). Septin ring size scaling and dynamics require the coiled-coil region of Shs1p. *Molecular Biology of the Cell*, 23(17), 3391–3406. https://doi.org/10.1091/mbc.E12-03-0207
- Michielse, C. B., Salim, K., Ragas, P., Ram, A. F. J., Kudla, B., Jarry, B., ... Hondel, C. A. M. J. J. van den. (2004). Development of a system for integrative and stable transformation of the zygomycete Rhizopus oryzae by Agrobacterium-mediated DNA transfer. *Molecular Genetics and Genomics*, 271(4), 499–510. https://doi.org/10.1007/s00438-004-1003-y
- Miller, M., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES science gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop* (pp. 1–8). New Orleans, LA. https://doi.org/10.1109/GCE.2010.5676129
- Misof, B., & Misof, K. (2009). A monte carlo approach successfully identifies randomness in multiple sequence alignments: A more objective means of data exclusion. *Systematic Biology*, 58(1), 21–34. https://doi.org/10.1093/sysbio/syp006
- Momany, M., & Talbot, N. J. (2017). Septins Focus Cellular Growth for Host Infection by Pathogenic Fungi. *Frontiers in Cell and Developmental Biology*, *5*, 33. https://doi.org/10.3389/fcell.2017.00033
- Montenegro-Montero, A., Canessa, P., & Larrondo, L. F. (2015). Around the fungal clock: Recent advances in the molecular study of circadian clocks in Neurospora and other fungi. *Advances in Genetics*, *92*, 107–184. https://doi.org/10.1016/bs.adgen.2015.09.003
- Mostowy, S., & Cossart, P. (2012). Septins: the fourth component of the cytoskeleton. *Nature Reviews Molecular Cell Biology*, *13*(3), 183–194. https://doi.org/10.1038/nrm3284

- Nishihama, R., Onishi, M., & Pringle, J. R. (2011). New insights into the phylogenetic distribution and evolutionary origins of the septins. *Biological Chemistry*, *392*(8–9), 681–687. https://doi.org/10.1515/BC.2011.086
- Nyilasi, I., Papp, T., Csernetics, Á., & Vágvölgyi, C. (2008). Agrobacterium tumefaciens-mediated transformation of the zygomycete fungus Backusella lamprospora. *Journal of Basic Microbiology*, 48(1), 59–64. https://doi.org/10.1002/jobm.200700221
- Oh, Y., & Bi, E. (2011). Septin structure and function in yeast and beyond. *Trends in Cell Biology*, 21(3), 141–148. https://doi.org/10.1016/j.tcb.2010.11.006
- Pan, F., Malmberg, R. L., & Momany, M. (2007). Analysis of septins across kingdoms reveals orthology and new motifs. *BMC Evolutionary Biology*, 7, 103. https://doi.org/10.1186/1471-2148-7-103
- Papadopoulos, J. S., & Agarwala, R. (2007). COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics*, 23(9), 1073–1079. https://doi.org/10.1093/bioinformatics/btm076
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics (Oxford, England), 20*(2), 289–290. https://doi.org/10.1093/bioinformatics/btg412
- Parfrey, L. W., Lahr, D. J. G., Knoll, A. H., & Katz, L. A. (2011). Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proceedings of the National Academy of Sciences*, 108(33), 13624–13629. https://doi.org/10.1073/pnas.1110633108
- Pattengale, N. D., Alipour, M., Bininda-Emonds, O. R. P., Moret, B. M. E., & Stamatakis, A. (2009). How many bootstrap replicates are necessary? In *Research in Computational Molecular Biology* (pp. 184–200). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-02008-7\_13
- Rambaut, A., Suchard, M., Xie, D., & Drummond, A. (2014). Tracer (Version 1.6). Retrieved from http://beast.bio.ed.ac.uk/Tracer
- Revuelta, J. L., & Jayaram, M. (1986). Transformation of Phycomyces blakesleeanus to G-418 resistance by an autonomously replicating plasmid. *Proceedings of the National Academy of Sciences*, *83*(19), 7344–7347.
- Richards, T. A., Leonard, G., & Wideman, J. G. (2017). What defines the "Kingdom" Fungi? *Microbiology Spectrum*, 5(3). https://doi.org/10.1128/microbiolspec.FUNK-0044-2017
- Ruiz-Trillo, I., Burger, G., Holland, P. W. H., King, N., Lang, B. F., Roger, A. J., & Gray, M. W. (2007). The origins of multicellularity: a multi-taxon genome initiative. *Trends in Genetics*, *23*(3), 113–118. https://doi.org/10.1016/j.tig.2007.01.005
- Sali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 234(3), 779–815. https://doi.org/10.1006/jmbi.1993.1626
- Salichos, L., & Rokas, A. (2010). The diversity and evolution of circadian clock proteins in fungi. *Mycologia*, 102(2), 269–278.
- Sanderson, M. J. (2003). r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics (Oxford, England)*, 19(2), 301– 302. https://doi.org/10.1093/bioinformatics/19.2.301
- Sellin, M. E., Sandblad, L., Stenmark, S., & Gullberg, M. (2011). Deciphering the rules governing assembly order of mammalian septin complexes. *Molecular Biology of the Cell*, 22(17), 3152–3164. https://doi.org/10.1091/mbc.E11-03-0253

 Serrão, V. H. B., Alessandro, F., Caldas, V. E. A., Marçal, R. L., D'Muniz Pereira, H., Thiemann, O. H., & Garratt, R. C. (2011). Promiscuous interactions of human septins: The GTP binding domain of SEPT7 forms filaments within the crystal. *FEBS Letters*, *585*(24), 3868–3873. https://doi.org/10.1016/j.febslet.2011.10.043

Shimodaira, H., & Hasegawa, M. (2001). CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics*, *17*(12), 1246–1247. https://doi.org/10.1093/bioinformatics/17.12.1246

Sirajuddin, M., Farkasovsky, M., Hauer, F., Kühlmann, D., Macara, I. G., Weyand, M., ... Wittinghofer, A. (2007). Structural insight into filament formation by mammalian septins. *Nature*, 449(7160), 311–315. https://doi.org/10.1038/nature06052

Sirajuddin, M., Farkasovsky, M., Zent, E., & Wittinghofer, A. (2009). GTP-induced conformational changes in septins and implications for function. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(39), 16592–16597. https://doi.org/10.1073/pnas.0902858106

Sjöstrand, J., Sennblad, B., Arvestad, L., & Lagergren, J. (2012). DLRS: gene tree evolution in light of a species tree. *Bioinformatics*, *28*(22), 2994–2995. https://doi.org/10.1093/bioinformatics/bts548

Soll, D. R., & Sonneborn, D. R. (1971). Zoospore Germination in Blastocladiella emersonii: Cell Differentiation without Protein Synthesis? *Proceedings of the National Academy of Sciences*, *68*(2), 459–463.

Spatafora, J. W., Aime, M. C., Grigoriev, I. V., Martin, F., Stajich, J. E., & Blackwell, M. (2017). The Fungal Tree of Life: from Molecular Systematics to Genome-Scale Phylogenies. *Microbiology Spectrum*, 5(5). https://doi.org/10.1128/microbiolspec.FUNK-0053-2016

Spatafora, J. W., Chang, Y., Benny, G. L., Lazarus, K., Smith, M. E., Berbee, M. L., ... Stajich, J. E. (2016). A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia*, *108*(5), 1028–1046. https://doi.org/10.3852/16-042

Spiliotis, E. T., & Gladfelter, A. S. (2012). Spatial Guidance of Cell Asymmetry: Septin GTPases Show the Way. *Traffic*, *13*(2), 195–203. https://doi.org/10.1111/j.1600-0854.2011.01268.x

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Sukumaran, J., & Holder, M. T. (2010). DendroPy: a Python library for phylogenetic computing. *Bioinformatics*, *26*(12), 1569–1571. https://doi.org/10.1093/bioinformatics/btq228

Takizawa, P. A., DeRisi, J. L., Wilhelm, J. E., & Vale, R. D. (2000). Plasma membrane compartmentalization in yeast by messenger RNA transport and a septin diffusion barrier. *Science*, 290(5490), 341–344. https://doi.org/10.1126/science.290.5490.341

The PyMOL Molecular Graphics System. (n.d.). PyMOL (Version 1.8). Schrödinger, LLC.

- Torruella, G., de Mendoza, A., Grau-Bové, X., Antó, M., Chaplin, M. A., del Campo, J., ... Ruiz-Trillo, I. (2015). Phylogenomics reveals convergent evolution of lifestyles in close relatives of animals and fungi. *Current Biology*, 25(18), 2404–2410. https://doi.org/10.1016/j.cub.2015.07.053
- Vernot, B., Stolzer, M., Goldman, A., & Durand, D. (2008). Reconciliation with non-binary species trees. *Journal of Computational Biology*, *15*(8), 981–1006. https://doi.org/10.1089/cmb.2008.0092

- Versele, M., Gullbrand, B., Shulewitz, M. J., Cid, V. J., Bahmanyar, S., Chen, R. E., ... Thorner, J. (2004). Protein–protein interactions governing septin heteropentamer assembly and septin filament organization in Saccharomyces cerevisiae. *Molecular Biology of the Cell*, 15(10), 4568–4583. https://doi.org/10.1091/mbc.E04-04-0330
- Vieira, A. L. G., & Camilo, C. M. (2011). Agrobacterium tumefasciens-mediated transformation of the aquatic fungus Blastocladiella emersonii. *Fungal Genetics and Biology*, 48(8), 806– 811. https://doi.org/10.1016/j.fgb.2011.02.006
- Vogel, H. J. (1961). Lysine Synthesis and Phytogeny of Lower Fungi : Some Chytrids versus Hyphochytrium. *Nature*, *189*(4769), 1026–1027. https://doi.org/10.1038/1891026a0
- Walker, W. F., & Doolittle, W. F. (1982). Nucleotide sequences of 5S ribosomal RNA from four oomycete and chytrid water molds. *Nucleic Acids Research*, 10(18), 5717–5721. https://doi.org/10.1093/nar/10.18.5717
- Wang, Y., DiGuistini, S., Wang, T.-C. T., Bohlmann, J., & Breuil, C. (2010). Agrobacteriummeditated gene disruption using split-marker in Grosmannia clavigera, a mountain pine beetle associated pathogen. *Current Genetics*, 56(3), 297–307. https://doi.org/10.1007/s00294-010-0294-2
- Warren, D. L., Geneva, A. J., & Lanfear, R. (2017). RWTY (R We There Yet): An R package for examining convergence of bayesian phylogenetic analyses. *Molecular Biology and Evolution*, 34(4), 1016–1020. https://doi.org/10.1093/molbev/msw279
- Weirich, C. S., Erzberger, J. P., & Barral, Y. (2008). The septin family of GTPases: architecture and dynamics. *Nature Reviews Molecular Cell Biology*, *9*(6), 478–489. https://doi.org/10.1038/nrm2407
- Word, J. M., Lovell, S. C., LaBean, T. H., Taylor, H. C., Zalis, M. E., Presley, B. K., ... Richardson, D. C. (1999). Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. *Journal of Molecular Biology*, 285(4), 1711–1733. https://doi.org/10.1006/jmbi.1998.2400
- Word, J. M., Lovell, S. C., Richardson, J. S., & Richardson, D. C. (1999). Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *Journal of Molecular Biology*, 285(4), 1735–1747. https://doi.org/10.1006/jmbi.1998.2401
- Wöstemeyer, J., Burmester, A., & Weigel, C. (1987). Neomycin resistance as a dominantly selectable marker for transformation of the zygomycete Absidia glauca. *Current Genetics*, *12*(8), 625–627. https://doi.org/10.1007/BF00368066
- Yelton, M. M., Hamer, J. E., Souza, E. R. de, Mullaney, E. J., & Timberlake, W. E. (1983). Developmental regulation of the Aspergillus nidulans trpC gene. *Proceedings of the National Academy of Sciences*, 80(24), 7576–7580.
- Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T.-Y. (2017). ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, 8(1), 28–36. https://doi.org/10.1111/2041-210X.12628

#### **Appendix A: Supplemental Figures for Chapter 1**

Figure A.1: Maximum likelihood phylogenetic analysis with RAxML software. Topology is the most likely tree from 2000 independent searches, node values represent bootstrap support. Protein names are shown for those taxa with experimental evidence. *Aspergillus* and *Drosophila* sequences used to recognize septin groups are in bold. Coiled-coil domain predictions, black representing p<0.05 and grey p<0.10, found to the right of names. Domain predictions for proteins longer than 600 residues have been shortened with diagonal lines.



Figure A.2: Bayesian phylogenetic analysis with MrBayes software. Topology represents consensus tree after 200 million generations, node values represent posterior probability. Protein names are shown for those septins with experimental evidence. *Aspergillus* and *Drosophila* sequences used to recognize septin groups are in bold. Coiled-coil domain prediction, black representing p<0.05 and grey p<0.10, shown to the right of names. Coiled-coil predictions for proteins longer than 600 residues have been shortened with diagonal lines.



Figure A.3: Results of Bayesian phylogenetic analysis with jPRIME software. Topology represents maximum clade credibility tree, node values represent bootstrap support. Protein names are shown for those taxa with experimental evidence. *Aspergillus* and *Drosophila* sequences used to recognize septin groups are in bold. Coiled-coil domain prediction, black representing p<0.05 and grey p<0.10 to the right of septin names. Coiled-coil predictions for proteins longer than 600 residues have been shortened with diagonal lines.



Figure A.4: Ancestral reconstructions for presence of septin groups inferred with Mesquite with the MK1 symmetrical model. Shading of pie charts at nodes represent proportional likelihood of a node containing a member of that septin group. Statistical test supporting use of MK1 found at below state reconstructions.



# Figure A.5: COBALT alignment of representative septins from the 7 groups, showing location of conserved regions and interface regions

1A CDC10 NP 009928 1	[ 14] EDTITNOTEHBLIKKGEOENIN	AVEOSELGKSTLINTLEASHLIDSA	TGDDI-SALPVTK
1B Sep5 NP 651961.1	[ 28] FDTLPDOLVNKSVONGESENTI	CIGETALGKSTLMDTLENTSFGST -	PSPHNLP
2A NP 013418.2 CDC3	[101] FANLPKOWHRRSIKNGFSFNLI	LCVGPDGIGKTTLMKTLFNNDDIEAN	LVKDYEEELANDOEEEEGOGEGHENOSOEORH
2B Sep4 NP 573147.2	[ 67] FATLPEQVHRKSVKRGFEFTLM	NVVGESGLGKSTLINSLFLGDLYKNR	QMPNVEERIEK
3 CDC11 NP 012610.1	[ 4] IDAS <mark>SALRKRKH</mark> LKRGITFTVN	11VGQS <mark>G</mark> SGRSTFINTLCGQQVVDTS	TTILLPTDTSTEI
4 CDC12 NP 011975.1	[ 16] ISNLPNQRYKIVNEEGGTFTVN	1LCGES <mark>G</mark> LGKTTFINTLFQTVLKRA-	DGQQH-RQEPIRK
5_AspE_XP_662271.1	[ 4] YNPI <mark>K-LRR</mark> KKNVKKGIQFCLM	1VC <mark>GAS<mark>G</mark>TGRTT</mark> FVNTLCGKQVLQGK	DADDA-TNAHLEE
	_	G1	
	h2 h3	a2	
1 CDC10 ND 009928 1			
18 Sep5 NP 651961 1	NVKLKANTYELOESNVELKLTVCOT	GYGDOVNK-ADSYKALVEYV-DSOF	YEAVLOEELKIORAMASAHDORVHACLYFICPT
2A NP 013418 2 CDC3	KVKIKSYESVIE ENGVKININVIDT	EGEGDELNNDOKSWDPLIKEI-DSBE	DOYLDAENKINGHSINDKRIHACLYFIEPT
2B Sep4 NP 573147.2	TTKVEKKTMDIE ERGVRLRLTVVDTI	PGFGDAINC-EDSWRVCTOYI-DEOF	BOY FTDE SGLNB BNTODNRVHCCLYFVPPW
3 CDC11 NP 012610.1	DLOLREETVELE-DDEGVKIOLNIIDTI	PGFGDSLDN-SPSFEIISDYI-RHOY	DEILLEESRVRRN-PRFKDGRVHCCLYLINPT
4 CDC12 NP 011975.1	TVEIDITRALLE EKHFELRVNVIDT	PGFGDNVNN-NKAWQPLVDFI-DDQH	IDSYMRQEQQPYRTKKFDLRVHAVLYFIRPT
5 AspE XP 662271.1	GVRIKPVTVELELDDEGTRISLTIVDT	PGF <mark>GDQIDN</mark> -EARHVVTVNKIWERQY	DILAEESRIKRN-PRFRDNRVHVLLYFITPT
	G	3	<u>\$1</u> <u>\$2</u>
	a3 () b5	<b>a4</b> ()	<u>a5</u> (
1A CDC10 NP 009928 1	GKELSRLDVEALKRITETANVTPVTCK	DTLTLDERTEFRELIONEFEKYNFK	IYPYDSEELTDEEI.ELNRSVR
LB Sep5 NP 651961.1	GHGLKAMDLVCMKOLDTRVNIIPVTAK	ADTISKSELSGFKERIMDELRRNNVS	IYOFPMDDETVSETNAAMN
2A NP 013418.2 CDC3	GHYLKPLDLKFMQSVYEKCNLIPVIAK	DILTDEEILSFKKTIMNOLIOSNIE	LFKPPIYSNDDAENSHLSERLF
2B Sep4 NP 573147.2	GHSLROMDLDLIRRLHRKVNIVLVIGK	ADCLNKQEVRKLKERILQDLEDNHIQ	LYQFP-ECDSDEDDDFKQQDRELK
3 CDC11 NP 012610.1	GHGLKEIDVEFIRQLGSLVNIIPVISK	DSLTRDELKLNKKLIMEDIDRWNLP	IYNFPFDEDEISDEDYETNMYLR
1_CDC12_NP_011975.1	GHGLKPIDIETMKRLSTRANLIPVIAK	DTLTAQELQQFKSRIRQVIEAQEIR	IFTPPLDADSKEDAKSGSNPDSAAVEHARQLI
5_AspE_XP_662271.1	GHGLRELDIELMKRLSPRVNVIPVIGK/	DSLTPAELAESKKLIMEDIEHYRIP	VYNFPYDIEEDDEDTVEENAELR
	S3 G4		
		N	
	() <u>b6</u> <u>b9</u> <u>b10</u> <u>b7</u>		a7 ()
1A_CDC10_NP_009928.1	SIIPFAVVGSENEIEIN-GETFRGR	KTRWSAINVEDINQCDFVYLREFLIR	THLQDLIETTSYIHYEGFRARQLIAL
1B_Sep5_NP_651961.1	GHLPFAVVGSTEFVK-VAGKQVRA	QYP <mark>WGAV</mark> HIENEA <mark>H</mark> CDFVKLREMLI <mark>R</mark>	TNMEDLREQTHTRHYELFRQRRLQQM
2A_NP_013418.2 CDC3	SSLPYAVIGSNDIVENYSGNQVRGR:	SYP <mark>WGVI</mark> EVDNDN <mark>H</mark> SDFNLLKNLLI <mark>K</mark>	QFMEELKERTSKILYENYRSSKLAKL
2B_Sep4_NP_573147.2	ASIPFAVVGSNTILEVA-GKKVRGRG	2YP <mark>WGVV</mark> NVEDPE <mark>H</mark> SDFIKLRTFLI <mark>S</mark>	THMQDLKDTTQDVHYENFRAQCISQISQHALR
3_CDC11_NP_012610.1	TLLPFAIIGSNEVYEMG-GDVGTIRG <mark>R</mark> I	KYPWGILDVEDSS <mark>I</mark> SDFVILRNALL <mark>I</mark>	SHLHDLKNYTHEILYERYRTEALSGESVAAES
4_CDC12_NP_011975.1	EAMPFAIVGSEKKFDNGQGTQVVAR	KYPWGLVEIENDSHCDFRKLRALLL <mark>R</mark>	TYLLDLISTTQEMHYETYRRLRLEGH
5_AspE_XP_662271.1	GLMPFAIVGSDDFVEID-GRKVRAR	2YPWGVVEVENPR <mark>H</mark> SDFLAIRSALL <mark>H</mark>	SHLADLKEITHDFLYENYRTEKLSK
		34	
A_CDC10_NP_009928.1		KENANSRSS	AHMSSNAIQR
.B_Sep5_NP_651961.1	GFVDVDSNNQPVSF	QQTFESKRS	DHLACLQAKEEE
A_NP_013418.2 CDC3	GIKQDNSVFKEFDPISK	QQEEKI	LHEAKLAKLEIE
B_Sep4_NP_573147.2	ERGKLKRDSISSTNGFDAAI	SETDRI	LLQKDEEIRRMQDM
_CDC11_NP_012610.1	IRPNLTKLNGSSSSSTTTRRNTNPFKQS	SNNINNDVLNPASDMHGQSTGENNET	YMTREEQIRLR
_CDC12_NP_011975.1		ENTGEGNED	OFTLPAIAPARKLSHNPRYKEEENALKKYFTDQ
_AspE_XP_662271.1	SVDGAATTGTDSS	INP-EDLATQSV	-RLKEEQLRRK
	Π		
	W		
A CDC10 NP 009928.1			
LB Sep5 NP 651961.1	VRQMFVQRVKQKENELKDNEKELH	TKFDRLKREHLEEKAQLEEARRQLEE	DCQELQRRRLQMANGSHTLTLGRGKKK
2A_NP_013418.2 CDC3	MKTVFQQKVSEKEKKLQKSETELF	ARHKEMKEKLTKQLKALEDKKKQLEL	SINSASPNVNHSPVPTKKKGFLR
2B_Sep4_NP_573147.2	LTQMQEKLKQTHLMEMKKNDSVID	7	
3_CDC11_NP_012610.1	LKAFEERVQQELLLKRQELLQREKELR	IEARLEKEAKIKQEE	
4_CDC12_NP_011975.1	VKAEEQRFRQWEQNIVNERIRLNGDLEE	LIQGKVKKLEEQVKSLQVKKSHLK	
5_AspE_XP_662271.1	LREIELKVQREIAEKRQELLARESQLRE	LIEARMAREASQGEAAEAEQ	
		polybasic domain	Helix A-Sheet
		G domain	G interface NC interface
		SUE element	

Table A.1: Sources for proteomes used in this study.

Organism	Genome Version Info	Source
Allomyces macrogynus	GCA_000151295.1 / A_macrogynus_V3	NCBI
Aspergillus nidulans	GCF_000149205.1 / ASM14920V1	NCBI
Basidiobolus meristosporus	CBS 931.73 v1.0	JGI
Batrachochytrium dendrobatidis	GCA_000203795.1 / v1.0	NCBI
Catenaria anguillulae	PL171 v2.0	JGI
Capsaspora owczarzaki	GCA_000151315.2 / C_owczarzaki_V2	NCBI
Conidiobolus coronatus	GCA_001566745.1	NCBI
Creolimax fragrantissima	Accessed Sep 26, 2016 DOI: 10.6084/m9.figshare.1403592	FigShare
Cryptococcus neoformans	GCA_000149245.3 / CNA3	NCBI
Drosophila melanogaster	GCA_000001215.4 / Release 6 plus ISO1 MT	NCBI
Encephalitozoon cuniculi	GCA_000091225.1 / ASM9122V1	NCBI
Fonticula alibcans	GCA_000388065.2 / Font_alba_ATCC_28817_V2	NCBI
Gonapodya prolifera	GCA_001574975.1 / Ganpr1	NCBI
Monosiga brevicollis	GCA_000002865.1 / V1.0	NCBI
Paramicrosporidium saccamaboe	draft	Dr. Alicia Quandt
Phycomyces blakesleeanus	GCA_001638985.2 / Phybl2	NCBI
Rhizophagus irregularis	GCA_001593125.1 / ASM159312v1	NCBI
Saccharomyces cerevisiae	GCA_000146045.2 / R64	NCBI
Salpingoeca rosetta	GCA_000188695.1 / Proterospongia_sp_ATCC50818	NCBI
Schizophyllum commune	GCA_000143185.1 / V1.0	NCBI
Sphaeroforma arctica	GCA_001186125.1 / Spha_artica_JP610_V1	NCBI
Trichoplax adhaerens	GCA_000150275.1 / v1.0	NCBI

Table A.2: Sequences used in homology modelling. Protein codes represent GenBank accession numbers except for *B. meritosporus* and *C. coronatus* which are not listed on GenBank and instead use Joint Genome Institute protein IDs.

Taxonomic Group	Таха	Group 1 Septin	Group 2 Septin
Dikarya	Aspergillus nidulans	XP_664292.1	XP_658998.1
	Cryptococcus neoformans	XP_012050631.1	XP_012049501.1
	Saccharomyces cerevisae	NP_013418.2	NP_009928.1
	Schizophyllum commune	XP_003037204.1	XP_003035685.1
Early-	Allomyces macrogynus	KNE60169.1	KNE63007.1
aiverging Fungi	Basidiobolus meritosporus	313179	315462
	Batrachochtytrium dendrobatidis	XP_006675869.1	XP_006678596.1
	Catenaria anguillulae	1173908	1512492
	Conidiobolus coronatus	KXN68162.1	KXN74403.1
	Gonapodya prolifera	KXS17957.1	KXS10614.1
	Phycomyces blakesleeanus	OAD67191.1	OAD79262.1
Animals	Drosophila melanogaster	NP_523430.1	NP_524417.1
	Caenorhabditis elegans	NP_001041198	NP_493388
	Homo sapiens	XP_011529619	XP_016859678
	Trichoplax adherans	XP_002112086.1	XP_002112436.1
Early-	Monosiga brevicola	XP_001742566.1	XP_001745238.1
aiverging Metazoa	Capsaspora owczarski	XP_004349897.2	XP_004346425.2

## Appendix B: Agrobacterium-Mediated Transformation Protocol

Part 1: Overlap PCR Conditions

- 1. Extract ~3 kb 5' and 3' flanking sequence from the genome
- 2. Design primers to amplify flanking sequence and hph sequence separately.
- 3. Use ~20 bp overlap, try and select sequences with minimum hairpins

Table B.1: First Round PCR Fragment Amplification

Components	Volume/25 µL	PCR conditions:
Water	14 µL	98 <sup>o</sup> C for 2 min
5x Phusion HF buffer	5 µL	
2.5 mM dNTPs	2 µL	30 cycles of:
Primers FW (10 um)	1 µL	98 °C 15sec,68 °C 30 sec,72 °C 1:30 min
Primers RV (10 um)	1 µL	
Template	1 µL	72 °C: 10 min extension
DMSO	0.7 µL	
Phusion	0.3 µL	4°C hold

If gel of first step shows strong single band, then cleanup with standard PCR cleanup kit. Can also gel extract if desired.

Table B.2. Bebena Roana Ovenap i Orc Extension				
Components	Volume/25 µL	PCR conditions		
Water	12 µL	98 <sup>o</sup> C for 2 min		
5x Phusion HF buffer	5 µL			
2.5 mM dNTPs	2 µL	30 cycles of:		
5' flank	1 µL	98 <sup>o</sup> C 15sec,68 <sup>o</sup> C 30 sec,72 <sup>o</sup> C 1:30 min		
3' flank	1 µL			
Hph fragment	3 µL	72 °C: 10 min extension		
DMSO	0.7 µL			
Phusion	0.3 µL	4°C hold		

#### Table B.2: Second Round Overlap PCR Extension

#### Table B.3: Final Product Amplification With Restriction Sites on Primers

Components	Volume/25 µL	PCR conditions
Water	14 µL	
5x Phusion HF buffer	5 µL	98 °C for 2 min
2.5 mM dNTPs	2 µL	30 cycles of:
5' flank nest (or FW)	1 µL	98 °C 15sec68 °C 30 sec. 72 °C 3:00 min
3' flank nest (or RV)	1 µL	
DMSO	0.7 µL	72 °C: 10 min extension
Phusion	0.3 µL	
Template	1 µL	4 <sup>o</sup> C hold

Part 2: Plasmid construction.

Clone PCR-amplified insert into pCambia0380 using standard protocols. Amplify plamid by transformation in *E. coli* DH5 $\alpha$  competent cell. Sequence plasmid after transformation into *E. coli*.

Part 3:

Growth and storage of Agrobacterium tumefaciens

- Strain GV3101 is resistant to gentamicin and rifampicin
- 25-50 µg/mL Gentamycin in media
- 10 µg/mL Rifampicin in media
- Grow at 28-30°C
- Store as -80°C glycerol stock (0.8 mL of overnight culture + 0.2 mL sterile 80% glycerol)

## Freeze/thaw shock transformation of *A. tumefaciens* with plasmid DNA (binary vector system)

1. Pick a single colony of GV3101 from Gentamicin + Rifampicin plate and inoculate into 3 ml of LB (with 3  $\mu$ L of Gentamycin 1000x stock). Grow @ 28°C overnight at 200rpm.

2. Inoculate 50 mL of LB (no antibiotics) in a 250 mL flask with 1 mL of the overnight culture and grow @ 28°C until mid-log phase (OD<sub>600</sub> is between 0.5 and 1.0), approximately 5-6hrs.

3. Chill culture 5-10 minutes on ice. Carefully transfer 1.5 mL culture to chilled, sterile 1.7 mL microfuge tubes and centrifuge @ 10 x g for 1 minutes @ 4°C.

4. Discard supernatant, drain inverted for 30-60 seconds, and resuspend pellet in 100  $\mu$ L of ice cold 20 mM CaCl<sub>2</sub>. By gentle stirring. Be careful not to touch the bottom of tube with your hands. Let sit 5 mins on ice.

5. Add 7  $\mu$ L of 150 ng/ $\mu$ L plasmid to each tube, very gently mix and leave on ice. Freeze tubes in liquid N2, then thaw tubes for 5 min. @ 37 °C.

6. Add 1 mL of LB to each tube, incubate for ~2 hours with shaking at 200 rpm @ 28°C.

7. Spin tubes. Remove supernatant and resuspend pellet in 200 uL of LB.

8. Plate all of the suspension on appropriate antibiotic-LB (50  $\mu$ g/mL Kanamycin, Gentamicin) plates and incubate for two days @ 28 °C. Transformed colonies should be visible on the third day of incubation.

9. Streak one or two colony on Ka/Gm/Rif plate for storage, then store at -80 °C.

## Part 3 - Transform Blastocladiella with Agrobacterium

- 4-8 days before transformation: inoculate plates of PYG media with zoopores to synchronize culture.
- 2-4 days before transformation: Harvest zoopores, and inoculate onto fresh PYG to synchronize cultures
- 2-4 days before transformation: streak *Agrobacterium* on Ka/Gm/Rif plate, incubate @28 °C.

Transformation procedure:

Day 1.

Make Minimal Media, aliquot 20 mL into 100 mL flask, inoculate single colony of *Agrobacterium* into each flask, incubate with shaking 200 rpm @ 28-30 °C.

<u>Minimal Medium (MM)</u> – can be stored in fridge for several months without antibiotics Add the followings to 946 ml of sterilized water

10 mL K-buffer (pH 7.0): 200 g/L K<sub>2</sub>HPO<sub>4</sub> 145 g/L KH<sub>2</sub>PO<sub>4</sub> 20 mL M-N buffer 30 g/L MgSO<sub>4</sub>.7H<sub>2</sub>O 15 g/L NaCl 1 mL 1% CaCl<sub>2</sub>.2H<sub>2</sub>O (w/v) 10 mL 20% glucose (w/v) 2 mL 0.05% FeSO<sub>4</sub> (w/v) 5 mL Spore Elements 100 mg/L ZnSO<sub>4</sub>.7H<sub>2</sub>O 100 mg/L CuSO<sub>4</sub>.5H<sub>2</sub>O 100 mg/L H<sub>3</sub>BO<sub>3</sub> 100 mg/L MnSO<sub>4</sub>.H<sub>2</sub>O 100 mg/L Na<sub>2</sub>MoO<sub>4</sub>.2H<sub>2</sub>O 2.5 mL 20 % NH<sub>4</sub>NO<sub>3</sub> (w/v) 1mL Ka and Gm stock solution (add to 20 µL to 20 mL flask aliquots just before shaking).

\*Autoclave K-buffer, M-N buffer ,  $CaCl_2 \bullet 2H_2O$  , all the others are filter sterilized.

Day 3.

Check the OD of *Agrobacterium*. As long as OD = 0.8-1.0, you can go on. Spin down in refrigerated centrifuge for 20 min @ 4 °C @ 4000 rpm.

While you are waiting, add 200  $\mu$ M acetosyringone to IM for the final incubation (40 mg/100 mL, dissolve in 1 mL of DMSO before adding to IM). Resuspend cells in equal volume of induction medium. Dilute to OD ~ 0.2. Shake for 6h @ 200 rpm @ 28-30 °C.

Harvest *Blastocladiella* zoospores, adjust concentration to 1x10<sup>7</sup> zoospores/mL, store on ice while *Agrobacterium* matures.

Pipette 5 mL of induced *Agrobacterium* culture on the plate. Induction Medium (IM) Add the followings to 904 mL of sterilized water: 10 mL K-buffer (pH 7.0): 20 mL M-N solution 1 mL 1% CaCl<sub>2</sub>.2H<sub>2</sub>O 2 mL 0.01% FeSO<sub>4</sub> (w/v) 5 mL Spore Elements 2.5 mL 20% NH<sub>4</sub>NO<sub>3</sub> (w/v)

10 mL 50% glycerol

40 mL 1 M MES (2-[N-Morpholino]ethanesulfonic acid), pH 5.3, use NaOH to adjust pH 5 mL 2 M glucose

2 mL 100 mM acetosyringone (3',5'-Dimethoxy-3'-hydroxyacetophenone), stored at – 20°C

\*Add 13 g/L agar for solid medium.

\*Prepare the acetosyringone stock solution with ethanol.

Next day.

Scrape sporangia from IM plates, and transfer to selective media plates.

Selective Medium for *Blatocladiella*.

PYG supplemented:

-with hygromycin B at 150  $\mu$ g/mL for selection of *Blastocladiella* transformants -and with cefotaxime 200  $\mu$ M + moxalactam 100  $\mu$ g/mL to kill off *Agrobacterium* 

Table B.4: *Blastocladiella* Primers and genomic regions. Smal restriction site sequence in bold.

Name	Sequence
Be1568TermF	5-TCTCCTAGGGATTCTTGTGACTGACACCTGG-3'
Be1568TermF	5-TCT <b>CCTAGG</b> GTTTTCGATTATGGAACAGCGCG-3'
SBhphBDF	5'-ATGAAAAAGCCTGAGCTTACCGC-3'
SBhphBDR	5'-TTCCTTTGCCCTCGGTCTAGT-3'
SBhphBDR Blastocladiella 1568 Genomic Region	5'-TTCCTTTGCCCTCGGTCTAGT-3' 5'-GTTTTCGATTATGGAACAGCGCGTCGTGCTCGCGCGCGAGCCGAGCCGGAGCCGGAGCCGAGCCGAGCCGGA TATCAAGCCGGGCGTGCTCCTGCAGCGCGGCGCG
	GTACAGAAACTTTTGCGAAACCAGGTGTCAGTCACAAGAATC-3'