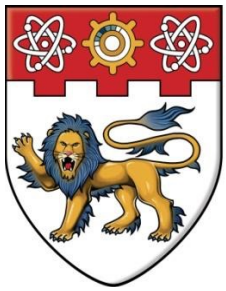


NANYANG TECHNOLOGICAL UNIVERSITY

CHARACTERIZATION OF CANDIDATE GENES GOVERNING COMPLEX MULTICELLULARITY IN *NEUROSPORA CRASSA*

WONG JIE YUN

SCHOOL OF BIOLOGICAL SCIENCES

2017

Characterization of candidate genes governing complex multicellularity in *Neurospora crassa*

Wong Jie Yun

School of Biological Sciences

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of Master of Science

2017

Acknowledgements

First and foremost, I would like to thank my Principle Investigator, Dr. Gregory Jedd, for giving me the opportunity to pursue my Master of Science degree, for his guidance, motivation and valuable teachings without which this project would not have been possible.

I would like to extend my sincerest gratitude to my supervisor, Dr. Mark Featherstone for accepting me as his student and for his guidance on my reports.

In addition, appreciation to the current and past members of the Jedd lab, Dr. Jamie Greig, Dr. Zheng Peng, Ms Yang Jing, Ms Peh Yee Xin, Dr. Julian Lai, Ms Rachel Loh and Ms Sherry Goh whom have provided me with laboratory assistance as well as imparted technical skills when I needed help.

Most importantly, I would like to thank my loved ones for their encouragement, company and endless support.

Table of contents

Acknowledgements ..... **i**
Table of contents ..... **ii**
List of figures ..... **iv**
Abbreviations ..... **v**
Abstract ..... **vii**
1 Introduction **1**
Unicellular versus multicellular organisms 1
Simple versus complex multicellular organisms 2
Neurospora crassa as a model organism 4
2 Material and methods **8**
Identification of candidate CM-associated genes 8
Neurospora strains and methods 8
Racetube assay 8
Marker Fusion Tagging (MFT) 9
N. crassa transformation 9
gDNA extraction 10
Genotyping / Polymerase chain reaction (PCR) 11
Crossing 11
Ascospore germination 12
Microscopy 12
Subcellular fractionation 12
Western Blotting 13
Detergent treatment 14
Immunoprecipitation 14
RNA extraction and cDNA synthesis 15
Reverse transcription polymerase chain reaction (RT-PCR) 16
Molecular Cloning 17
E. coli transformation and plasmid extraction 18
Screening of colonies and sequencing 18
Protein expression and purification 19
3 Results **21**
Bioinformatics search revealed 147 CM-associated genes 21
ROGDI is a cytoplasmic protein that associates with the RAVE complex 21
Hep2 is a membrane protein 30
Vez is an integral protein 32
NCU02049 is associated with the Spitzenkörper 33

4 Discussion.....36
5 Conclusions 40
6 References 41
7 Appendix.....44
1 List of genes from the bioinformatics search.....44
2 List of genes from ROGDI-HA MS results.....47
3 List of DNA oligonucleotides used in this study.....48

List of figures

Figure 1. Summary of the *Neurospora crassa* life cycle 5

Figure 2. Evolution of CM in fungi.....22

Figure 3. Immunoprecipitation results revealed association of ROGDI with V-ATPase and RAVE 24

Figure 4. Localization of V-ATPase and PMA in the *N. crassa* hyphae.26

Figure 5. Expression of ROGDI and Rav1 constructs in *E.coli* 29

Figure 6. Immunoprecipitation results revealed no distinct interacting partners of Hep2.....31

Figure 7. Immunoprecipitation results revealed no distinct interacting partners of Vez.....32

Figure 8. Investigation of the importance of coiled coil domains in SPZ1, SPZ2 and SPA2.....35

Abbreviations

Abbreviated form	Extended form
3D	Three-dimensional
Amp	Ampicillin
BD	Big Dye
CM	Complex Multicellularity
DTT	Dithiothreitol
EPS	Extracellular Polymeric Substance
ETDA	Ethylenediaminetetraacetic Acid
FGSC	Fungal Genetics Stock Center
GFP	Green Fluorescent Protein
HA	Hemagglutinin
HEPES	4-(2-hydroxyethyl)-1-Piperazineethanesμlfonic acid
Hyg	Hygromycin
IP	Immunoprecipitation
IPTG	Isopropyl β-D-1-Thiogalactopyranoside
LB	Luria Bertani
MFT	Marker Fusion Tagging
MS	Mass Spectrometry
NEB	New England Biolabs
Pan	Calcium Pantothenate
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PIC	Protease Inhibitor Cocktail
PMA	Plasma membrane ATPase
PMSF	Phenylmethylsulfonyl Fluoride
pTRPC-Hyg	pTRPC-Hygromycin
RAVE	Regulator of ATPase of Vacuoles and Endosomes
RT-PCR	Reverse Transcription Polymerase Chain Reaction

Abbreviated form	Extended form
SB	Sonication Buffer
SC	Synthetic Crossing medium
SDS-PAGE	Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis
Spitzenkörper	Spz
TMD	Transmembrane Domain
V-ATPase	Vacuolar H^+-translocating ATPase
V_B	Vogel's Plating Medium, Basal
V_N	Vogel's N medium
V_T	Vogel's Plating Medium, top
WT	Wild Type
β-me	Beta-Mercaptoethanol

Wong Jie Yun

Abstract

Complex multicellularity (CM) occurs exclusively in the Eukarya domain and emerged independently within six distinct eukaryotic clades including ascomycete fungi such as *Neurospora crassa*. All CM organisms are characterized by three-dimensional (3D) cellular organization, cell differentiation and intercellular communication. It was hypothesized that new aspects of CM could be discovered by searching for genes present in CM fungi but absent in unicellular yeast counterparts. To this end, we performed bioinformatics analysis to identify genes expressed in Pezizomycotina and *Neolecta irregularis* but absent in budding and fission yeast. In total, we identified 147 genes that fit our criteria and we characterized four of which displayed an observable growth defect in *N. crassa* knockout strains. Interestingly, the protein product of one of the candidate gene, ROGDI, was found to interact with a subunit of the Vacuolar H^+-translocating ATPase (V-ATPase) complex as part of the Regulator of ATPase of Vacuoles and Endosomes (RAVE) complex and could potentially function to maintain vacuolar pH levels. Another candidate protein, Spz1, was shown to interact with proteins known to function in cell polarity through its coiled-coil domain. Altogether, the results from this study provide clues on the extant of genes required for CM.

1 Introduction

Unicellular versus multicellular organisms

A cell is the most basic form of life on earth. The existence of the first single-celled organism was documented as far back as 3.5 billion years ago (Choi, 2015). Unicellular organisms are present in all the domains: Bacteria, Archae, and Eukarya. Several recurring characteristics of single-celled species include (1) having only a single compartment where all cellular activities are performed in the cytoplasm, (2) absorbing food and nutrients through diffusion/osmosis and (3) replicating through asexual reproduction. These features enable them to thrive independently in almost any environment and quickly expand their population.

Although unicellular organisms are self-sufficient and are present till today, the congregation of two or more unicellular species have been repeatedly shown to be evolutionary advantageous (Niklas, 2014). One excellent example of this assembly include the formation of biofilms. Essentially, biofilms compose of a group of microorganisms which adhere to each other to form a large sheet (Berne et al., 2015). The integrity of the structure is maintained by the secretion of an extracellular polymeric substance (EPS) by the microbes themselves (Decho, 2000). Interestingly, biofilms are ubiquitously found in nature and are highly responsive to changes in the local environment (Battin et al., 2016). This complex system is believed to increase the survivability of the lineage (Donlan and Costerton, 2002). The driving force for such

adaptation could arise from the disadvantages associated with the direct contact of unicellular organisms with the external environment.

Over the course of time, multicellular organisms began to emerge and many have postulated that these complex organisms evolved from unicellular ancestors. This transition has been stimulated by many factors including environmental changes, genomic innovation and natural selection (Carroll, 2001; Knoll and Carroll, 1999). In addition, studies have shown an increase in genes involved in cell differentiation, cell-cell communication, and adhesion (Rokas, 2008). In contrast to unicellular organisms, multicellular organisms consist of highly differentiated cell types that allow the expansion of the organism. This eventually leads to the formation of macroscopic tissues which require intricate systems to coordinate multiple processes. The characteristics of multicellular organisms include: (1) consist of more than one cell (2) cell adherence (3) division of labour among specialized cells for specific processes (4) coordinated cell signalling for communication (Niklas and Newman, 2013). Multicellularity comes with several advantages. These include (1) reducing predation by forming macroscopic structures (Stanley, 1973), (2) facilitating food consumption (Bonner, 1988), (3) improving methods of dispersal for reproduction (Bonner, 1988) and (4) dividing labor (Baldauf, 2003).

Simple versus complex multicellular organisms

Multicellularity can be further subcategorized into simple and complex multicellularity (CM). The former include organisms which undergo

differentiation programs to form simple filamentous or cluster structures held together by adhesive molecules. Although these structural adaptations benefit them by deterring predators and facilitate feeding, such organisms have limited intercellular communication and lack the transfer of food resources between cells. On the other hand, CM organisms form 3D structures that not only encourage cell-to-cell communication but also exposes a minimum amount of cells to the external environment. Not surprisingly, CM occurs exclusively in the Eukarya domain and arose relatively late in the history of life, entering the fossil record during the Ediacaran Period, more than three billion years after microbial life began to diversify (Knoll, 2011). It has been observed that CM emerged independently within six distinct eukaryotic clades: animals, embryophytic land plants, florideophyte red algae, laminarialean brown algae, basidiomycete fungi, and ascomycete fungi (Knoll, 2011). The sheer size and number of complex multicellular organisms allowed them to dominate the space in which we live.

It was hypothesized that the emergence of CM began with the co-option of existing genes, and ongoing research has begun to understand the nature of these genes. Important genes required for CM would encode adhesion proteins that allow the development of tissues with a defined 3D architecture and specialized cell types (King, 2004; Knoll, 2011; Rokas, 2008). Secondly, the fact that complex multicellular organisms exist exclusively in the Eukarya domain (Knoll, 2011) may suggest that the unique features expressed by an eukaryotic cell are necessary for the transition. Among other traits, eukaryotes contain a dynamic cytoskeletal and an endomembrane system. These features may allow

would-be complex multicellular eukaryotes to better respond to intercellular molecular signals by changing its morphology (Geiger et al., 2009) or secreting signalling molecules in a precise manner (Draber et al., 2012). The third important feature that complex multicellular organisms possesses are microscopic passageways across adjacent cell walls and membranes (Knoll, 2011). This connection between cells facilitate coordinated signaling in the form of molecular channels (Bloemendal and Kuck, 2013) to facilitate cell-cell transfer of nutrients and signaling molecules. Such channels include plasmodesmata in plants (Lucas and Lee, 2004), gap junctions in animals (Sosinsky and Nicholson, 2005) and septal pores in fungi (Jedd and Pieuchot, 2012). These features, and the expression of many other proteins, are believed to be required for the switch to complex multicellularity. However, this list of genes is far from complete and more research can be carried out to uncover more of such essential genes.

Neurospora crassa as a model organism

Neurospora crassa is an ideal organism in the study of complex multicellularization. Being categorized as an ascomycete fungi, *N. crassa* displays characteristics of CM including a complex 3D organization held together by adhesion molecules and a septal pore which allows cell to cell communication. It can undergo both asexual and sexual reproduction. In the former, the asexual cells, called conidia, germinate and grow into a mass of multinucleated, filamentous, branched threads termed hyphae. The hyphae are partitioned by crosswalls called septa (Harris, 2001) and the presence of a septal pore

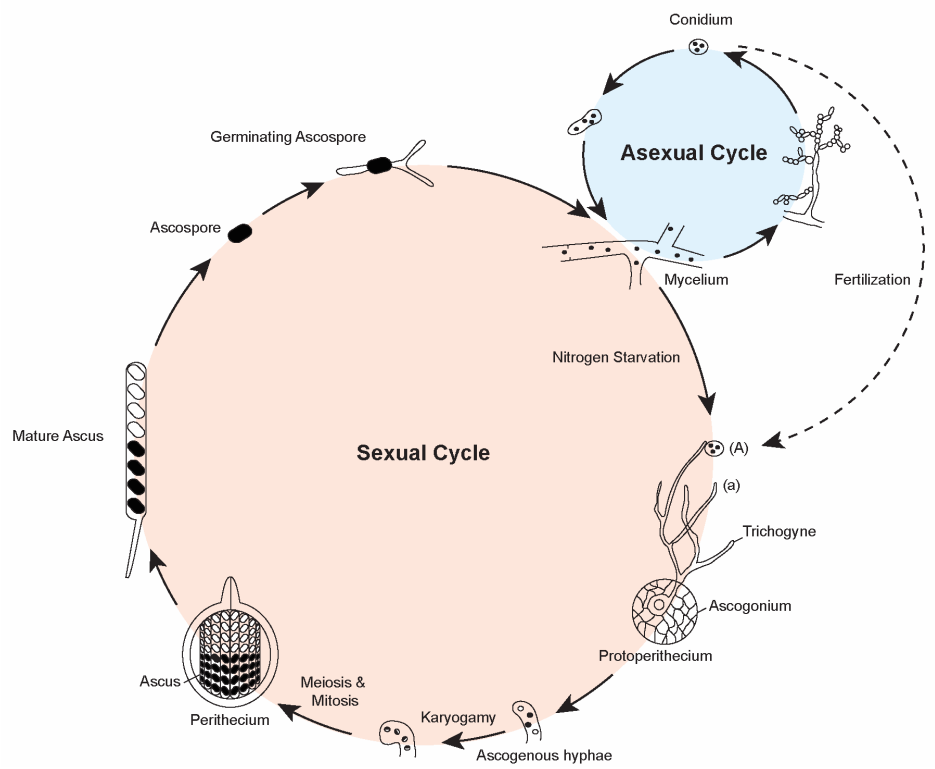


Figure 1. Summary of the *Neurospora crassa* life cycle. *N. crassa* undergoes both sexual and asexual cycle depending on the environmental conditions, depicted by beige and blue respectively. The sexual cycle is initiated upon nitrogen starvation as well as the presence of two opposite mating strains, denoted by (A) and (a).

allows intercellular communication. In the event of favourable conditions for sexual reproduction and the presence of strains of two different mating type, the organism can undergo sexual reproduction. This is stimulated in the laboratory by nitrogen starvation (Westergaard and Mitchell, 1947) which up regulates genes involved in sexual reproduction (Nelson and Metzenberg, 1992). The female reproductive structure termed protoperethecia is a complex structure consisting of a coiled, knot-like hyphae structure called the ascogonium. At the tip of the ascogonium is the trichogyne. Fertilization occurs when cells of the opposite mating type comes into contact with the trichogyne. After which the cell walls and nuclei of both parental strains fuses, producing a “dikaryon” in which nuclei from both parents proliferate as the sexual fruiting body called the perithecium develops. The nuclei of opposite

mating types undergo karyogamy in ascogenous hyphae and then undergo meiosis forming four haploid products. These undergo a further mitosis, producing eight ascospores enclosed in a sac called an ascus (Borkovich et al., 2004), which is the defining feature of Ascomycetes (Fig. 1).

The relatively short life cycle of *N. crassa* makes it favourable to be used as a model organism. In addition, its haploid genetics is extremely useful in studying gene function. It allows identification of phenotype associated with gene mutation after a sexual cross. This is facilitated by the removal of the nonhomologous end-joining DNA repair gene, *mus51/52*, that allows homologous recombination to occur at higher frequency (Colot et al., 2006). Due to this discovery, deletion strains are easily constructed and the collection of single gene deletions is available from the Fungal Genetics Stock Center (FGSC). In addition, techniques such as Marker Fusion Tagging (Lai et al., 2010) allows efficient fusion of fluorescent-protein tag to endogenous proteins of interest. These extensive genetic manipulation techniques make *N. crassa* an ideal model organism. Last but not least, the genome of the organism has been fully sequenced in 2004 (Galagan et al., 2003).

This project aims to use *N. crassa* as a model organism to uncover more genes essential for CM. To this end, we employed a systematic bioinformatics analysis to identify orthologous genes that are found exclusively in ascomycota multicellular fungi, including *N. crassa*, but not in unicellular fungi (Nguyen et al., 2017). This analysis focused on a subset of 147 of such genes. Four of these were chosen for further

characterization based on growth phenotypes exhibited in *N. crassa* knockout strains.

2 Material and methods

Identification of candidate CM-associated genes

Genes were selected for functional characterization if they encode a protein that is (1) present in *N. irregularis* and the Pezizomycotina, and (2) not detected in *S. cerevisiae*, *S. pombe* and non-fungal species using BLASTp (NCBI nr database, e-value = 1e-3) (Nguyen et al., 2017).

Neurospora strains and methods

Neurospora crassa strains used in this study were grown on synthetic Vogel's N medium (V_N) solid and liquid media (Davis and Serres, 1970). Strains grown on solid V_N slants were grown in the presence of light. Strains grown in liquid V_N were grown at 28^{o}C at 150 rpm. Genetic crosses were carried out on synthetic crossing medium (SC) as previously described (Davis and Serres, 1970) and left in the dark. *N. crassa* conidia were transformed by electroporation (Turner et al., 1997) and plated in the V_T layer on top of V_B (Davis and Serres, 1970). Ascospores were germinated by heat shock treatment and similarly plated in the V_T layer on top of V_B. Both type of plates were kept in a 30^{o}C incubator. *N. crassa* deletion strains (Colot et al., 2006) were obtained from the Fungal Genetics Stock Center (www.fgsc.net) unless otherwise stated. For transformation, FGSC 9719 *(mus-52::bar* mat a*)* was used. For crossing, FGSC 465 (pan-2 mat A) was used.

Racetube assay

Racetubes were made by dispensing 10 ml of V_N supplemented with 20 µg/ml Calcium pantothenate (Pan) into a 25 ml serological pipette and

laid flat to solidify. After which, the tip of the pipette was removed by mechanical force. Strains of interest were grown on V_N supplemented with 20 µg/ml Pan plates one day prior to the start of the assay. For each sample, three spots were excised at three distinct locations, 2 mm behind the growth front. These were placed on the 23rd mark on the pipette and growth rate was measured at 0, 24 and 48 h.

Marker Fusion Tagging (MFT)

The Hygromycin-Green Fluorescent Protein (Hyg-GFP), Hygromycin-mCherry (Hyg-mCherry) and 3X Hemagglutinin (HA) tags were constructed based on previous work (Lai et al., 2010). Briefly, the Hyg-GFP and Hyg-mCherry cassettes were constructed by fusing three sets of PCR fragments amplified using three primer sets (Appendix 3). The first and third pair of primers amplifies the homologous arms for recombination, whereas the second primer pair amplifies the Hyg-GFP or Hyg-mCherry. The three fragments contain approximately twenty base pairs of overlapping regions with the neighbouring fragment, which will allow them to conjugate using fusion PCR. As for the HA cassettes, four sets of amplified fragments were used. The 3X HA sequence is synthesized on the forward primer of the second set of primers. The pTRPC-Hygromycin (pTRPC-Hyg) cassette is inserted downstream of the HA tag. As such, primer set one, two and four amplifies the genomic region whereas set three on pTRPC-Hyg.

***N. crassa* transformation**

Purified donor cassettes were transformed into FGSC 9719 strains via electroporation. Briefly, the strain was grown on 5 ml V_N slants for 7-10 days. The conidia were resuspended in water and passed through 40

micron cell strainer, spun at 3 k rpm, 4^{o}C for 5 min. The pellet was washed for five times with 1 ml cold sterile water followed by another five washes with 1 ml cold 1 M Sorbitol and centrifuged at 3 k rpm, 4^{o}C for 5 min. After which the spores were left to recover on ice for 2 h. For each reaction, 50 µl of conidia were mixed with approximately 3 µl of DNA, depending on the quality of the donor DNA. Reactions were left on ice for 20 min, followed by electroporation at 2.1 kV. The transformed conidia were recovered in 1 ml of 1 M sorbitol at 30^{o}C for 1 h. Finally, the cells were plated at varying volume of 10, 50 and 100 µl on V_B and V_T (Davis and Serres, 1970) plates containing 100 µg/ml hygromycin B (Invitrogen, USA) (Hyg). After incubation for 4 days at 30^{o}C, 12 colonies were transferred onto 1 ml V_N slants containing 10 µg/µl Hyg.

gDNA extraction

The strains were left to conidiate for 5 days. After which, genomic DNA extraction was carried out using MasterPure™ Yeast DNA Purification Kit (Epicentre, USA). Generally, 100 µl of resuspended conidia was used for each extraction. The conidia were centrifuged at 16.2 k x *g* for 5 min. The pellet was then resuspended in 100 µl of Yeast Lysis buffer followed by incubation at 65^{o}C for 15 min. The samples were then incubated on ice for 5 min. Fifty microlitres of MPC Protein Precipitator was added to each samples, vortex and centrifuged at 16.2 k x *g* for 10 min. The supernatant was transferred to a new tube with 200 µl of isopropanol to precipitate the DNA, mixed well and centrifuged at 16.2 k x *g* for 10 min. The pellet was then washed with 70% ethanol. Finally, the DNA pellet was dissolved in 50 µl of TE buffer.

Genotyping / Polymerase chain reaction (PCR)

The genotype of the strains were analyzed by polymerase chain reaction (PCR) using a set of primer pair that flanks the site of integration. Briefly, reactions were carried out in 20 µl reaction volume in the presence of 3 µl of template, 0.5 µM primer mixture, 0.5 mM dNTPs mix (Promega, USA), and 5 U/µl Taq Polymerase (TLL, Singapore). Reactions were carried out with an initial denaturation step at 95^{o}C for 2 min, followed by 35 cycles of denaturation at 98^{o}C for 30 s, annealing at 56^{o}C for 30 s, and extension at 72^{o}C for 1 min per kilobase pair of the PCR product, and a final extension step at 72^{o}C for 10 min. PCR amplicons were resolved on 1% Agarose (Vivantis, USA) gel to determine if the insertion is present. This genotyping assay allows the validation of correctly inserted donor cassettes as well as distinguishing between homokaryon or heterokaryon strains.

Crossing

To ensure the insertions were homokaryons, the transformed strains with the correct integration were crossed with FGSC 465. Two microlitres of each strains with opposite mating types were spotted on opposite ends of a 20 ml SC plate supplemented with 20 µg/ml Pan. The plates were sealed with micropore tape and left in the dark. After 10 days, a fresh lid was introduced to replace the old one and the plates were left turned over, with the lid on the underside. After another 10 days, the ascospores were harvested by washing the lid with 10 ml of sterile water, and kept in 4^{o}C.

Ascospore germination

The ascospores were left at 4^{o}C overnight before use. The old water in which the ascospores were harvested was replaced with fresh sterile water. Generally, 5 µl of ascospores were used for each heat shock, depending on the density of the ascospores from the harvest. The spores were subjected to heat shock at 60^{o}C for 15 min. After which, they were plated onto V_B and V_T plates supplemented with 20 µg/ml Pan and 100 µg/ml Hyg. The plates were left in a 30^{o}C incubator for 2 days. Twenty-four colonies were excised and transferred onto 1 ml V_N slants containing 20 µg/ml Pan and 10 µg/ml Hyg. The slants were left to conidiate for 5 days and gDNA extraction was performed followed by genotyping.

Microscopy

Microscopy images were performed using the Leica TCS SP8 inverted confocal microscope with the HC PL APO 63x/1.40 OIL objective and acquired by Leica Application Suite X. White light laser of 50% was used. For GFP signal, emitted wavelength collected was between 505-555nm. As for mCherry signal, emitted wavelength collected was between 592 to 700nm.

Subcellular fractionation

The strains of interest were grown on 1 ml V_N slants for 5 days. The conidia were resuspended in water and grown in 20 ml of liquid V_N at 28^{o}C, 150 rpm for 30 h. The mycelium suspension was then filtered through a miracloth, washed thoroughly with cold sterile water and finally dried using filter paper. It was then immediately submerged into

liquid nitrogen and grounded using a mortar and pestle to produce a fine powder. IP buffer (20 mM HEPES pH 7.4, 150 mM KAc, 2 mM $MgCl_2$, 1 mM DTT, 0.5 mM PMSF, 1 cOmplete EDTA free tablet (Roche, Germany) per 10 ml) was added to the powder in a 1:1 ratio and incubated on ice for 10 min. The suspension was then centrifuged through a 50 micron cell strainer at 200 x *g* for 1 min at 4˚C followed by 500 x *g* for 1.5 min at 4˚C. The supernatant was distributed equally into three tubes and spun at 2 k, 10 k, 100 k x *g* respectively for 30 min at 4˚C. The pellet was washed twice with IP buffer. Finally, the pellet was dissolved in 2X Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) loading dye and IP buffer in a 1:1 ratio, based on the volume of cell extract used per centrifugation speed.

Western Blotting

Proteins were separated via SDS-PAGE followed by transferring onto polyvinylidene fluoride (PVDF) membrane at 30V, overnight. The blot was probed with the corresponding primary and secondary antibodies. Protein bands were detected using Immobilon Western Chemiluminescent HRP Substrate (Merck, Germany). The primary antibodies used in this study were as follows: rat anti-HA HRP (1:2500; Roche, Germany), mouse anti-GFP (1:1000; Roche, Germany), mouse anti-mCherry (1:2500; Chromotek, Germany), rabbit anti GAPDH (1:2500; Abcam, USA). The secondary antibodies used were as follows: sheep anti-mouse (1:5000; GE Healthcare, USA), donkey anti-rabbit (1:5000; GE Healthcare, USA).

Detergent treatment

350 µl of sample powder was incubated with IP buffer in a 1:1 ratio and incubated on ice for 10 min. The suspension was then centrifuged through a 50 micron cell strainer at 200 x *g* for 1 min at 4˚C followed by 500 x *g* for 1.5 min at 4˚C. One hundred microlitres of the supernatant was distributed into three tubes and spun at 10 k x *g* for 10 min at 4˚C. The pellet was then incubated with 150 µl of 1% Tween20, 1% Digitonin, 1% Triton X-100 or IP buffer on ice for 15 min. This was followed by centrifuging at 10 k x *g* for 30 min at 4˚C. The pellet was washed twice with IP buffer. Finally, the pellet was dissolved in 100 µl of 2X SDS-PAGE loading dye and 100 µl IP buffer. Proteins were detected using western blot assay as previously described.

Immunoprecipitation

The experiment was carried out using the Dynabeads® Co-Immunoprecipitation Kit (Novex, USA) according to the manufacturer's protocol. For each reaction, 15 mg of Dynabeads and 75 µg of Anti-HA High affinity antibody (Roche, Germany) were coupled together at 37˚C, overnight.

To immunoprecipitate cytosolic proteins, the proteins were extracted from the powder of the strain of interest by firstly adding IP buffer supplemented with 1% digitonin to the powder of interest in a 1:1 ratio followed by thawing on ice on 15 min. The suspension was centrifuged through 40 micron cell strainer at 500 x *g* and the crude cell extract spun at 16 k x *g* for 30 min at 4˚C. The supernatant was then incubated with the HA conjugated Dynabeads for 2 h. To elute the HA conjugated

complexes, the beads were first washed in the following order; once with IP buffer + 1% digitonin, twice with IP buffer, once with 1x LWB + 0.02% Tween20 and finally thrice with 1x LWB. Samples were eluted twice with HPH-EB by incubating at room temperature for 10 min for the first elution and 20 min for the second elution. Proteins were dried using Speed vac followed by electrophoresis separation using SDS-PAGE. Coomassie staining was performed and the bands were compared between the HA tagged strain and Wild-type (WT). Bands that were present in HA but absent in WT were excised and sent for analysis via Mass Spectrometry.

For membrane proteins, the proteins were extracted from the strain of interest by firstly adding IP buffer to the powder of interest in a 1:1 ratio followed by thawing on ice for 15 min. The suspension was centrifuged through 40 micron cell strainer at 500 x *g* and the crude cell extract spun at 16 k x *g* for 30 min at 4˚C. The pellet was then incubated with 1.5 ml IP buffer + 1% Triton X-100 or 1% digitonin, and incubated on ice for 20 min. The suspension was centrifuged at 16 k x *g* for 30 min at 4˚C and supernatant containing the extracted protein was then incubated with HA conjugated Dynabeads for 2 h. Washing and eluting was the same as above for the cytosolic proteins.

RNA extraction and cDNA synthesis

RNA extraction was performed on FGSC 465 sample powder. The RNA was extracted using RNeasy Plant Mini Kit (Qiagen, USA) according to the manufacturer's protocol. Briefly, 450 µl of buffer RLC + 4.5 µl β-mercaptoethanol was added to 100 µl of sample powder, vortex and

incubated at 56^{o}C for 1-3 min. The extract was then passed through the Qiashredder and centrifuged at maximum speed for 2 min. The supernatant was then transferred to the gDNA Eliminator spin column and centrifuged at maximum speed for 30 s. The flow through was transferred to a new tube and 0.5 volume of 100% ethanol was added and mixed immediately by pipetting. The solution was transferred to the RNeasy spin column and spun at 8 k x *g* for 15 s. The column was then washed once with 700 µl of Buffer RW1 and twice with 500 µl of Buffer RPE by centrifuging at 8 k x *g* for 15 s. Finally, the RNA was eluted with 30 µl of RNase-free water.

The extracted RNA was used as a template for cDNA synthesis using AffinityScript Multiple Temperature cDNA Synthesis Kit (Agilent Technologies, USA). Two and a half micrograms of RNA was incubated with 300 ng of random primers in a 15.7 µl reaction at 65^{o}C for 5 min and then 4^{o}C for 5 min in the thermocycler. The reaction mix was then supplemented with 1X AffinityScript RT buffer, 0.4 mM dNTP, 20 U RNase Block Ribonuclease Inhibitor and 1 µl AffinityScript Multiple Temperature Reverse Transcriptase to a final volume of 4.3 µl. The mixtures were then incubated at 25^{o}C for 10 min, 55^{o}C for 60 min and finally, 70^{o}C for 15 min.

Reverse transcription polymerase chain reaction (RT-PCR)

RT-PCR reactions were carried out using exon-specific primers. Essentially, reactions were done in a 20 µl volume in the presence of 5 µl cDNA template, 0.5 µM each of forward and reverse primer, 0.5 mM dNTPs, 5 U/µl Taq DNA polymerase (TLL, Singapore) and an

appropriate reaction buffer. PCR cycling conditions were as follows: initial denaturation at 94^{o}C for 2 min, followed by 35 cycles of denaturation at 94^{o}C for 30 s, annealing at 56^{o}C for 30 s, and extension at 72^{o}C for 1 min per kilo base pair of product, and a final extension step at 72^{o}C for 10 min. PCR amplicons were resolved on 2% agarose (Vivantis, USA) gel to determine if the cDNA synthesis was successful.

Molecular Cloning

The Rav constructs were cloned into pDUET vector at the BamHI/ EcoRI, for pDUET His-ROGDI, AscI/PmeI for pDUET His-Rav1 and BamHI/KpnI for pDUET His-ROGDI, S-Rav1. The inserts were amplified using KAPA HiFi DNA Polymerase (KAPA Biosystems, South Africa), in a 50 µl reaction with 500 ng of template, 0.4 µM each of forward and reverse primer, 0.4 mM dNTP, 0.5 U KAPA HiFi. PCR cycling conditions were as follows: initial denaturation at 95^{o}C for 3 min, followed by 35 cycles of denaturation at 98^{o}C for 20 s, annealing at 58^{o}C for 30 s, and extension at 68^{o}C for 1 min per kilo base pair of product, and a final extension step at 68^{o}C for 5 min. As for the backbone, 3 µg of pDUET vector was digested with 10 U of BamHI/EcoRI, AscI/PmeI, BamHI/KpnI enzymes (New England Biolabs, UK) with the appropriate buffer in a 50 µl reaction at 37^{o}C for 2 h. PCR amplicons and vector backbones were resolved on 1% agarose (Vivantis, USA) gel and band of interest were excised and gel purified using illustra GFX PCR DNA and Gel Band Purification Kits (GE Healthcare, USA). The concentration of the eluted DNA were quantified using nanodrop. The fragments were ligated with In-Fusion® HD Cloning Kit (Clontech, USA) using 25 ng and 50 ng of

vector and insert respectively, in a 5 µl reaction volume. The reaction mix were incubated at 50°C for 15 min and then 5 min on ice.

E. coli transformation and plasmid extraction

2.5 µl of plasmid were incubated with 25 µl of XLI-Blue cell (TLL, Singapore) on ice for 20 min. When using BL21 cells (TLL, Singapore), 1 µl of plasmid was used instead. Heat shock was performed at 42°C for 90 s and returned to ice for 2 min. One millilitre of Luria-Bertani (LB) broth (TLL, Singapore) was then added and the cells were left to recover at 37°C for 1 h. Lastly, 10% of the cells were plated onto LB agar plates supplemented with 100 µg/ml Ampicillin (Amp) and incubated at 37°C overnight. For each construct, 6 colonies were picked and grew in 4 ml of LB+Amp at 37°C, 200 rpm for 20 h. Plasmid extraction was performed on 3 ml of the culture using Axygen Plasmid Miniprep Kit (Axygen, USA). To pellet the bacteria, the LB culture was centrifuged at 12 k x *g* for 1 min. The pellet was then resuspended in 250 µl of Buffer S1 by vortexing. This was followed by addition of 250 µl of Buffer S2, tubes inverted gently to mix and then 350 µl of Buffer S3 and inverted gently to mix as well. To clarify the lysate, the samples were centrifuged at 12 k x *g* for 20 min, 4°C. The supernatant was then transferred to the spin column and centrifuged at 12 k x *g* for 1 min. Seven hundred microlitres of Buffer W2 was added and centrifuged at 12 k x *g* for 1.5 min and then 3 min. Finally, the DNA was eluted in 50 µl of eluent.

Screening of colonies and sequencing

To check for constructs with the successful inserts, 500 ng of plasmid were used in each reaction, supplemented with the appropriate buffer

with 2.5 U of enzyme used at the cloning site in a 20 µl reaction volume and incubated at 37ºC for 1 h. Bands were separated on a 0.5% agarose gel. Constructs showing correct number and size of fragments were selected for sequencing. One hundred nanograms of template was used with 6 µl of Big Dye (BD) (TLL, Singapore), 0.5 µM primer in a 20 µl reaction. The reaction was placed in a thermocycler under the following cycling conditions: initial denaturation at 96ºC for 1 min, followed by 25 cycles of denaturation at 96ºC for 10 s, annealing at 50ºC for 10 s, and extension at 60ºC for 1 min.

Protein expression and purification

The strain of interest were picked and grown in 5 ml of LB+Amp at 37˚C, 200 rpm for 16 h. The cultures were then upscaled to 50 ml and incubated at 37˚C, 200 rpm for another hour. Induction was performed by adding 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) to each culture and incubated at 26˚C, 200 rpm, 6 h. Finally, the cultures were centrifuged at 3 k x *g* for 15 min at 4˚C. The pellet was washed with 1 ml 1x Phosphate Buffered Saline (PBS) and centrifuged at 4.5 k x rpm for 15 min at 4˚C and then let dry. To sonicate the cells, the pellet was firstly resuspended in 2 ml of Resuspension Buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 8) with 10 mM imidazole, 1 cOmplete EDTA free tablet per 20 ml. Two millimolar Phenylmethylsulfonyl Fluoride (PMSF) was added to the samples prior to sonication, on ice. The lysate was then centrifuged at 18 k x *g* for 15 min at 4˚C. To allow binding of the His-tagged proteins with the beads, the supernatant was then transferred to a new tube and 20 mM imidazole and 100 µl of HisPur Ni-NTA Resin (Thermo Scientific, USA) was added and incubated at 4˚C, rolling for 1 h. This was followed

by centrifuging at 4.5 k x rpm for 5 min at 4˚C. The beads were washed with Wash buffer (Resuspension buffer + 20 mM imidazole) thrice and finally eluted twice using Elution buffer (Resuspension Buffer + 300mM imidazole).

3 Results

Bioinformatics search revealed 147 CM-associated genes

We used BLASTp (NCBI nr database, e-value = 1e-3) to select for *N. crassa* genes that are evolutionary conserved in (1) *N. irregularis* and the Pezizomycotina, but (2) not present in *S. cerevisiae, S. pombe* and non-fungal species. We reasoned that since *N. irregularis* displays features of CM, this organism should possess CM-related proteins while unicellular fungi and non-fungal species should not (Fig. 2A). The analysis generated 147 genes that fit this criteria (Appendix 1). The deletion strains of 147 genes were retrieved from the deletion collection purchased from FGSC and tested for an observable phenotype. Specifically, we looked for defects in hyphae growth. Hyphae consists of multiple cells interconnected via septal pores. Intercellular communication is one of the defining features of complex multicellularity. Therefore, hyphae growth defect was used to associate with complex multicellularity in this study. In total, seven strains displayed signs of hyphae growth defect (Fig. 2C) and in this report, we characterized four of those genes. The genes of interests are (1) *NCU08091*, (2) *NCU06509*, (3) *NCU09240* and (4) *NCU02049*.

ROGDI is a cytoplasmic protein that associates with the RAVE complex

The first protein that we were interested in from the bioinformatics study was the protein encoded by *NCU08091*. This gene encodes an uncharacterized protein but contains only a highly conserved Rogdi

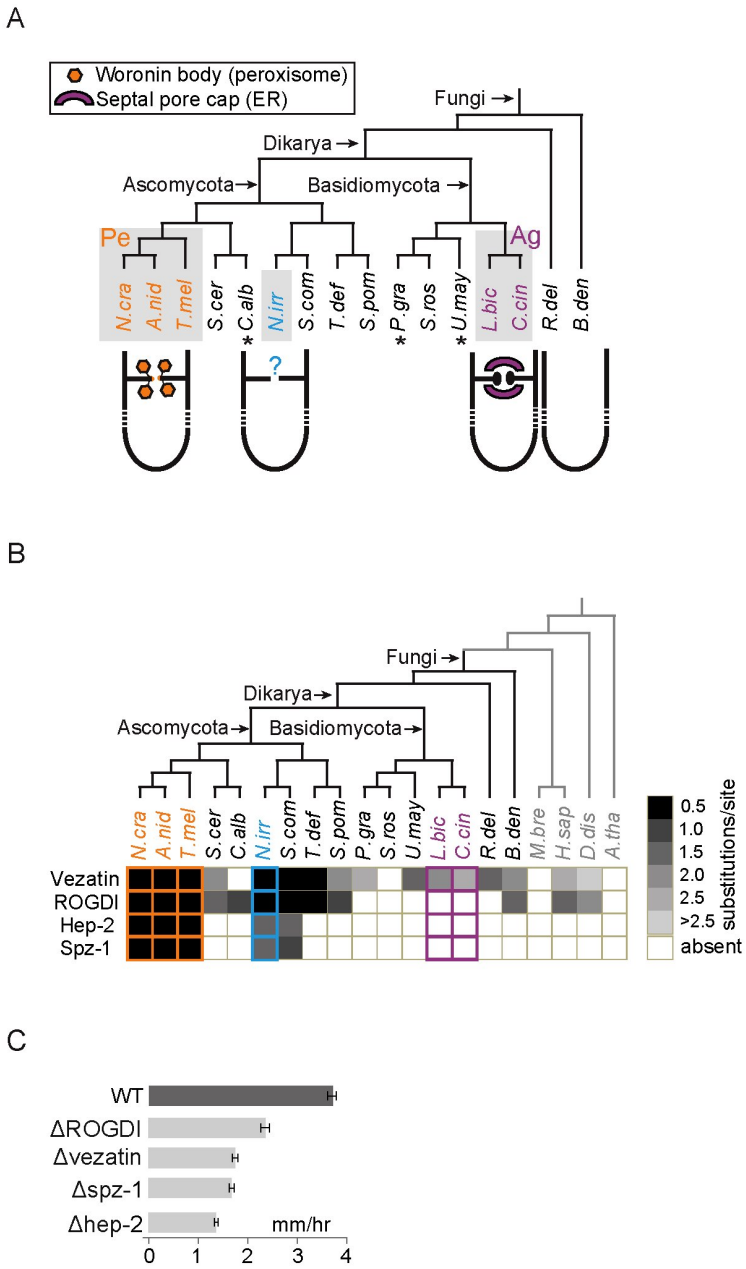


Figure 2. Evolution of CM in fungi. (A) The phylogeny of representative fungal species with sequenced genomes. CM taxa are shown with grey background. Abbreviations: Pe = Pezizomycotina, Ag = Agaricomycotina, Sa = Saccharomycotina, Ta = Taphrinomycotina, *N.cra* = *Neurospora crassa*, *A.nid* = *Aspergillus nidulans*, *T.mel* = *Tuber melanosporum*, *S.cer* = *Saccharomyces cerevisiae*, *C.alb* = *Candida albicans*, *N.irr* = *Neolecta irregularis*, *S.com* = *Saitoella complicata*, *T.def* = *Taphrina deformans*, *S.pom* = *Schizosaccharomyces pombe*, *P.gra* = *Puccina graminis*, *S.ros* = *Sporobolomyces roseus*, *U.may* = *Ustilago maydis*, *L.bic* = *Laccaria bicolor*, *C.cin* = *Coprinopsis cinerea*, *R.del* = *Rhizopus delemar*, B.den = *Batrachochytrium dendrobatis* (B) Phylogenetic distribution of the fours genes of interest in this study. (C) Growth rate of the four deletion strains of interest in comparison with WT.

leucine zipper domain. As it has two plausible mRNA transcipts, the predicted number of amino acids of the translated protein is 320 or 372 amino acids and as such, the predicted molecular weight is either 34.1 kDa or 39.4 kDa. The Rogdi leucine zipper domain spans almost the entire protein, at amino acid positions 26-311 or 25-363 of the two transcripts respectively. Thus, we refer the protein encoded by

NCU08091 as ROGDI. Since it is an uncharacterized protein, the nature of this protein has not been studied. Using a bioinformatics tool [(TMpred) (http://embnet.vital-it.ch/software/TMPRED_form.html)], it was predicted that ROGDI does not contain any transmembrane domain (TMD) or signal peptide. This might suggest that it functions as a cytosolic protein. To this end, we tagged the endogenous protein with hemagglutinin (HA) at the C-terminus using the MFT technique. Practically, a ROGDI-HA donor DNA was transformed into *N. crassa* to replace the endogenous gene via homologous recombination. Successfully-transformed strains were backcrossed to generate homokaryons, ensuring that the pool of strains were carrying only the HA tagged copy. To determine the subcellular distribution of the ROGDI protein, a cellular extract was subjected to incremental centrifugation at 2 k, 10 k and 100 k x *g*. Both the pellet as well as the supernatant from each spin were kept for western blot analysis. These results showed that ROGDI is predominantly found in the supernantant (Fig. 3A). Interestingly, we observed higher amounts of protein in the pellet fraction at higher centrifugation speeds (Fig. 3A). These results suggest that ROGDI primarily functions as a cytoplasmic protein but may occasionally associate with the membrane.

To determine the protein interacting partners of ROGDI protein, we carried out an *in vivo* immunoprecipitation assay using the ROGDI-HA strain. As shown previously from our subcellular fractionation results, the protein was found mainly in the supernatant fraction. Therefore, we immunoprecipitated ROGDI-HA from the 10 k x *g* supernatant fraction in the presence of 1% digitonin. The extracted ROGDI-HA complex were

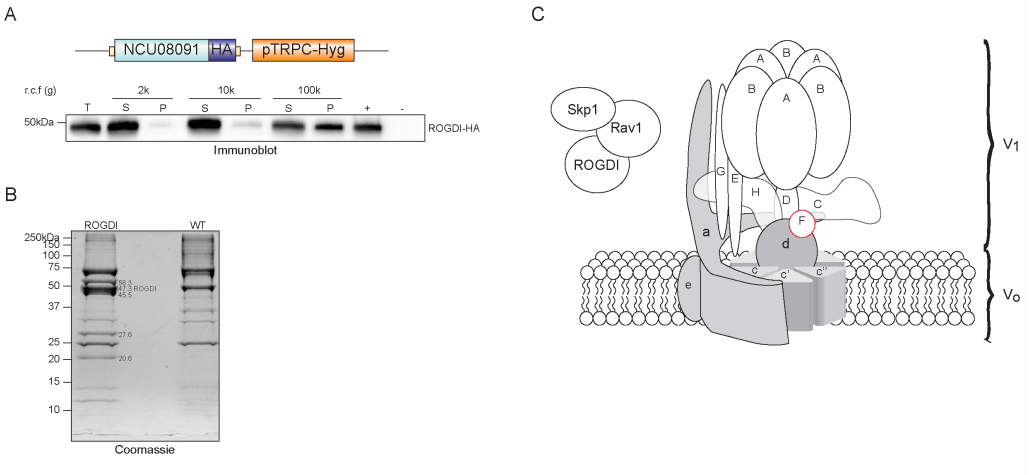


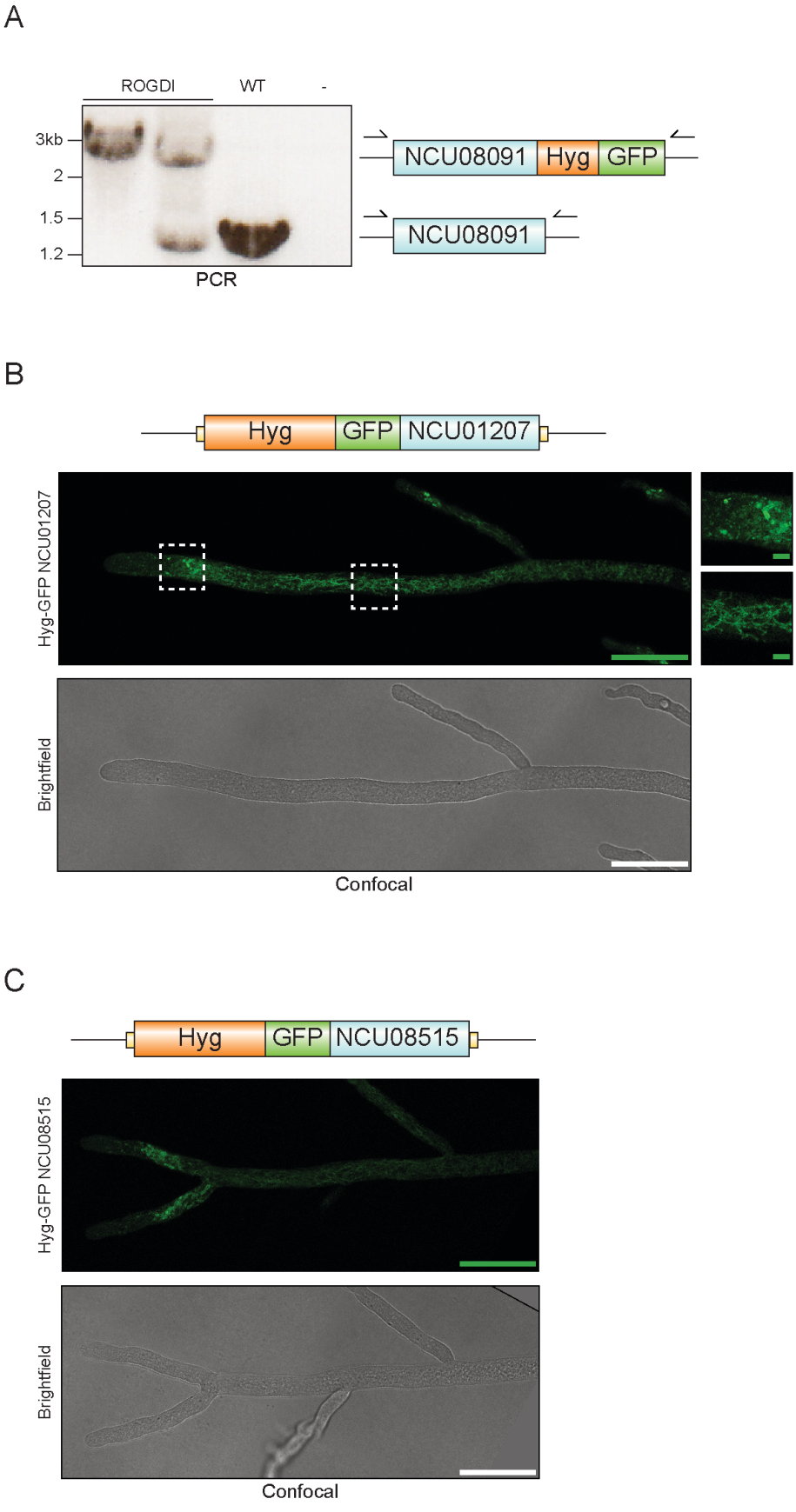
Figure 3. Immunoprecipitation results revealed association of ROGDI with V-ATPase and RAVE. (A) Top: Fusion cassette of ROGDI with 3X HA and the downstream pTRPC-Hyg used in the generation of ROGDI-HA via MFT. Bottom: Immunoblot of subcellular fractionation of ROGDI-HA construct at 2 k, 10 k , 100 k x *g*. T=Total, S=Supernatant, P=Pellet,+=ROGDI-HA strain which serves as a positive control, -=WT strain that does not express HA, which serves as a negative control. (B) Results of IP. Coomassie staining of the immunoprecipitated ROGDI-HA strain with a WT untagged strain. Estimated molecular weight of bands present in ROGDI-HA but absent in WT strain are written next to the band on the gel picture. (C) IP results revealed targets of the whole RAVE complex as well as the V_1 subunit of the V-ATPase (white) except for F (circled in red).

separated using SDS-PAGE and the protein profile was compared to a control strain without the HA tagged using coomassie staining. From the results, we found four protein bands that were enriched in the presence of ROGDI-HA protein (Fig. 3B). To identify the nature of these proteins, we excised these gel segments, together with the corresponding segments of the control lane, and analyzed using Mass Spectrometry (MS). The results revealed that the complex contains components of the Vacuolar H^+-translocating ATPase (V-ATPase) as well as the regulator of ATPase of vacuoles and endosomes (RAVE) complex (Appendix 2, Fig. 3C). Interestingly, the data show that ROGDI-HA protein interacts with almost all subunits of the periphery subunit (V_1) of the V-ATPase, with the exception of subunit F, but not with the membrane subunit (V_o). Therefore, our immunoprecipitation results clearly show that ROGDI associates with V-ATPase via the V_1 subunit as well as the cytoplasmic RAVE complex and this could account for its localization in the cytoplasm. Interestingly, the yeast orthologue of ROGDI has been

previously found to interact with V-ATPase and the RAVE complex (Smardon et al., 2002).

To visualize the subcellular localization of ROGDI, we fused Green Fluorescence Protein (GFP) sequences to both the N- and C-terminus of our target gene using the MFT technique. However, we were unable to detect any fluorescence signal, despite the successful GFP integration (Fig. 4A). This lack of signal might be due to its low mRNA expression at the mycelium stage or due to the misfolding of the fusion protein into an unstable structure. Next, we tagged both proteins A (*NCU01207*) (Fig. 4B) and B (*NCU08515*) (Fig. 4C) of the V_1 subunit at the N-terminus with GFP. Microscopy results revealed that both proteins localized in a similar pattern, despite a clear difference in intensity. Interestingly, the V-ATPase complex distributes in a segmental manner in the cell. Towards the hyphal tip, the complex forms a vacuole-like structure, while at the sub-apical region, it forms a strand-like structure (Fig. 4B). This may suggest that V-ATPase localizes within a vacuolar tubular compartment. Since ROGDI functions in a complex with the V_1 subunit of V-ATPase, we may postulate that it co-localizes in a similar segmental manner. However, further optimizations on the tagging of the protein have to be done to validate this.

The interaction between ROGDI and V_1 subunit of the V-ATPase complex suggests that the latter component also fractionates in the supernatant of the cell extract. To test this, we spun the *N. crassa* cell extract expressing N-terminus tagged protein A of the V_1 V-ATPase subunit at 2 k x *g*, 10 k x *g* and 100 k x *g*. Western blot analysis


A
ROGDI
WT
-
3kb
2
1.5
1.2
PCR
NCU08091
Hyg
GFP
NCU08091
B
Hyg
GFP
NCU01207
Hyg-GFP NCU01207
Brightfield
Confocal
C
Hyg
GFP
NCU08515
Hyg-GFP NCU08515
Brightfield
Confocal

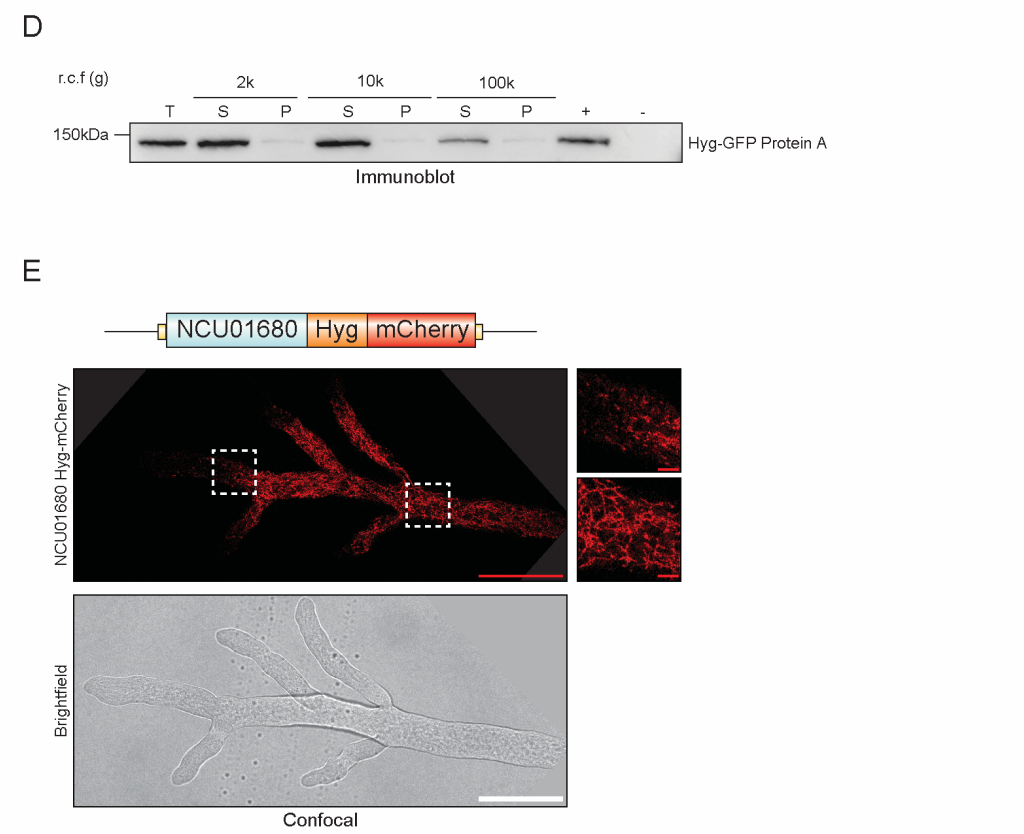


Figure 4. Localization of V-ATPasea and PMA in the *N. crassa* hyphae. (A) Gel electrophoresis results of PCR to genotype ROGDI-HygGFP. Lane 1 is an example of a homokaryon whereas lane 2 is that of a heterokaryon. Lane 3 uses WT as a template which serves as a control with no successful integration. -=Non-template control that serves as a negative control. (B) Top: Fusion cassette of Protein A of V_1 subunit of the V-ATPase with Hyg-GFP used in the generation of Hyg-GFP NCU01207 via MFT. Bottom: Confocal imaging of the strain harbouring this construct. Scale bar = 50μm; Scale bar of square zoomed images = 5μm. (C) Top: Fusion cassette of Protein B of V_1 subunit of the V-ATPase with Hyg-GFP used in the generation of Hyg-GFP NCU08515 via MFT. Bottom: Confocal imaging of the strain harbouring this construct. Scale bar = 50μm. (D) Immunoblot of subcellular fractionation of Hyg-GFP Protein A at 2 k, 10 k ,100 k *x g*. T=Total, S=Supernatant, P=Pellet, +=Hyg-GFP Protein A strain which serves as a positive control, -=WT strain that does not express Hyg-GFP, which serves as a negative control. (E) Top: Fusion cassette of PMA with Hyg-mCherry used in the generation of PMA Hyg-mCherry via MFT. Bottom: Confocal imaging of the strain harbouring this construct. Scale bar = 50μm; Scale bar of square zoomed images = 5μm.

confirmed that the protein A was predominantly found in the supernatant fraction, similar to the subcellular fractionation of ROGDI (Fig. 4D). This is likely due to the nature of the V_1 subunit which forms a peripheral component of the V-ATPase complex, which is not embedded in the cell membrane (Kane, 2006).

Eukaryotic cells express two distinct proton pumps, Vacuolar H^+-translocating ATPase (V-ATPase) and plasma membrane ATPase (PMA). Both pump H^+ out of the cytoplasm, so we wondered how PMA distributes itself in the hyphae of *N. crassa*. To this end, we tagged

Pma-1 (*NCU01680*) with a Hyg-mCherry sequence at the C-terminus via the MFT method. Surprisingly, the protein also distributes in a segmental manner throughout the whole hyphae. However, as compared to V-ATPases, Pma-1 does not form circular structures near the apical end (Fig. 4E). This is in agreement to the exclusive function of V-ATPase complex on vacuoles. Although both V-ATPase and PMA localizes segmentally in the hyphae, it would be interesting to determine whether these complexes are in close proximity to each other. In the future, we plan to express both tagged-proteins in a single *N. crassa* strain and test their co-localization.

Our MS results indicate that *Neurospora* ROGDI is a component of the RAVE complex and interacts with the V_1 subunit of the cellular V-ATPase. To test whether these interactions are direct, we carried out an *in vitro* binding assay. We cloned in the open-reading frames of these proteins into the pETDUET vector. The product of the cloning include (1) Rav1 in-frame with the N-terminus His tag present on the plasmid itself, (2) ROGDI with N-terminus His tag (3) Rav1 with His tag and ROGDI with S tag in the same vector. These constructs were transformed into BL21 expression vector. However, only the construct harbouring the His-tagged ROGDI was successfully expressed (Fig. 5A). Therefore, we bound the purified His-tagged ROGDI protein to the Ni-NTA beads and incubated the reaction with *N. crassa* cell extract that expressed protein A of the V-ATPase complex fused to a GFP sequence. However, our data showed no interaction between these two proteins (Fig. 5B).

The manner in which RAVE regulates the V-ATPase remains unclear. We wondered whether the RAVE function is dependent on pH. To test this, we repeated the experiment and modified the pH of the reaction to pH 6, 7 or 8. However, we were still not able to observe any interaction between His-ROGDI and protein A-GFP (Fig. 5C). It is possible that this was due to the competition by the endogenous RAVE complex for His-ROGDI binding. To overcome this, we plan to engineer a strain that expresses a GFP-tagged protein A of the V-ATPase complex in a ROGDI deletion background. This may reduce the competition by the RAVE complex and enrich binding of protein A to the His-ROGDI protein. We can also co-express His-ROGDI and protein A-GFP. Lastly, we may change the GFP position of protein A to C-terminus as the current tagging strategy might interfere with interaction or structure with His-ROGDI.

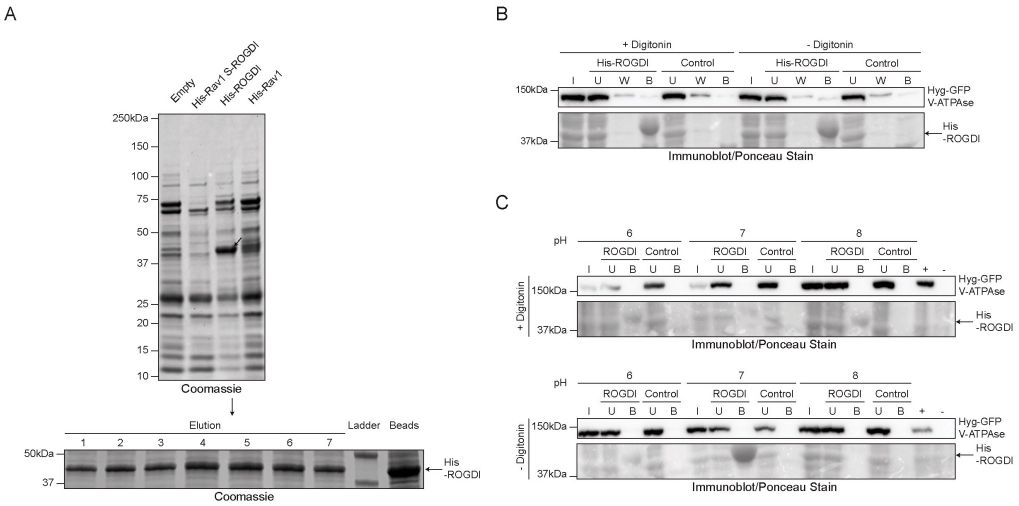


Figure 5. Expression of ROGDI and Rav1 constructs in *E.coli*. (A) Top: Coomassie staining of the BL21 cell lysate expressing the four different pETDUET constructs after 6 hours of induction. Bottom: Coomassie staining of the eluted His-ROGDI after binding to HisPur Ni-NTA Resin for 1h at 4^{o}C, rolling. Arrows indicate His-ROGDI which is 41.5kDa. (B) Top: Immunoblot of V-ATPase cell lysate incubated with His-ROGDI beads for 1 h at 4^{o}C, rolling, to determine if GFP fused V-ATPase interacts with His-ROGDI with and without digitonin treatment. I=Input, U=Unbound, W=Wash, B=Beads. Bottom: Ponceau stain to prove the presence of His-ROGDI as indicated by the arrow. (C) Top: Immunoblot of V-ATPase cell lysate treated with or without digitonin, incubated with His-ROGDI beads for 1 h at 4^{o}C, rolling, to determine if V-ATPase interacts with His-ROGDI with and without digitonin treatment, and at varying pH conditions of 6, 7 and 8. +=ROGDI-HA strain which serves as a positive control, -=WT strain that does not express HA, which serves as a negative control. Bottom: Ponceau stain to prove the presence of His-ROGDI as indicated by the arrows.

Hep2 is a membrane protein

The second analyzed protein is encoded by *NCU06509* and we refer to it as Hep2 (Nguyen et al., 2017). This gene encodes a novel protein of unknown function. We first tested the subcellular fraction of this protein by fusing it with a 3X HA tag using the MFT technique. Cell extracts were separated by centrifugation at 2 k, 10 k and 100 k x *g* and both the pellet and supernatant were analyzed by western blotting. The results indicated that the protein was found mainly in the pellet (Fig. 6A). This suggests that Hep2 functions as an integral or peripheral membrane protein. A distinctive feature of integral membrane proteins include the TMD that functions to anchor the structure to the lipid membrane bilayer. Interestingly, Hep2 is predicted to contain such transmembrane domain. To validate this, we treated the 10 k x *g* pellet fraction where Hep2 in enriched in with various detergents (1% Tween20, 1% digitonin and 1% TritonX-100) followed by centrifugation. These chemicals serve to solubilize the lipid bilayer and allow transmembrane proteins to shift towards the supernatant fraction. Indeed, our analysis showed that incubation with 1% digitonin and 1% TritonX-100, but not 1% Tween20, weakened the interaction of the Hep2 from the pellet as seen from the increase in levels in the supernatant fraction (Fig. 6B). Therefore, we conclude that Hep2 is a transmembrane protein.

Next, we were interested to determine whether Hep2 interacts with other protein partners. To test this, we carried out an immunoprecipitation assay using the previously generated Hep2-HA strain. The complex was isolated from the pellet of the cell extract treated with 1% digitonin or 1%

TritonX-100. Gel separation followed by coomassie staining of the purified proteins showed no enrichment of pulled-down proteins as compared to a control strain that does not express a HA-tag protein bait (Fig. 6C). One possibility for the lack of enrichment could be that the Hep2-HA complex remain strongly bounded to the beads. Indeed, western blot analysis revealed high amounts of bound Hep2-HA protein (Fig. 6D). Therefore, to test for Hep2 interacting partners, we collected the beads and separated the bound proteins on SDS-PAGE. However, we were still unable to detect an enrichment of protein interacting partners using coomassie staining. In addition MS analysis failed to detect specific interacting partners of Hep2, which suggests that further optimization of the immunoprecipitation protocol is required. We plan to look into this in the future.

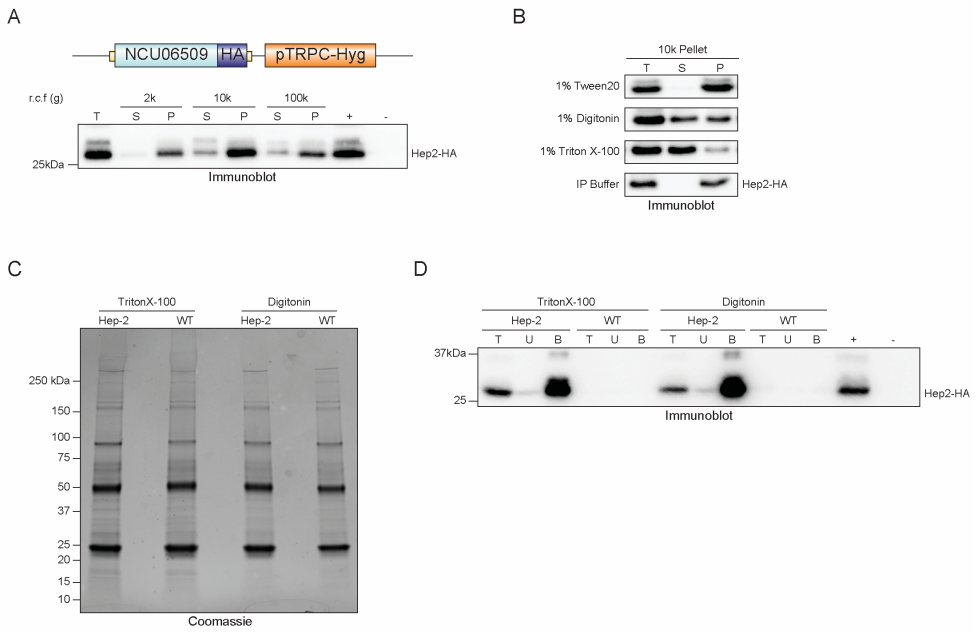


Figure 6. Immunoprecipitation results revealed no distinct interacting partners of Hep2. (A) Top: Fusion cassette of Hep2 with 3X HA and the downstream pTRPC-Hyg used in the generation of Hep2-HA via MFT. Bottom: Immunoblot of subcellular fractionation of Hep2-HA construct at 2 k, 10 k ,100 k x *g*. T=Total, S=Supernatant, P=Pellet,+=Hep2-HA strain which serves as a positive control, -=WT strain that does not express HA, which serves as a negative control. (B) Immunoblot of detergent treatment of the 10 k *x g* pellet. (C) Coomassie staining of the immunoprecipitated Hep2-HA strain with a WT untagged strain using TritonX-100 and Digitonin. (D) Immunoblot of the samples from the different IP steps to determine where Hep2-HA has been lost. T=Total, U=Unbound, P=Pellet,+=Hep2-HA strain which serves as a positive control, -=WT strain that does not express HA, which serves as a negative control.

Vez is an integral protein

NCU09240 encodes an uncharacterized protein that contains a vezatin domain. As such, we will refer to this protein as vezatin (Vez). We first tested the subcellular fraction of this protein by fusing it with a HA tag. The cell extract was spun at 2 k, 10 k and 100 k x *g* and both the pellet and supernatant fractions were analyzed by western blotting. The results indicated that the protein was predominantly found in the pellet when spun at 2 k, 10 k x *g* and exclusively in the pellet when spun at 100 k x *g* (Fig. 7A). TM prediction analysis revealed the presence of four TM domains and this strongly suggest it being an integral protein. To validate this, we collected the 10 k x *g* pellet and treated it with various detergents (1% Tween20, 1% digitonin and 1% TritonX-100) followed by a similar speed centrifugation. Our analysis showed that incubation with all the detergent weakened the interaction of the Vez from the pellet, with the least shift with 1% Tween20 (Fig. 7B). Therefore, we conclude that Vez is a transmembrane protein.

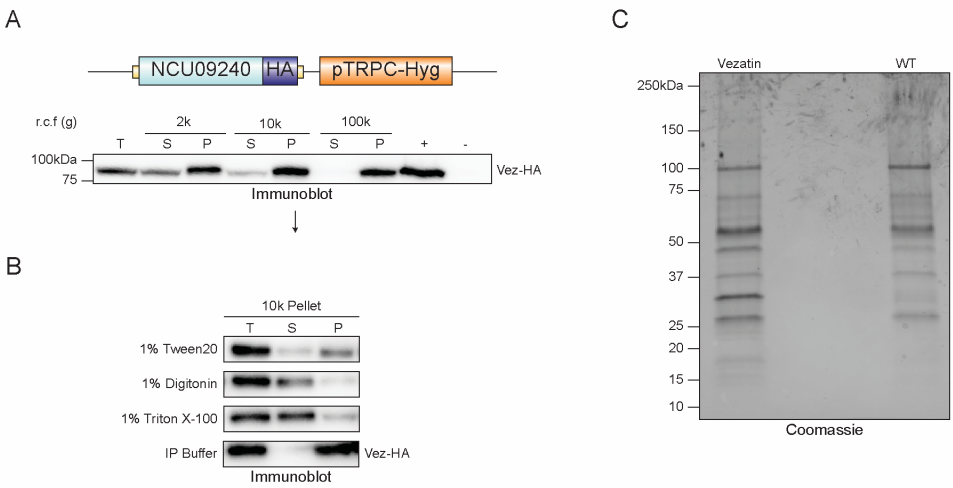


Figure 7. Immunoprecipitation results revealed no distinct interacting partners of Vez. (A) Top: Fusion cassette of Vez with 3X HA and the downstream pTRPC-Hyg used in the generation of Vez-HA via MFT. Bottom: Immunoblot of subcellular fractionation of Vez-HA construct at 2 k, 10 k ,100 k x *g*. T=Total, S=Supernatant, P=Pellet, +=Vez-HA strain which serves as a positive control, -=WT strain that does not express HA, which serves as a negative control. (B) Immunoblot of detergent treatment of the 10 k *x g* pellet. (C) Coomassie staining of the immunoprecipitated Vez-HA strain with a WT untagged strain.

To uncover Vez interacting partners, we performed immunoprecipitation using the previously generated Vez-HA strain. The complex was isolated from the pellet of the cell extract treated with 1% TritonX-100. Gel separation followed by coomassie staining of the purified proteins showed no enrichment of pulled-down proteins as compared to a control strain (Fig. 7C).

NCU02049 is associated with the Spitzenkörper

The protein encoded by *NCU02049* was another hit revealed from the bioinformatics study. The MFT tagging of this protein revealed its localization at the hyphae tip, near the Spitzenkörper (Spz) vesicle supply centre, which is associated with hyphal growth and morphogenesis. As such, we name the protein product encoded by this gene Spz1. Previous immunoprecipitation data from our lab revealed that Spz1 interacts with several other proteins including another novel protein, Spz2, and the known polarity proteins, Spa2 and Myo-2. Protein architecture analysis revealed conserved coiled coil domains in the Spz1, Spz2 and Spa2 proteins (Fig. 8A). To begin to dissect the function of these proteins, we systematically replaced segments of each protein with the Hyg-mCherry cassette using the MFT method. Phenotypic changes that occur with these deletions were determined by comparing the localization and growth rate of the mutant protein as compared to the N-terminally-tagged proteins, which are all fully functional as assessed by growth rate and localization to the Spz. Deletion of segments 1 and 2 of the Spz1 protein showed no change in the localization of the protein, while deletion of the C-terminal third of the

protein (segment 3) showed a loss of tip localization. This suggests that segment 3 is required for transport of Spz1 to the tip (Fig. 8B), while segment 2 executes a distinct function and is possibly related to vesicle binding.

Three deletions were also introduced in the Spz2 protein. Deletion of segments 1 and 3 resulted in protein mislocalization as well as hyphal growth defect (Fig. 8B). Further inspection of the protein sequence revealed that the replacement of segment 1 of Spz2 protein deleted a highly conserved region (the UDS domain), which has an unknown function. Surprisingly, the middle segment of the protein appeared to be dispensable to the function of the protein, as seen from the unchanged phenotype. These data define segments 1 and 3 as being required for Spz2 function and localization. Current work is aimed at determining how these deletions affect the ability of Spz2 to oligomerize with Spa2.

Spa2 is predicted to contain two coiled coil domains in the centre portion of the protein. We generated four deletion segments in Spa 2 as indicated in Figure 8A. These data define segments 1 and 3 as playing essential roles, while segments 2 and 4 appear to be largely dispensible for function. Segment 3 coincedes with a coiled-coil domain and appears to be required for tip-localization. By contrast, segment 1 coincides with a predicted N-terminal globular domain and does not appear to be required for localization to the Spz. These data define two functionally distinct regions of Spa2. Current work is aimed at determining how these regions affect overall oligomer formation.

To rule out the possibility that the mislocalization of the mutant protein was due to protein misexpression or protein degradation, we carried out western blot analysis and probed for the mCherry tag. The results showed that in general, all mutant proteins were expressed at comparable amounts to the tagged full length protein (Fig. 8C). It is noteworthy that expression of Spz1 which lacked the first segment appeared to be unstable. This was also consistent with the low intensity of the fluorescent signal in the previous experiment. We conclude that most of the mutant proteins that fail to localize at the hyphal tip do so due to loss of targeting information and not due to protein instability.

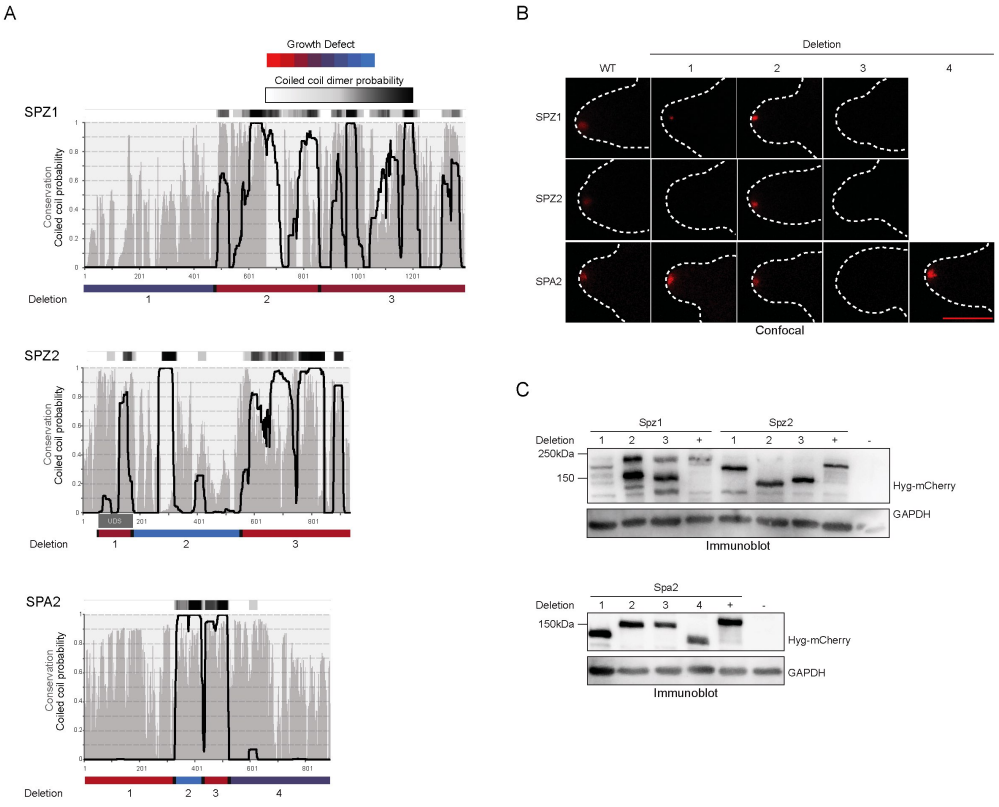


Figure 8. Investigation of the importance of coiled coil domains in SPZ1, SPZ2 and SPA2. (A) Plot of conservation (grey) and coiled coil probability (black) of SPZ1, SPZ2 and SPA2. Each protein is divided into 3 or 4 segments based on coiled coil probability and replaced with Hyg-mCherry cassette via MFT. Growth rate of each strain harbouring the deleted segment was determined via race tube assay. Red regions represent deletion areas that showed the most hyphal growth defect whereas blue, the least growth defect, comparable to a terminally tagged strain. (B) Confocal imaging of the respective deletion strains. Scale bar=10μm. (C) Immunoblot of the different deletion strains. +=Hyg-mCherry full length N-terminus tagged strain which serves as a positive control, -=WT strain which does not express Hyg-mCherry, which serves as a negative control.

4 Discussion

In an attempt to uncover genes essential for complex multicellularity in fungi, we screened for highly conserved orthologous genes expressed in multicellular fungal species but absent in unicellular ones. This work focuses on a 147 of 1050 genes (Nguyen et al., 2017). Four of these displaying growth defect in knockout *N. crassa* strains were selected for additional investigation. Although the other candidate genes did not contribute to fungal growth, more work needs to be done to investigate their role in other important CM-related processes.

ROGDI, encoded by the *NCU08091* gene in *N. crassa*, is a highly conserved protein that has yet to be characterized. One interesting piece of information about this protein is that mutation of its orthologue in human causes a disease known as Kohlschütter-Tönz syndrome, characterized by nervous system-related phenotypes including epilepsy, psychomotor regression and amelogenesis imperfecta (Schossig et al., 2012). In this study, we found that ROGDI functions as a cytoplasmic protein. The pulldown experiment indicated that ROGDI interacts with at least one component of the V_1 subunit, and the remaining components of the V_1 subunit interacts with each other. It would be interesting to investigate the direct mechanism of ROGDI and V-ATPAse interaction. This highly convincing interaction points to the possibility that ROGDI performs similar functions as Yeast Rav2, which functions in the RAVE complex to regulate the assembly of the V_1 subunit of the V-ATPase by dissociating it from the membrane (Smardon et al., 2014). Owing to the fact that V-ATPases function by pumping protons from the cytoplasm

into the vacuoles, we postulate that ROGDI may assist to regulate the pH environment of organelles by influencing V-ATPase assembly. Experiments that measure changes in the pH of organelles would allow us to address this possibility. The importance of ROGDI in vacuolar pH maintenance could explain the hyphae growth defect observed in its knockout *N. crassa* strain. We hope that in the future, our data may provide insights into the molecular mechanisms underlying Kohlschütter-Tönz Syndrome and may lead to therapeutic possibilities.

We showed that the ROGDI-interacting complex, V-ATPase, distributes in a segmental manner hyphae structures of *N. crassa*. At the periphery of the hyphae, the complexes are predominantly associated with circular vacuoles while at the sub-apical region, they localize with vacuolar tubular compartments. Similarly, we showed that Pma-1 localizes in a segmental manner, but does not associate with circular vacuoles. In the future, it will be interesting to determine if these proton pumps are localized to the same membranous structures, or define distinct sub-compartments.

To test the interaction between Rav1 and ROGDI, we initially planned to fuse each of the proteins with Histidine tags by expressing their open reading frame in a pET-DUET vector. However, only His-ROGDI was successfully expressed. Several reasons may explain for the failure to express the other tagged constructs. This include the large protein size and the probable inability of the protein to fold properly in *E.coli*. The His-tagged ROGDI was abundantly expressed and it binds to the Ni-NTA beads with great affinity. Incubation of the bound His-ROGDI

protein with *N. crassa* cell extract expressing the protein A Hyg-GFP tagged V-ATPase showed no interaction between these proteins. We believe that this lack of observable interaction is due to technical difficulties associated with *in vitro* binding of these proteins. One way to overcome this problem, and conclusively demonstrate the direct interaction between ROGDI and V_1 V-ATPase, would be to express purified forms of both proteins and incubate them *in vitro.*

Based on the work done from Kane's lab (Kane, 2006), the association of the V_1 and V_o subunits of V-ATPase is glucose dependent in *S. cerevisiae*. To maintain the homeostasis of the cellular H^+, the assembly of the V-ATPase by the RAVE complex helps to maintain the cytosolic levels of H^+. Therefore, it might be worthwhile to explore the significance of glucose levels in V-ATPase association in *N. crassa*.

In this study, we showed that both Hep2 and Vez proteins function as integral membrane proteins, as seen from our subcellular fractionation data. However, we were not successful in determining the interacting partners of these protein using our current immunoprecipitation protocol. We believe that these proteins may interact with other protein partners in a transient manner. Therefore, one possible change that can be made to the protocol would be to increase the amount of cell extract to be used for immunoprecipitation. This may in turn increase the detection of the interacting proteins.

In mammalian epithelial cells, Vezatin was firstly discovered to function as a complex with E-cadherin-catenin as well as actin cytoskeleton

(Kussel-Andermann et al., 2000). These focal adhesion molecules are required for maintaining cell to cell contacts and intercellular communication in a complex multicellular organism. Interestingly, Vezatin was found to be essential for maturation of neuronal synapses and dendritic spines, a developmental system that relies on intricate cell communication (Danglot et al., 2012). Thus, these findings provide clues on the contribution of Vezatin protein to development of CM. Future work in *Neurospora* can focus on whether Vezatin plays a role in cell-cell adhesion that occurs during fruiting body formation.

A coiled coil domain structure occurs when two or more alpha helices are wound around each other (Burkhard et al., 2001). The predicted prevalence of coiled coils in Spz1 as well as its interacting partners, Spz2 and Spa2 indicates a functional importance of this structure. By replacing segments of the protein with an mCherry tag, we were able to determine the importance of these coiled coil domains in the localization of the proteins. Apart from the coiled coil domains, we found other highly conserved protein segments on Spz1, Spz2 and Spa2 proteins, which have yet to be characterized. Future work will make use of these variants to dissect domain function and determine how this complex functions to organize vesicles. This can be done by crossing these variants with strains expressing GFP-tagged vesicle markers.

5 Conclusions

Accumulating evidence suggests that the emergence of complex multicellularization (CM) is associated with an increase in organelle complexity. To uncover the list of genes required for CM, we systematically searched for genes which are encoded in ascomycete multicellular fungal species including *Neurospora crassa*, but absent in unicellular yeasts. In total, we examined 147 potential CM-related genes that derived from our bioinformatics analysis. Four genes which displayed hyphal growth defects in knockout strains were further characterized using subcellular fractionation experiments, immunoprecipitation assay, binding assay, fluorescent-tagging and deletion analysis. These proteins play a role in diverse aspects of cellular processes and may indicate that the onset of CM requires an interplay between a wide range of different protein families. The intricate involvement of these different classes of protein for the onset of CM is intriguing to uncover. In conclusion, the results gathered from this study may provide further clues on the genetic requirements for CM and shed light on the evolutionary transition to multicellularity.

6 References

Baldauf, S.L. (2003). The deep roots of eukaryotes. Science *300*, 1703-1706.

Battin, T.J., Besemer, K., Bengtsson, M.M., Romani, A.M., and Packmann, A.I. (2016). The ecology and biogeochemistry of stream biofilms. Nat Rev Microbiol *14*, 251-263.

Berne, C., Ducret, A., Hardy, G.G., and Brun, Y.V. (2015). Adhesins Involved in Attachment to Abiotic Surfaces by Gram-Negative Bacteria. Microbiol Spectr *3*.

Bloemendal, S., and Kuck, U. (2013). Cell-to-cell communication in plants, animals, and fungi: a comparative review. Naturwissenschaften *100*, 3-19.

Bonner, J. (1988). The Evolution of Complexity by Means of Natural Selection.

Borkovich, K.A., Alex, L.A., Yarden, O., Freitag, M., Turner, G.E., Read, N.D., Seiler, S., Bell-Pedersen, D., Paietta, J., Plesofsky, N.*, et al.* (2004). Lessons from the genome sequence of Neurospora crassa: tracing the path from genomic blueprint to multicellular organism. Microbiol Mol Biol Rev *68*, 1-108.

Burkhard, P., Stetefeld, J., and Strelkov, S.V. (2001). Coiled coils: a highly versatile protein folding motif. Trends Cell Biol *11*, 82-88.

Carroll, S.B. (2001). Chance and necessity: the evolution of morphological complexity and diversity. Nature *409*, 1102-1109.

Choi, C.Q. (2015). How Did Multicellular Life Evolve? In Astrobiology Magazine.

Colot, H.V., Park, G., Turner, G.E., Ringelberg, C., Crew, C.M., Litvinkova, L., Weiss, R.L., Borkovich, K.A., and Dunlap, J.C. (2006). A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. Proc Natl Acad Sci U S A *103*, 10352-10357.

Danglot, L., Freret, T., Le Roux, N., Narboux Neme, N., Burgo, A., Hyenne, V., Roumier, A., Contremoulins, V., Dauphin, F., Bizot, J.C.*, et al.* (2012). Vezatin is essential for dendritic spine morphogenesis and functional synaptic maturation. J Neurosci *32*, 9007-9022.

Davis, R., and Serres, F.d. (1970). Genetic and microbiological research techniques for Neurospora crassa. Methods Enzymol, 79-143.

Decho, A.W. (2000). Microbial biofilms in an intertidal system: an overview. Cont Shelf Res *20*, 1257-1273.

Donlan, R.M., and Costerton, J.W. (2002). Biofilms: survival mechanisms of clinically relevant microorganisms. Clin Microbiol Rev *15*, 167-193.

Draber, P., Sulimenko, V., and Draberova, E. (2012). Cytoskeleton in mast cell signaling. Front Immunol *3*, 130.

Galagan, J.E., Calvo, S.E., Borkovich, K.A., Selker, E.U., Read, N.D., Jaffe, D., FitzHugh, W., Ma, L.J., Smirnov, S., Purcell, S.*, et al.* (2003). The genome sequence of the filamentous fungus Neurospora crassa. Nature *422*, 859-868.

Geiger, B., Spatz, J.P., and Bershadsky, A.D. (2009). Environmental sensing through focal adhesions. Nat Rev Mol Cell Biol *10*, 21-33.

Harris, S.D. (2001). Septum formation in Aspergillus nidulans. Curr Opin Microbiol *4*, 736-739.

Jedd, G., and Pieuchot, L. (2012). Multiple modes for gatekeeping at fungal cell-to-cell channels. Mol Microbiol *86*, 1291-1294.

Kane, P.M. (2006). The where, when, and how of organelle acidification by the yeast vacuolar H+-ATPase. Microbiol Mol Biol Rev *70*, 177-191.

King, N. (2004). The unicellular ancestry of animal development. Dev Cell *7*, 313-325.

Knoll, A.H. (2011). The Multiple Origins of Complex Multicellularity. Annu Rev Earth Planet Sci *39*, 217-239.

Knoll, A.H., and Carroll, S.B. (1999). Early animal evolution: emerging views from comparative biology and geology. Science *284*, 2129-2137.

Kussel-Andermann, P., El-Amraoui, A., Safieddine, S., Nouaille, S., Perfettini, I., Lecuit, M., Cossart, P., Wolfrum, U., and Petit, C. (2000). Vezatin, a novel transmembrane protein, bridges myosin VIIA to the cadherin-catenins complex. EMBO J *19*, 6020-6029.

Lai, J., Ng, S.K., Liu, F.F., Patkar, R.N., Lu, Y., Chan, J.R., Suresh, A., Naqvi, N., and Jedd, G. (2010). Marker fusion tagging, a new method for production of chromosomally encoded fusion proteins. Eukaryot Cell *9*, 827-830.

Lucas, W.J., and Lee, J.Y. (2004). Plasmodesmata as a supracellular control network in plants. Nat Rev Mol Cell Biol *5*, 712-726.

Nelson, M.A., and Metzenberg, R.L. (1992). Sexual development genes of Neurospora crassa. Genetics *132*, 149-162.

Nguyen, T.A., Cissé, O.H., Wong, J.Y., Zheng, P., Hewitt, D., Nowrousian, M., Stajich2, J.E., and Jedd, G. (2017). Innovation and

constraint leading to complex multicellularity in the Ascomycota. Nature Communications, in press.

Niklas, K.J. (2014). The evolutionary-developmental origins of multicellularity. Am J Bot *101*, 6-25.

Niklas, K.J., and Newman, S.A. (2013). The origins of multicellular organisms. Evol Dev *15*, 41-52.

Rokas, A. (2008). The origins of multicellularity and the early history of the genetic toolkit for animal development. Annu Rev Genet *42*, 235-251.

Schossig, A., Wolf, N.I., Fischer, C., Fischer, M., Stocker, G., Pabinger, S., Dander, A., Steiner, B., Tonz, O., Kotzot, D.*, et al.* (2012). Mutations in ROGDI Cause Kohlschutter-Tonz Syndrome. Am J Hum Genet *90*, 701-707.

Smardon, A.M., Diab, H.I., Tarsio, M., Diakov, T.T., Nasab, N.D., West, R.W., and Kane, P.M. (2014). The RAVE complex is an isoform-specific V-ATPase assembly factor in yeast. Mol Biol Cell *25*, 356-367.

Smardon, A.M., Tarsio, M., and Kane, P.M. (2002). The RAVE complex is essential for stable assembly of the yeast V-ATPase. J Biol Chem *277*, 13831-13839.

Sosinsky, G.E., and Nicholson, B.J. (2005). Structural organization of gap junction channels. Biochim Biophys Acta *1711*, 99-125.

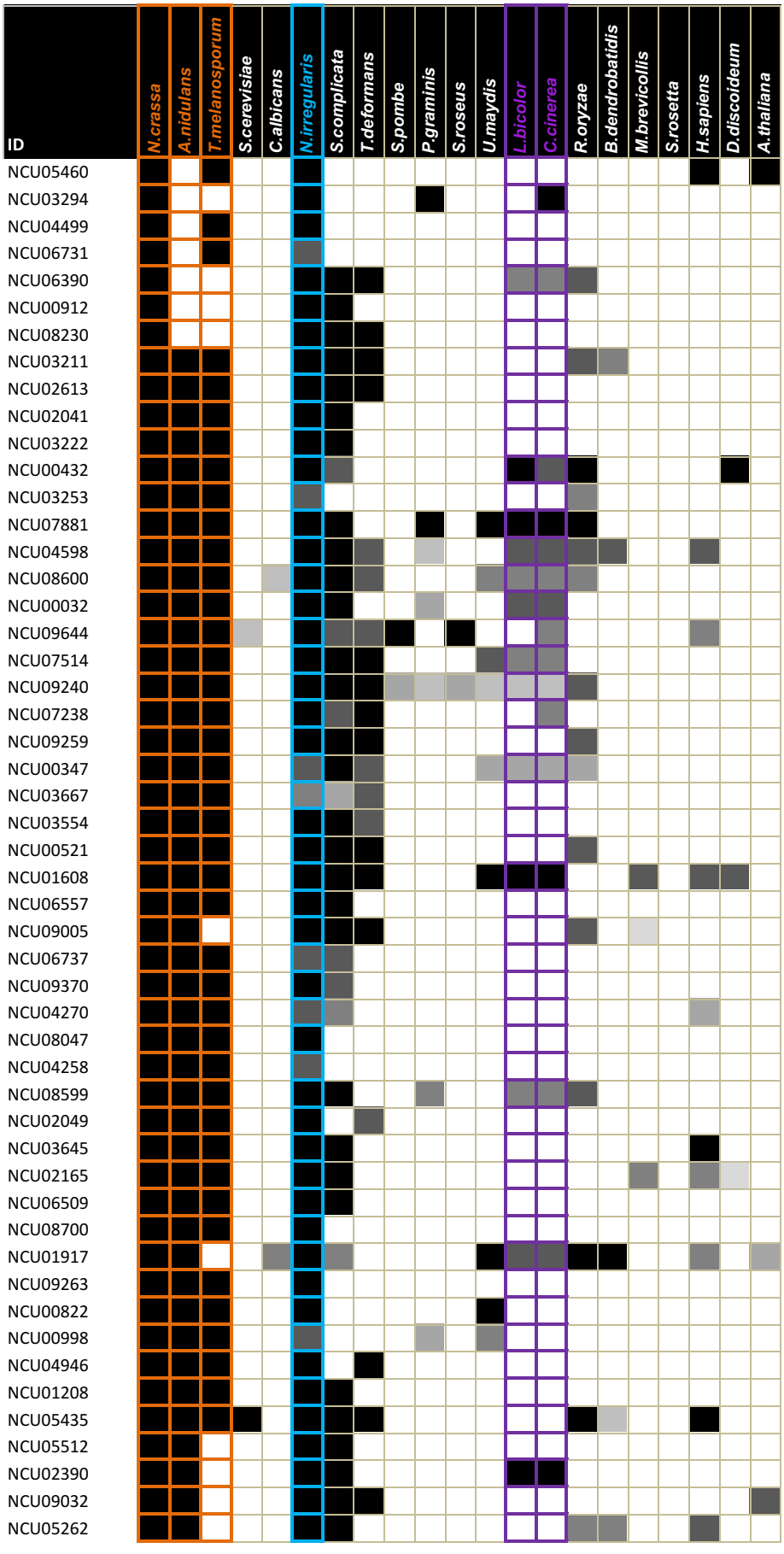
Stanley, S.M. (1973). An ecological theory for the sudden origin of multicellular life in the late precambrian. Proc Natl Acad Sci U S A *70*, 1486-1489.

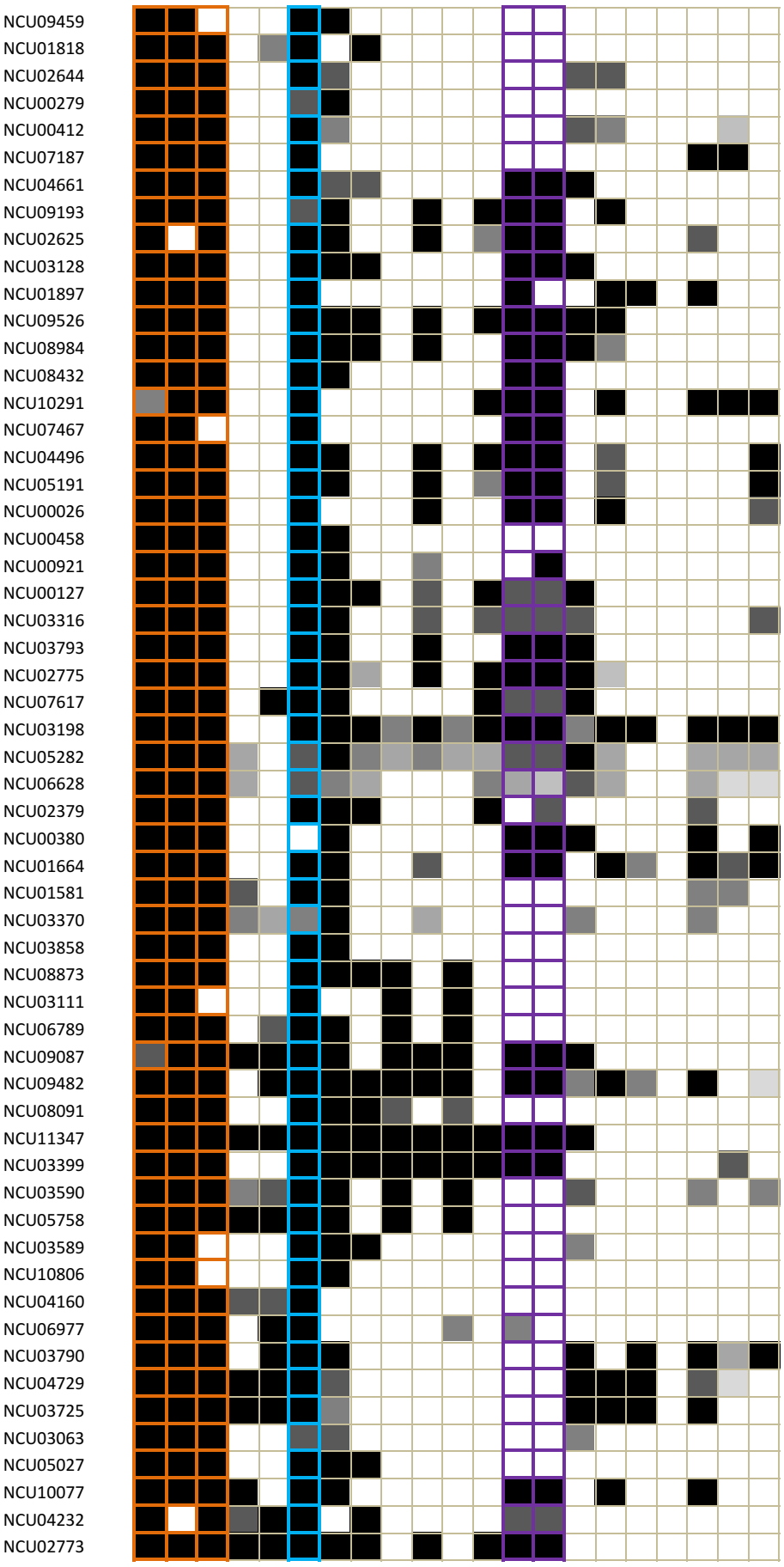
Turner, G., Jimenez, T., Chae, S., Baasiri, R., and Borkovich, K. (1997). Utilization of the Aspergillus nidulans pyrG gene as a selectable marker for transformation and electroporation of Neurospora crassa. Fungal Genetics Newsletter *44*, 55-59.

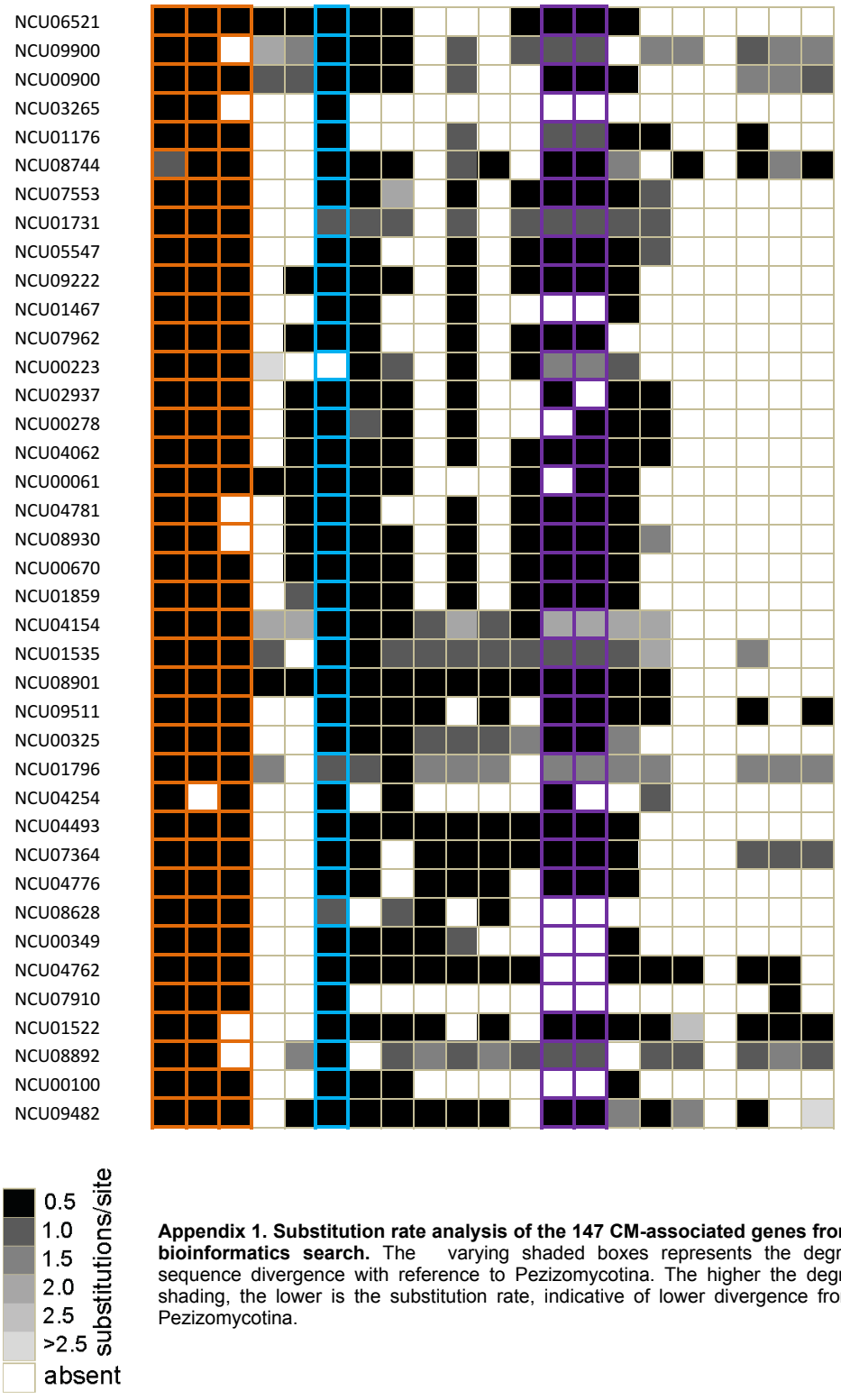
Westergaard, M., and Mitchell, H.K. (1947). Neurospora V. A Synthetic Medium Favoring Sexual Reproduction. American Journal of Botany *34*, 573-577.

7 Appendix

1 List of genes from the bioinformatics search




NCU09459
NCU01818
NCU02644
NCU00279
NCU00412
NCU07187
NCU04661
NCU09193
NCU02625
NCU03128
NCU01897
NCU09526
NCU08984
NCU08432
NCU10291
NCU07467
NCU04496
NCU05191
NCU00026
NCU00458
NCU00921
NCU00127
NCU03316
NCU03793
NCU02775
NCU07617
NCU03198
NCU05282
NCU06628
NCU02379
NCU00380
NCU01664
NCU01581
NCU03370
NCU03858
NCU08873
NCU03111
NCU06789
NCU09087
NCU09482
NCU08091
NCU11347
NCU03399
NCU03590
NCU05758
NCU03589
NCU10806
NCU04160
NCU06977
NCU03790
NCU04729
NCU03725
NCU03063
NCU05027
NCU10077
NCU04232
NCU02773



Appendix 1. Substitution rate analysis of the 147 CM-associated genes from our bioinformatics search. The varying shaded boxes represents the degree of sequence divergence with reference to Pezizomycotina. The higher the degree of shading, the lower is the substitution rate, indicative of lower divergence from the Pezizomycotina.

2 List of genes from ROGDI-HA MS results

Gene number	Gene name
NCU01207	vacuolar membrane ATPase-1 / Subunit A
NCU01658	WD repeat protein / RAVE 1
NCU02181	40S ribosomal protein S4
NCU02746	vacuolar membrane ATPase-13 / Subunit H
NCU03023	phenol 2-monooxygenase
NCU03901	peroxin 14
NCU05118	vacuolar membrane ATPase-10
NCU05347	histone H2Az
NCU07446	vacuolar membrane ATPase-4 / Subunit E
NCU08035	vacuolar ATPase subunit D
NCU08091	ROGDI
NCU08389	60S ribosomal protein L20
NCU08500	40S ribosomal protein S8
NCU08515	vacuolar membrane ATPase-2 / Subunit B
NCU08991	sulfur control-3
NCU09602	heat shock protein 70-1
NCU09760	hypothetical protein
NCU09897	vacuolar membrane ATPase-5

3 List of DNA oligonucleotides used in this study

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
27GFPpHyg P1	CCTCGACCTCCTAACCATGA		Construct
27HApHyg P2	ACATCGTAAGGGTATGCCATATCCTCCAGCTCCTCTCCCA		Construct
27HApHyg P5	AGGATATGGCATACCCTTACGATGTTCCTGACTATGCGGGCTATCCCTATGACGTCCCGGACTATGC CATGGGCTACCCTTACGACGTTCCAGATTACGCTTAGAGGGCCCGACGGCTGCT		Construct
27GFPpHyg P6	TGCTCCTTCAATATCATCTTGCGTTCTCGAGAAGCTACCATACTAGTG		Construct
27GFPpHyg P7	CACTAGTATGGTAGCTTCTCGAGAACGCAAGATGATATTGAAGGAGCA		Construct
27HApHyg P8	CATGTCTGGAACTTGATAAACAAAAACAATCGTATTTGTACCCACTATTCCT	NCU08091-HA	Construct
24HApHyg P9	AGGAATAGTGGGTACAAATACGATTGTTTTTGTTTATCAAGTTCCAGACATG		Construct
27GFPpHyg P10	AACAAAGGGAAACAGTCCACA		Construct
27GFPpHyg P11	AACAAAGGGAAACAGTCCACA		Genotyping
27GFPpHyg P12 (NEW)	GGGGGATGTTGATACGTTTG		Genotyping

27GFPpHyg P1	CCTCGACCTCCTAACCATGA		Construct
27HGFP P2	GGTGAGTTCAGGCTTTTTCATATCCTCCAGCTCCTCTCCCAT		Construct

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
27HGFP P3	ATGGGAGAGGAGCTGGAGGATATGAAAAAGCCTGAACTCACC	NCU08091-HygGFP	Construct
27GFPpHyg P4	AGCAGCCGTCGGGCCCTCTACTTGTACAGCTCGTCCATGC		Construct
27GFPpHyg P5	GCATGGACGAGCTGTACAAGTAGAGGGCCCGACGGCTGCT		Construct
27HGFP P6	GAACTTGCTCGAGTGGGAGA		Construct
27GFPpHyg P11	ACAAACCCTCACCATCAACC		Genotyping
27HGFP P8	CCACGTCAACTAGGCGGTAT		Genotyping

21HA P1	GACGAAGGGACCAGTTGCTA	NCU06509-HA	Construct
21HA P2	ACATCGTAAGGGTATGCCATACGGTACTCCTTCTCCTTGA		Construct
21HA P5	ACCGTATGGCATACCCTTACGATGTTCCTGACTATGCGGGCTATCCCTATGACGTCCCGGACTATGC CATGGGCTACCCTTACGACGTTCCAGATTACGCTTAAAGTGTAGATATAAACGGCACAAG		Construct
21HA P6	TGCTCCTTCAATATCATCTTGCGATCAGATGTTTCAGCTTGATC		Construct
21HA P7	CTTTGATCAAGCTGAAACATCTGATCGCAAGATGATATTGAAGGAGCA		Construct
21HA P8	AAACCATGGAAGAAGCTGTCTTAATTAAGTTCGGTCGGCATC		Construct

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
21HA P9	GATGCCGACCGAACTTAATTAAGACAGCTTCTTCCATGGTTT		Construct
21HA P10	TTCTTCGATGTCGGTTTTGTC		Construct
21HApHyg P11	CCAGGAATGGCAGGAGACTA		Genotyping
21HApHyg P12	GGCGAACTTGGCAGCAAT		Genotyping
54GFPpHyg P1	CATCCCGTATTCTTGCACAC		Construct
54HA P2	GAACATCGTAAGGGTATGCCATTAAAGAAATCCTCGTCGGC		Construct
54HA P5 (NEW)	CTTTAATGGCATACCCTTACGATGTTCCTGACTATGCGGGCTATCCCTATGACGTCCCGGACTATGCC ATGGGCTACCCTTACGACGTTCCAGATTACGCTTGACGACAGACAAACTCTG		Construct
54GFPHA P6 (NEW)	TGCTCCTTCAATATCATCTTGCGGGAACCACCAATACCACAGGTGGT		Construct
54GFPpHyg P7	CCTGTGGTATTGGTGGTTCCCGCAAGATGATATTGAAGGAGCA		Construct
54GFPpHyg P8	GGCAATGGACGATAAAAAGAGACTATTCCTTTGCCCTCGGA	NCU09240-HA	Construct
54GFPpHyg P9	TCCGAGGGCAAAGGAATAGTCTCTTTTTATCGTCCATTGCC		Construct
54GFPpHyg P10	TCGGAAGATTCGAAGACGAG		Construct

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
54GFPpHyg P11	CAGCTGAGCAGTCTGACAGG		Genotyping
54GFPpHyg P12	AGGTGCTCCTGGTCAATGTC		Genotyping
HGFP VMA-A P1	GGGGCAGCTTGTTTACTTCA	HygGFP NCU01207	Construct
HGFP VMA-A P2	GGTGAGTTCAGGCTTTTTCATCTTGGGCGAAAGGGCGATT		Construct
HGFP VMA-A P3	AATCGCCCTTTCGCCCAAGATGAAAAAGCCTGAACTCACC		Construct
HGFP VMA-A P4	CAAGGCTTACGGGAGCCATCTTGTACAGCTCGTCCATG		Construct
HGFP VMA-A P5	CATGGACGAGCTGTACAAGATGGCTCCCGTAAGCCTTG		Construct
HGFP VMA-A P6	CATTTCTTCTTGCGGTCCAG		Construct
HGFP VMA-A P7	ATGCTGGGTTCCTGATGAAG		Genotyping
HGFP VMA-A P8	ACAGTACCCCAGACGTCACC		Genotyping
HGFP VMA-B P1	GAGTTTCGGCAAAAGAGTGG		Construct

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
HGFP VMA-B P2	GGTGAGTTCAGGCTTTTTCATTGTGGAAGAGTGCGGACGT	HygGFP NCU08515	Construct
HGFP VMA-B P3	ACGTCCGCACTCTTCCACAATGAAAAAGCCTGAACTCACC		Construct
HGFP VMA-B P4	TACCCGGGGGTCGGCCATCTTGTACAGCTCGTCCATG		Construct
HGFP VMA-B P5	CATGGACGAGCTGTACAAGATGGCCGACCCCCGGGTA		Construct
HGFP VMA-B P6	CATGGTGTCGATGGCAGAGAT		Construct
HGFP VMA-B P7	GGTGTGTGATTACGCGGTTA		Genotyping
HGFP VMA-B P8	GCAATTTCATTGTGAGGCAGA		Genotyping

Duet Rav2-1 P1	ACCACAGCCAGGATCCGATGTCTGTCGAGATCTGGCCACCT	pDUET His-ROGDI	Construct
Duet Rav2-1 P2	TGCAGGCGCGCCGAGCTCGAATTCCTAATCCTCCAGCTCCTCTCCCA		Construct
pET Upstream	ATGCGTCCGGCGTAGA		Sequencing
Rav2 Seq F	GATCCATCGCCTCTTCACAT		Sequencing
DuetDOWN1	GATTATGCGGCCGTGTACAA		Sequencing

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
Duet Rav1 P1	TTCGAGCTCGGCGCGCCTGATGAAAGCTGTCCTGCCCGGAA	pDUET His-Rav1	Construct
Duet Rav1 P2	TTGTCGACCTGCAGGCGCGGTTTAAACGTTAAAACCCAAAGCTATCCAACAGC		Construct
pET Upstream	ATGCGTCCGGCGTAGA		Sequencing
Rav1 Seq F1	GCGCTTCGATTTTCTCTACC		Sequencing
Rav1 Seq F2	TTGTGTCGCGGACAGACT		Sequencing
Rav1 Seq F3	CGATCCGAACTCACAATCTG		Sequencing
Rav1 Seq F4	AAATTGGCGAGAAACTGACG		Sequencing
T7 terminator	GCTAGTTATTGCTCAGCGG		Sequencing
Duet Rav2-1 P1	ACCACAGCCAGGATCCGATGTCTGTCGAGATCTGGCCACCT		Construct
Duet Rav2-1 P2	TGCAGGCGCGCCGAGCTCGAATTCCTAATCCTCCAGCTCCTCTCCCA		Construct
Duet Rav2-1 P3	TGGGAGAGGAGCTGGAGGATTAGGAATTCGAGCTCGGCGCGCCTGCA		Construct

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
Duet Rav2-1 P4	TTCCGGGCAGGACAGCTTTCATGTGGCCGGCCGATATCCAATTGAGATCTG		Construct
Duet Rav2-1 P5	CAGATCTCAATTGGATATCGGCCGGCCACATGAAAGCTGTCCTGCCCGGAA		Construct
Duet Rav2-1 P6	TTACCAGACTCGAGGGTACCAAACCCAAAGCTATCCAACAGCGA		Construct
pET Upstream	ATGCGTCCGGCGTAGA		Sequencing
Rav2 Seq F	GATCCATCGCCTCTTCACAT	pDUET His-Rav2 Rav1-S	Sequencing
DuetDOWN1	GATTATGCGGCCGTGTACAA		Sequencing
DuetUP2	TTGTACACGGCCGCATAATC		Sequencing
Rav1 Seq F1	GCGCTTCGATTTTCTCTACC		Sequencing
Rav1 Seq F2	TTGTGTCGCGGACAGACT		Sequencing
Rav1 Seq F3	CGATCCGAACTCACAATCTG		Sequencing
Rav1 Seq F4	AAATTGGCGAGAAACTGACG		Sequencing
T7 terminator	GCTAGTTATTGCTCAGCGG		Sequencing

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
PMA1 Hmch P1	TGGAGATCTTCCTCGGTCTC	PMA1 Hyg mCherry	Construct
PMA1 Hmch P2	GGTGAGTTCAGGCTTTTTCATTTGCGACTTCTCATGCTGAGTA		Construct
PMA1 Hmch P3	TACTCAGCATGAGAAGTCGCAAATGAAAAAGCCTGAACTCACC		Construct
PMA1 Hmch P4	TGCACCCGGAAAATCATTATATCGCTTACTTGTACAGCTCGTCCATGC		Construct
PMA1 Hmch P5	GCATGGACGAGCTGTACAAGTAAGCGATATAATGATTTTCCGGGTGCA		Construct
PMA1 Hmch P6	TTGCTCCCGCGATAGGTGTTG		Construct
PMA1 Hmch P7	CTTCTCGCCAGATTTTCCAC		Genotyping
PMA1 Hmch P8	GGCCAGCCAAAAGTAAGTGT		Genotyping
Spz1 Del1 P1	ACAGGGCAGGTGAAGAGAGA		Construct
Spz1 Del1 P2	GGTGAGTTCAGGCTTTTTCATAATGGTGGCTGAGGCTAGG		Construct
Spz1 Del1 P3	CCTAGCCTCAGCCACCATTATGAAAAAGCCTGAACTCACC		Construct
Spz1 Del1 P4	GATGGCTTTCAATTCGTCCTCCTTGTACAGCTCGTCCATGC		Construct

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
Spz1 Del1 P5	GCATGGACGAGCTGTACAAGGAGGACGAATTGAAAGCCATC		Construct
Spz1 Del1 P6	ATCGCATCCTGTGTGTTCAG	Spz1 Del1 (Del 1-492)	Construct
Spz1 Del1 P7	AGAGCGCGACTGCTACTGG		Genotyping
Spz1 Del1 P8	GTGAAGCTCTTCCTGCAACC		Genotyping
SK17	GAGCTGCATCAGGTCGGAGAC		Genotyping
Sk17 com	GTCTCCGACCTGATGCAGCTC		Genotyping
Spz1 Del1 WT P8	CCTGTCTGAGCCGATGAGAT		Genotyping
Spz1 Del2 P1	TACAACTTCAAGGCCCCTGT		Construct
Spz1 Del2 P2	GGTGAGTTCAGGCTTTTTCATCATTTCGTCGATCTGATACTG		Construct
Spz1 Del2 P3	CAGTATCAGATCGACGAAATGATGAAAAAGCCTGAACTCACC		Construct
Spz1 Del2 P4	TGTTGAACGTGAAAGGACCATCTTGTACAGCTCGTCCATGC	Spz1 Del2 (Del 493-855)	Construct
Spz1 Del2 P5	GCATGGACGAGCTGTACAAGATGGTCCTTTCACGTTCAACA		Construct

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
Spz1 Del2 P6	TATCCTCACGAGCCATGTCA		Construct
Spz1 Del2 P7	TGCGGGTGAACAATAAACTG		Genotyping
Spz1 Del2 P8	CGTCAATGTGTTGCTTTTCG		Genotyping
Spz1 Del3 P1	CCAAGAGGGCTTCGATGATTG	Spz1 Del3 (Del 583-934)	Construct
Spz1 Del3 P2	GGTGAGTTCAGGCTTTTTCATCGACATGCGGGTTCCTGTCT		Construct
Spz1 Del3 P3	AGACAGGAACCCGCATGTCGATGAAAAAGCCTGAACTCACC		Construct
Spz1 Del3 P4	CATCACAAACAAGTTCAGTCTCTTACTTGTACAGCTCGTCCAT		Construct
Spz1 Del3 P5	ATGGACGAGCTGTACAAGTAAGAGACTGAACTTGTTTGTGATG		Construct
Spz1 Del3 P6-5	GGAACTGAATGCCGGGAAGC		Construct
Spz1 Del3 P7	TGCAAGATGAGCGGATTAGG		Genotyping
Spz1 Del2 P6-6	TGCGCTTTTGGATTCCTCACAG		Genotyping

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
Spz2 Del1 P1	ATCCCGTTCGTCGCTTCT	Spz2 Del1 (Del 55-175)	Construct
Spz2 Del1 P2	GGTGAGTTCAGGCTTTTTCATGTGGATCTGGATAGGGTCAT		Construct
Spz2 Del1 P3	ATGACCCTATCCAGATCCACATGAAAAAGCCTGAACTCACC		Construct
Spz2 Del1 P4	GCATTTCTCCTTCTCGCTCCATGCTTGTACAGCTCGTCCATGC		Construct
Spz2 Del1 P5	GCATGGACGAGCTGTACAAGCATGGAGCGAGAAGGAGAAATGC		Construct
Spz2 Del1 P6	CATGGAGCGAGAAGGAGAAATGC		Construct
Spz2 Del1 P7	AGTAGACCGGGCCATCATC		Genotyping
Spz2 Del1 P8	CTCGTCCATGGATCCTTCAG		Genotyping

Spz2 Del2 P1	ATACCAAACCAACGCCAGAC		Construct
Spz2 Del2 P2	GGTGAGTTCAGGCTTTTTCATCGTCAACTGCAAGATGGC		Construct
Spz2 Del2 P3	GCCATCTTGCAGTTGACGATGAAAAAGCCTGAACTCACC		Construct
Spz2 Del2 P4	AGCTCGGCATCCTTTTCTGACTTGTACAGCTCGTCCATGC	Spz2 Del2 (Del	Construct

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
Spz2 Del2 P5	GCATGGACGAGCTGTACAAGTCAGAAAAGGATGCCGAGCT	176-582)	Construct
Spz2 Del2 P6	ATTCCTCCTTGAGCCTGCTA		Construct
Spz2 Del2 P7	CTGTCTGTTTTTGCGTCTTCG		Genotyping
Spz2 Del2 P8	ATCCGCCAGCAGCACTAC		Genotyping

Spz2 Del3 P1	CCGAGCAAAGACTGGAGATG	Spz2 Del3 (Del 583-934)	Construct
Spz2 Del3 P2	GGTGAGTTCAGGCTTTTTCATCCCGCTCTTGTTGTTGAGCT		Construct
Spz2 Del3 P3	AGCTCAACAACAAGAGCGGGATGAAAAAGCCTGAACTCACC		Construct
Spz2 Del3 P4	ATTCCAGAGGGCCCGGTCATTACTTGTACAGCTCGTCCA		Construct
Spz2 Del3 P5	TGGACGAGCTGTACAAGTAATGACCGGGCCCTCTGGAAT		Construct
Spz2 Del3 P6	TGAGCAACGAAACACCCACTC		Construct
Spz2 Del3 P7	GGACAGCAGGAACTCTCAGC		Genotyping
Spz2 Del3 P8	TGGATCTCCCTCCCCTCAAC		Genotyping

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
Spa2 Del1 P1	ATACCCTACCCGGCGCTACG	Spa2 Del1 (Del 1-330)	Construct
Spa2 Del1 P2	GGTGAGTTCAGGCTTTTTCATGGTATTCGTTCACCGCCCGCTT		Construct
Spa2 Del1 P3	AAGCGGGCGGTGAACGAATACCATGAAAAAGCCTGAACTCACC		Construct
Spa2 Del1 P4	ACCTCGAACGCCTGGCCCATCTTGTACAGCTCGTCCATGC		Construct
Spa2 Del1 P5	GCATGGACGAGCTGTACAAGATGGGCCAGGCGTTCGAGGT		Construct
Spa2 Del1 P6	CACAAGCCCAGAGTCTTCGACA		Construct
Spa2 Del1 P7	GCCATCAACCGTCTGGAT		Genotyping
Spa2 Del1 P8	TCGTCTATCGCAATCTGGAA		Genotyping
Spa2 Del2 P1	GTTGGTTCTGTCCGACCACT		Construct
Spa2 Del2 P2	GGTGAGTTCAGGCTTTTTCATAGACCTCCTGCCATTGCCAT		Construct
Spa2 Del2 P3	ATGGCAATGGCAGGAGGTCTATGAAAAAGCCTGAACTCACC		Construct

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
Spa2 Del2 P4	TCGAGTGGTACACCCTCTGACTTGTACAGCTCGTCCATGC	Spa2 Del2 (Del 331-520)	Construct
Spa2 Del2 P5	GCATGGACGAGCTGTACAAGTCAGAGGGTGTACCACTCGA		Construct
Spa2 Del2 P6	ATGACCGAGGTCTTGTCCTG		Construct
Spa2 Del2 P7	GGCATGTATGCGAGGAGTGT		Genotyping
Spa2 Del2 P8	CGCTCAGTTGGTTGATGTTG		Genotyping
Spa2 Del3 P1	AGCAACACAATCGTCCCAAAC	Spa2 Del3 (Del 521-888)	Construct
Spa2 Del3 P2	GGTGAGTTCAGGCTTTTTCATCGAATCACGCATGCTGCGTA		Construct
Spa2 Del3 P3	TACGCAGCATGCGTGATTCGATGAAAAAGCCTGAACTCACC		Construct
Spa2 Del3 P4	TGTCTTGTCTTATGTTGTTTTGGTATTACTTGTACAGCTCGTCCAT		Construct
Spa2 Del3 P5	ATGGACGAGCTGTACAAGTAATACCAAAACAACATAAGACAAGACA		Construct
Spa2 Del3 P6	CGCACATACATCCAGGAACA		Construct
Spa2 Del3 P7	ATGGTGGCTACGGCGTTC		Genotyping

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
Spa2 Del3 P8	ATGCACACGATGCGAGAAC		Genotyping
Spa2 Del2 P1	GTTGGTTCTGTCCGACCACT	Spa2 Del2-1 (Del 331-428)	Construct
Spa2 Del2 P2	GGTGAGTTCAGGCTTTTTCATAGACCTCCTGCCATTGCCAT		Construct
Spa2 Del2 P3	ATGGCAATGGCAGGAGGTCTATGAAAAAGCCTGAACTCACC		Construct
Spa2 Del2-1 P4	TTGCCGCCTGATTGTCGACCCTTGTACAGCTCGTCCATGC		Construct
Spa2 Del2-1 P5	GCATGGACGAGCTGTACAAGGGTCGACAATCAGGCGGCAA		Construct
Spa2 Del2-1 P6	TGAGCCGACGGGAGTTACTG		Construct
Spa2 Del2 P7	GGCATGTATGCGAGGAGTGT		Genotyping
Spa2 Del2-1 P8	ACACTCGAGGTCTGGCTTTG		Genotyping
Spa2 Del2-2 P1	AGTTGGAGAGGCGGTATCCT		Construct
Spa2 Del2-2 P2	GGTGAGTTCAGGCTTTTTCATATCCTCGAGTCCTTCCCGAA		Construct

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
Spa2 Del2-2 P3	TTCGGGAAGGACTCGAGGATATGAAAAAGCCTGAACTCACC	Spa2 Del2-2 (Del 429-520)	Construct
Spa2 Del2 P4	TCGAGTGGTACACCCTCTGACTTGTACAGCTCGTCCATGC		Construct
Spa2 Del2 P5	GCATGGACGAGCTGTACAAGTCAGAGGGTGTACCACTCGA		Construct
Spa2 Del2 P6	ATGACCGAGGTCTTGTCCTG		Construct
Spa2 Del2-2 P7	CCTCCTAGGTTTCGCGATCT		Genotyping
Spa2 Del2 P8	CGCTCAGTTGGTTGATGTTG		Genotyping
Hygmcherry Spz2 P1	CTTCCCGTCCCATTGCTCAG	HygmCherry Spz2	
Hygmcherry Spz2 P2	GGTGAGTTCAGGCTTTTTCATGATGATGATGGTGATGATGATATATTAGG		
Hygmcherry Spz2 P3	CCTAATATATCATCATCACCATCATCATCATGAAAAAGCCTGAACTCACC		
Hygmcherry Spz2 P4	CTAGTTGCGGCTCCATTCATCTTGTACAGCTCGTCCATGC		
Hygmcherry Spz2 P5	GCATGGACGAGCTGTACAAGATGAATGGAGCCGCAACTAG		
Hygmcherry Spz2 P6	CGCGAGTTGGTGTAGGTGTA		

Primer name	Nucleotide Sequence	MFT fragment / Plasmid	Purpose
Hygmcherry Spz2 P7	TCGCGTCACTTTATTGGAGA		
Hygmcherry Spz2 P8	GAATCCATCATCACCACCAA		

Spa2 Del1 P1	ATACCCTACCCGGCGCTACG		Construct
Spa2 Del1 P2	GGTGAGTTCAGGCTTTTTCATGGTATTCGTTCACCGCCCGCTT		Construct
Spa2 Del1 P3	AAGCGGGCGGTGAACGAATACCATGAAAAAGCCTGAACTCACC		Construct
Hygmcherry Spa2 P4	CAATGGTGCATTGCGAACATTCATCTTGTACAGCTCGTCCATGC	HygmCherry Spa2	Construct
Hygmcherry Spa2 P5	GCATGGACGAGCTGTACAAGATGAATGTTCGCAATGCACCATTG		Construct
Hygmcherry Spa2 P6	AGGGCGTCTGTGACTTTGAC		Construct
Spa2 Del1 P7	GCCATCAACCGTCTGGAT		Genotyping
Hygmcherry Spa2 P8	CATTCATCATAGGGGGCATC		Genotyping