

**Evolutionary history of clathrin-mediated
endocytosis and the eisosome**

Submitted by Luigi Cibrario

to the University of Exeter as a thesis for the degree of

Doctor of Philosophy in Biological Sciences

August 2011

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature:

Abstract

Endocytosis is both an ancient and a diverse feature of the eukaryotic cell. Studying how it evolved can provide insight into the nature of the last common eukaryotic ancestor, and the diversification of eukaryotes into the known extant lineages. In this thesis, I present two studies on the evolution of endocytosis. In the first part of the thesis I report results from a large-scale, phylogenetic and comparative genomic study of clathrin-mediated endocytosis (CME). The CME pathway has been studied to a great level of detail in yeast to mammal model organisms. Several protein families have now been identified as part of the complex set of protein-protein and protein-lipid interactions which mediate endocytosis. To investigate how such complexity evolved, first, I defined the modular nature of the CME interactome (CME-I) by literature review, and then I carried out a systematic phylogenetic and protein domain architecture analysis of the proteins involved. These data were used to construct a model of the evolution of the CME-I network, and to map the expansion of the network's complexity to the eukaryotic tree of life. In the second part of the thesis, I present results from evolutionary and functional studies of the eisosome, a protein complex which has been proposed to regulate the spatial distribution of endocytosis in *S. cerevisiae*. The phylogeny of eisosomes components Pil1 and Lsp1 reported here, suggests that eisosomes are likely to have originated at the base of the fungi, and then diversified significantly via multiple gene duplications. I thus studied the localisation and function of Pil1 and Lsp1 homologues in *Magnaporthe oryzae* to investigate the role of eisosomes in filamentous fungi. Results suggests that eisosomes are linked with septal formation and integrity in *M. oryzae*, and that the septal specific Pil2 paralogue was lost in budding yeasts. Together, the data presented in this thesis describe the evolutionary history of a complex biological system, but also highlights the problem of asymmetry in the understanding of endocytic diversity in the eukaryotes.

Table of Contents

Abstract	2
List of Figures	10
List of Tables	14
Acknowledgements	15
Abbreviations	16
1 General Introduction	17
1.1 Endocytosis is a hallmark of the eukaryotic cell	17
1.1.1 Endocytosis in eukaryogenesis.....	18
1.1.2 Endocytosis in the last common eukaryotic ancestor.....	25
1.1.3 Evolutionary history of endocytosis.....	27
1.2 Eukaryote diversification.....	28
1.2.1 Major eukaryote groups and rooting of the eukaryote tree.....	29
1.2.2 Outline of a consensus eukaryote tree of life.....	36
1.3 Endocytosis comprises multiple distinct pathway.....	39
1.3.1 Phagocytosis.....	39
1.3.2 Pinocytosis.....	42
1.3.2.1 Macropinocytosis.....	42
1.3.2.2 Caveolin-mediated endocytosis.....	43
1.3.2.3 Clathrin-mediated endocytosis.....	45
1.3.2.4 Clathrin independent endocytosis.....	46
1.3.2.5 Evolution of pinocytic diversity and functional overlaps among distinct pathways.....	48
1.3.3 Spatial regulation of endocytosis.....	51

1.3.3.1 Eisosomes.....	51
1.4 Thesis aims.....	53
2 Materials and methods.....	57
2.1 Bioinformatics	57
2.1.1 Genomic and proteomic data sampling.....	57
2.1.2 Sequence similarity searches.....	57
2.1.3 Conserved protein domain analysis.....	62
2.1.4 Preparing data sets for phylogenetic analysis.....	63
2.1.4.1 Renaming sequences.....	63
2.1.4.2 Multiple sequence alignment.....	64
2.1.4.3 Masking multiple sequence alignments.....	65
2.1.5 Phylogenetic analysis.....	65
2.1.5.1 Evolutionary model selection.....	65
2.1.5.2 Fast maximum likelihood phylogenetic inference.....	67
2.1.5.3 Bootstrap analysis.....	69
2.1.5.4 Bayesian phylogenetic analysis.....	70
2.2 Molecular biology.....	71
2.2.1 Laboratory methods.....	71
2.2.2 Growth and maintenance of fungal cultures.....	72
2.2.3 Nucleic acid extraction.....	73
2.2.3.1 DNA extraction.....	73
2.2.3.2 RNA extraction.....	74
2.2.3.2.1 Extraction of <i>Magnaporthe oryzae</i> total RNA.....	74
2.2.3.2.2 Extraction of <i>Saccharomyces cerevisiae</i> total RNA.....	75

2.2.4 DNA manipulations.....	76
2.2.4.1 DNA digestion with restriction enzymes.....	76
2.2.4.2 Ligation of DNA fragments.....	77
2.2.4.3 DNA gel electrophoresis	77
2.2.4.4 Gel purification of DNA fragments.....	78
2.2.5 Cloning of PCR product.....	78
2.2.6 Transformation of bacterial hosts.....	80
2.2.7 Plasmid DNA preparation.....	80
2.2.8 RNA manipulations	82
2.2.8.1 RNA gel electrophoresis.....	82
2.2.8.2 Reverse-transcription PCR.....	82
2.2.9 Treatment and reverse transcription of total RNA for rapid amplification of 5' and 3' cDNA ends.....	83
2.2.9.1 Dephosphorylation of non-mRNA and truncated mRNA from <i>M.</i> <i>oryzae</i> total RNA	84
2.2.9.2 RNA precipitation.....	84
2.2.9.3 Removal of cap structures from full length mRNA	85
2.2.9.4 Ligation of the RNA oligonucleotide to decapped full length mRNA..	85
2.2.9.5 Reverse-transcription of full length mRNA	85
2.2.10 DNA sequencing.....	86
2.2.11 Fungal transformation.....	86
2.2.11.1 Transformation of <i>Magnaporthe oryzae</i>	86
2.2.11.2 Transformation of <i>Saccharomyces cerevisiae</i>	88
3 Modular definition of the CME interactome network.....	90
3.1 Introduction.....	90

3.2	Materials and methods.....	95
3.2.1	Definition of the CME-I network.....	95
3.3	Results	96
3.3.1	Description of the CME-I network	101
3.3.2	The core module of the CME-I network.....	101
3.3.3	The membrane bending module	104
3.3.4	The vesicle scission module.....	107
3.3.5	The actin attachment module	107
3.3.6	The vesicle uncoating module	111
3.4	Discussion.....	112
4	Evolutionary history of the CME-I network.....	117
4.1	Introduction.....	117
4.2	Materials and methods	119
4.2.1	Identification of candidate CME-I protein homologues.....	119
4.2.2	Phylogenetic analyses of CME-I network protein families	121
4.2.3	Mapping CME-I network evolution to the eukaryotic tree of life.....	122
4.3	Results.....	123
4.3.1	Taxonomic distribution of CME-I proteins.....	123
4.3.1.1	The conserved core CME-I protein network	125
4.3.1.2	Epsins, dynamins, SNX9, synaptojanins, auxilin and EPS15/EPS15R/intersectin proteins are conserved in diverse eukaryotes but have complex evolutionary histories.....	135
4.3.1.3	AP180/CALM, ABP1 and N-WASP are eukaryote ancestral with secondary loss.....	135
4.3.1.4	HIP1/HIP1R, amphiphysins, endophilin and FCH protein are unikont-specific.....	144

4.3.1.5 Taxonomic distribution of tuba, β -arrestins and PTB proteins.....	147
4.3.2 Evolution of the CME functional repertoire.....	150
4.3.2.1 Mapping protein domain architectures of the CME-I network to the eukaryotic tree of life.....	153
4.3.2.2 Evolution of protein domain architectures in epsins, SNX9, dynamins and synaptojanins.....	154
4.3.2.3 Evolution of the EPS15/EPS15R and intersectin proteins domain architectures.....	157
4.3.2.4 Evolution of N-WASP and auxilin protein domain architectures.....	158
4.3.2.5 Mapping expansion of CME functional repertoire to the eukaryotic tree of life.....	160
4.3.3 Reconstructing expansions in complexity of the CME-I network.....	161
4.3.3.1 Model for studying changes in complexity of the CME-I network.....	161
4.3.3.2 The CME-I network in the LCEA.....	164
4.3.3.3 Expansion of network complexity in the LCUA CME-I.....	164
4.3.3.4 Expansion of network complexity in the LCOA CME-I.....	166
4.3.3.5 Expansion of network complexity in the LCHA CME-I.....	167
4.4 Discussion.....	169
5 Comparative genomic and phylogenetic study of eisosomes.....	177
5.1 Introduction.....	177
5.2 Materials and methods.....	181
5.2.1 Identifying Pil1 and Lsp1 homologues.....	181
5.2.2 Cloning and sequencing putative Pil1 homologue from a <i>Blastocladiella emersonii</i> cDNA library.....	182
5.2.3 Cloning and sequencing putative Pil1 homologue from <i>Capsaspora owczarzaki</i> cDNA library.....	183
5.2.4 Phylogenetic analyses of Pil1 and Lsp1.....	183

5.2.5 Secondary structure prediction of Pil1 and Lsp1 homologues	184
5.3 Results.....	185
5.3.1 Taxonomic distribution of Pil1 and Lsp1.....	185
5.3.2 Alignment and phylogenetic analysis of Pil1 and Lsp1.....	187
5.3.3 Secondary structure prediction of Pil1 and Lsp1 homologues.....	192
5.4 Discussion.....	193
6 Functional characterisation of <i>Magnaporthe oryzae</i> MoPil1 and MoPil2.....	197
6.1 Introduction.....	197
6.2 Materials and methods.....	202
6.2.1 Generating <i>M. oryzae</i> strains expressing MoPil1-GFP and MoPil2-RFP...202	
6.2.2 Microscopy.....	204
6.2.3 Cloning the MGG_00153 coding sequence from <i>M. oryzae</i> cDNA.....	205
6.2.4 Rapid amplification of 5' and 3' MGG_11731 cDNA ends.....	206
6.2.5 Cloning MGG_11731 coding sequence from <i>M. oryzae</i> cDNA.....	207
6.2.6 Building <i>Saccharomyces cerevisiae</i> transformation vectors.....	208
6.2.7 Phenotypic study of transformed <i>S. cerevisiae</i> strains	208
6.3 Results.....	209
6.3.1 Generating a <i>M. oryzae</i> strain expressing MoPil1-GFP.....	209
6.3.2 Sub-cellular localisation of MoPil1-GFP and MoPil2-RFP in <i>M. oryzae</i> conidia and hyphae	212
6.3.3 Functional complementation of <i>S. cerevisiae pill1</i> Δ mutant by <i>M. oryzae</i> MoPil1 and MoPil2.....	213
6.3.3.1 Differential stress resistance in <i>S. cerevisiae</i> wild type and <i>pill1</i> Δ strains.....	213
6.3.3.2 Construction of pYES2:MoPil1 and pYES2:MoPil2 transformation vectors.....	218

6.3.3.3 Expression of MoPil1 and MoPil2 in transformed <i>S. cerevisiae</i> cells	220
6.3.3.4 Analysis of <i>S. cerevisiae pill</i> Δ :MoPil1 and <i>pill</i> Δ :MoPil2 transformants.....	222
6.4 Discussion.....	224
7 General discussion.....	229
Appendix 1: Laboratory products and software suppliers used in this thesis.....	244
Appendix 2: Taxonomic distribution of protein domain architectures in CME-I	246
Appendix 3: Putative origin of protein-protein and protein-lipid interactions in CME-I.....	247
Appendix 4: Pil1 and Lsp1 phylogeny with <i>C. owczarzaki</i> NUL00001676.....	248
Appendix 5: Heat shock assay of transformed <i>S. cerevisiae pill</i> Δ mutants.....	249
Bibliography.....	250

List of Figures

Chapter 1

Figure 1.1 Syntrophic and autogenous scenarios of eukarote origin.....	20
Figure 1.2 The diversity of eukaryotes: from 4 kingdoms to 6 'supergroups'.....	35
Figure 1.3 Alternative hypotheses for the rooting of the eukaryotic tree of life.....	38
Figure 1.4 Endocytosis comprises mechanistically distinct pathways.....	40

Chapter 3

Figure 3.1 Novel interaction between two network components as a synapomorphy.....	93
Figure 3.2 Connectivity diagram depicting the CME-I network.....	102
Figure 3.3 Core module of the CME-I network.....	105
Figure 3.4 Membrane bending module of the CME-I network.....	106
Figure 3.5 Vesicle scission module of the CME-I network.....	108
Figure 3.6 Actin attachment module of the CME-I network.....	110
Figure 3.7 Vesicle uncoating module of the CME-I network.....	113

Chapter 4

Figure 4.1 Distinction of inparalogues from outparalogues when assessing homology to query protein.....	125
Figure 4.2 Taxonomic distribution of CME-I network proteins.....	127
Figure 4.3 Phylogenetic tree of clathrin heavy chain.....	129
Figure 4.4 Phylogenetic tree of clathrin light chain.....	131
Figure 4.5 Phylogenetic tree of AP2 α subunit.....	130
Figure 4.6 Phylogenetic tree of AP2 β subunits.....	132

Figure 4.7 Phylogenetic tree of AP2 μ subunits.....	133
Figure 4.8 Phylogenetic tree of AP2 σ subunits.....	134
Figure 4.9 Phylogenetic tree of epsins.....	138
Figure 4.10 Phylogenetic tree of synaptojanins.....	139
Figure 4.11 Phylogenetic tree of dynamins.....	140
Figure 4.12 Phylogenetic tree of SNX9.....	141
Figure 4.13 Phylogenetic tree of auxilin.....	142
Figure 4.14 Phylogenetic tree of EPS15, EPS15R and intersectins.....	144
Figure 4.15 Phylogenetic tree of ABP1.....	145
Figure 4.16 Phylogenetic tree of N-WASP.....	145
Figure 4.17 Phylogenetic tree of AP180 and CALM.....	146
Figure 4.18 Phylogenetic tree of HIP1 and HIPR.....	148
Figure 4.19 Phylogenetic tree of amphiphysins and endophilin.....	148
Figure 4.20 Phylogenetic tree of TOCA-1, FBP17 and CIP4	149
Figure 4.21 Phylogenetic tree of tuba.....	151
Figure 4.22 Phylogenetic tree of β arrestins.....	151
Figure 4.23 Phylogenetic tree of numb, disabled 2 and ARH.....	152
Figure 4.24 Putative origin of the CME protein domain repertoire.....	158
Figure 4.25 Connectivity diagram depicting the putative CME-I network topology in the last common eukaryotic ancestor	165
Figure 4.26 Connectivity diagram depicting the putative CME-I network topology in the last common unikont ancestor.....	166
Figure 4.27 Connectivity diagram depicting the putative CME-I network topology in the last common opisthokont ancestor	167
Figure 4.28 Connectivity diagram depicting the putative CME-I network topology in the last common holozoan ancestor.....	168

Chapter 5

Figure 5.1 Eisosomes assemble into punctate complexes at the plasma membrane of <i>S. cerevisiae</i> cells.....	178
Figure 5.2 Schematic model of the eisosome/MCC system.....	179
Figure 5.3 Taxonomic distribution of Pil1 and Lsp1 in Fungi and opisthokont protists.....	188
Figure 5.4 Phylogenetic tree of Pil1 and Lsp1.....	190
Figure 5.5 Schematic tree of Pil1 and Lsp1 evolutionary history.....	191
Figure 5.6 Secondary structure prediction of Pil1 and Lsp1 homologues.....	194

Chapter 6

Figure 6.1 The infective life cycle of <i>Magnaporthe oryzae</i>	199
Figure 6.2 Defining Pil1 and Pil2 paralogues in ascomycete fungi.....	201
Figure 6.3 Cloning strategy to generate MoPil1-GFP gene fusion	210
Figure 6.4 Generating a MoPil1-GFP gene fusion construct.....	211
Figure 6.5 Sub-cellular localisation of MoPil1-GFP at different stages of <i>M. oryzae</i> conidia germination and appressorium formation.....	214
Figure 6.6 Sub-cellular localisation of MoPil1-GFP in <i>M. oryzae</i> vegetative hyphae	215
Figure 6.7 Sub-cellular localisation of MoPil2-RFP in <i>M. oryzae</i> conidia	216
Figure 6.8 MoPil1-GFP and MoPil2-RFP do not colocalise in mature appressoria of <i>M. oryzae</i>	216
Figure 6.9 MoPil1-GFP and MoPil2-RFP co-localise in vegetative hyphae of <i>M. oryzae</i>	217
Figure 6.10 Stress resistance to Calcofluor White, Congo Red and SDS in <i>S. cerevisiae</i> wild type and <i>pil1</i> Δ strains.....	219
Figure 6.11 Restriction Enzyme digestion of pYES2:MoPil1.....	221
Figure 6.12 Cloning the MoPil2 coding sequence	221

Figure 6.13 Expression of MoPil1 and MoPil2 in <i>S. cerevisiae</i> pil1 Δ mutants	223
Figure 6.14 Functional complementation of <i>S. cerevisiae</i> pil1 Δ mutant by <i>M. oryzae</i> MoPil1 and MoPil2	225
Chapter 7	
Figure 7.1 Autogenous scenario of endomembrane origin based on RasGTPase evolution.	231
Figure 7.2 Evolution of actin modulation system in the last common unikont ancestor.....	233
Figure 7.3 Asymmetry in the knowledge of CME diversification across eukarotes.....	240
Appendix 2	
Taxonomic distribution of protein domains architectures in CME-I network proteins across 15 monophyletic eukaryote groups.....	246
Appendix 3	
Putative origins of protein-protein and protein-lipid interaction in the CME-I network.....	247
Appendix 4	
Phylogeny of Pil1 and Lsp1 with long branch sequence <i>C. owczarzaki</i> . NUL00001676	248
Appendix 5	
Heat shock assay of transformed <i>S. cerevisiae</i> strains	249

List of Tables

Chapter 2

Table 2.1 Predicted proteome and translated nucleotide databases of diverse eukaryotes used for phylogenomic studies in this thesis59

Table 2.2 Additional sampling of predicted proteome and translated nucleotide databases for the evolutionary study of the eisosome60

Table 2.3 EST libraries used in Pil1 and Lsp1 sequence similarity searches.....60

Chapter 3

Table 3.1 CME-I network protein97

Table 3.2 Table 3.2 List of CME proteins which also play roles in non-CME endocytic pathways and/or other cellular functions.115

Appendix 1

Laboratory products and software suppliers used in this thesis 244