An ancestral role in *de novo* peroxisome assembly is retained by the divisional peroxin Pex11 in the yeast *Yarrowia lipolytica*

by

Mary Jessica Klute

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Cell Biology
University of Alberta

Abstract

This thesis is on peroxisome evolution and reports my findings with regard to two major aspects of this process. First, I investigated the evolution of a peroxisomal protein family for which large expansions had been noted in diverse eukaryotes from a comparative genomic perspective and combined these findings with functional data delving into the evolutionary history of this protein family. Second, I studied the peroxisomal protein complement of three newly available eukaryotic genomes.

The peroxin (protein required for peroxisome biogenesis) Pex11p has a well recognized role in peroxisome division in diverse eukaryotes. Pex11p remodels and elongates the membranes of peroxisomes prior to the recruitment of dynamin-related GTPases that act in membrane scission to physically divide peroxisomes. I performed a comprehensive comparative genomics survey to understand the significance of the evolution of the Pex11 protein family in yeast and other eukaryotes. Pex11p itself is highly conserved and ancestral, and has undergone numerous lineage-specific duplications, while other Pex11 protein family members are fungal-specific innovations. Functional characterization of the *in silico* predicted Pex11 protein family members of the yeast *Yarrowia lipolytica*, *i.e.* Pex11p, Pex11Cp and Pex11/25p, demonstrated that Pex11Cp and Pex11/25p function in the regulation of peroxisome size and number characteristic of Pex11 protein family members. Unexpectedly, deletion of *PEX11* in *Y. lipolytica* produces cells that lack morphologically identifiable peroxisomes, mislocalize peroxisomal matrix proteins, and show preferential degradation of peroxisomal membrane proteins, *i.e.* they exhibit the classical pex mutant phenotype, which has not been reported for any other eukaryotic cell deleted for the *PEX11* gene. My results demonstrate an unprecedented role for Pex11p in *de novo* peroxisome assembly.

I also studied the peroxisomal protein complement of three medically and/or evolutionarily relevant eukaryotic genomes: the intestinal parasite *Blastocystis hominis*, *Bodo saltans*, a close relative of the parasitic trypanosomatids which contains a divergent peroxisome known as the glycosome,

and *Naegleria fowleri*, better known as the 'brain-eating amoeba'. I showed that *Blastocystis* lacks peroxisomal proteins, consistent with parasites generally lacking the peroxisome organelle. I demonstrated that *B. saltans* and *N. fowleri* encode a relatively complete *PEX* gene complement, with certain losses. I wrote a program, PTS Finder, to predict proteins targeted to the peroxisome, and used it during these analyses.

Preface

Chapters 1 and 3 of this thesis have been submitted for publication as:

Klute, M.J., Chang, J., Tower, R.J., Mast, F.D., Dacks, J.B., and R.A. Rachubinski. 2014. An ancestral role in *de novo* peroxisome assembly is retained by the divisional peroxin Pex11 in the yeast *Yarrowia lipolytica*. *J. Cell Sci.*

M.J. Klute, J. Chang, R.J. Tower, and F.D. Mast performed the experiments. M.J. Klute, J. Chang, R.J. Tower, F.D. Mast, J.B. Dacks, and R.A. Rachubinski provided a conceptual framework for the study, interpreted data and wrote the manuscript.

Dedication

For Mom and Dad.

Acknowledgements

Firstly, I would like to thank my co-supervisors, Dr. Richard Rachubinski and Dr. Joel Dacks. I appreciate all of the opportunities that you have given me. Thank-you for allowing me to share in your passion for basic science research.

Rick, it has been a privilege to be a part of your lab for the past three years. I feel like I am leaving my graduate program as a much more confident and opinionated scientist and person. Thank-you for allowing me the freedom to explore various aspects of this project.

Joel, I am truly grateful for your guidance not only throughout my graduate program, but also throughout my three summer research projects, three undergraduate research courses, and undergraduate degree. Your constant encouragement and words of faith in my abilities have been very much appreciated. I would not be where I am today without you.

I would also like to thank Dr. Kinga Kowalewska-Grochowska for serving on my supervisory committee, and Dr. Andrew Simmonds and Dr. Paul LaPointe for serving on my thesis defense committee.

Words cannot express how grateful I am for the friendship of Emily Herman. I consider myself very lucky to have worked alongside you throughout our undergraduate and graduate studies. Thank-you very much for all of your advice, whether it be science-related or not.

Fred Mast, I truly appreciate your involvement in and support throughout my graduate program, and everything you have taught me. Thank-you for always being willing to take the time to answer my questions, assist with my experiments, and provide helpful discussion and insights. Thank-you for being such an incredible mentor in both my academic and personal life.

I would like to thank all of the members of the Rachubinski lab, including Hiren Banerjee, Jenny Chang, Barbara Knoblach, Hanna Kroliczak, David Lancaster, Sandrine Lepine, Fred Mast, Rick Poirier, Elena Savidov and Rob Tower. Thank-you for your assistance with experimental troubleshooting and for keeping the lab running smoothly. In particular, I would like to thank Rob Tower for initial training and proposing many of the intial experiments for my graduate project, as

well as my benchmate, Jenny Chang, for answering many questions and being a great resource for everything *Yarrowia*-related.

I would also like to thank the members the Dacks lab, including Maria Aguilar González, Lael Barlow, Emily Herman, Chris Klinger, Alex Schlacht and Jeremy Wideman. Your enthusiam for science is both contagious and inspiring. It has been a pleasure working with all of you.

I would like to acknowledge the Department of Cell Biology, as well as the financial support provided by the Faculty of Medicine and Dentistry, the Faculty of Graduate Studies and Research, Alberta Health Services and the Government of Alberta.

My final thanks goes to my family: Mom, Dad and Stephen. I truly appreciate your support of all of my academic endeavours and your involvement in everything I do.

Technical Acknowledgements

The genome of *Bodo saltans* was analyzed with Dr. Fred Mast, Seattle Biomedical Research Institute. The PTS Finder program was written in collaboration with Emily Herman, Department of Cell Biology, University of Alberta. Affinity purification of antibodies against Pex2p was performed by Elena Savidov, Department of Cell Biology, University of Alberta.

Table of Contents

List of Tables

List of Figures

List of symbols and abbreviations

ATP	adenosine triphosphate
BLAST	basic local alignment search tool
C-terminal	carboxyl terminal
Da	Dalton
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
DRP	dynamin-related protein
ECL	enhanced chemi-luminescence
EM	electron microscopy
ER	endoplasmic reticulum
g	gram
g	gravitational acceleration
GFP	green fluorescent protein
GTP	guanosine triphosphate
h	hour
HGT	horizontal gene transfer
HRP	horse radish peroxidase
HMM	hidden Markov model
IgG	immunoglobulin G
IRD	infantile refsum disease
kb	kilobase
kDa	kilodalton
LECA	last eukaryotic common ancestor
LRE	light responsible element
M	molar concentration
mC	mCherry
min	minute
mPTS	peroxisomal membrane protein targeting signal
mRFP	monomeric red fluorescent protein
mRNA	messenger ribonucleic acid
NALD	neonatal adrenoleukodystrophy
N-terminal	amino terminal
OD	optical density
OPH	organelle paralogy hypothesis
ORE	oleate response element
ORF	open reading frame
PBD	peroxisome biogenesis disorder
PCR	polymerase chain reaction
PEX	peroxin
PMP	peroxisomal membrane protein
PNS	postnuclear supernatant
PPARα	peroxisome proliferator activated receptor α
PPRE	PPARα-responsive element
PTS	peroxisomal targeting signal
RCDP	rhizomelic chondrodysplasia punctata
ROS	reactive oxygen species
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec	second
SM	simplified media

UAS1	upstream activating sequence 1
UTR	untranslated region
V	volt
v/v	volume to volume ratio
VLCFA	very-long chain fatty acid
w/w	weight to weight ratio
w/v	weight to volume ratio
ZS	Zellweger syndrome

Chapter 1: Introduction

A version of this chapter has been submitted for publication as:
Klute, M.J., Chang, J., Tower, R.J., Mast, F.D., Dacks, J.B., and R.A. Rachubinski. 2014. An ancestral role in *de novo* peroxisome assembly is retained by the divisional peroxin Pex11 in the yeast *Yarrowia lipolytica*. *J. Cell Sci.*

1.1 Introduction to peroxisomes

Both the number and sophistication of internal compartments distinguish eukaryotic cells from their prokaryotic counterparts. One such example is the single-membrane-bound organelle known as the peroxisome. Christian de Duve and colleagues first used the term peroxisome to describe organelles containing a hydrogen peroxide-producing oxidase and catalase, which degrades the hydrogen peroxide into water and molecular oxygen (de Duve et al., 1960; Baudhuin et al., 1965; de Duve and Baudhuin, 1966). Hence, the hallmark metabolic pathways of peroxisomes are the β-oxidation of fatty acids coupled to the controlled decomposition of hydrogen peroxide by catalase. Peroxisomes contain over 50 enzymes, making them highly dynamic organelles that exhibit specialized biochemical functions depending on cell type and that are highly responsive to changing environmental cues. Other metabolic functions include the synthesis of bile acids and plamalogens, as well as purines and pyrimidines, α-oxidation of branched-chain fatty acids, and superoxide radical degradation (Wiese et al., 2007; see (Wanders and Waterham, 2006) for a comprehensive review of peroxisome biochemistry). Mammalian peroxisomes are specialized for different organs by having different enzymatic compositions to help carry out specific roles. For example, liver peroxisomes are specialized for bile acid synthesis (Islinger et al., 2010) as well as ethanol metabolism (Bradford et al., 1993; Orellana et al., 1998; Thurman and McKenna, 1975). Kidney peroxisomes were found to be confined to proximal tubule epithelial cells (Litwin et al., 1988). Peroxisomes are also specialized in brain cells; peroxisomes are found at the highest concentrations during myelination (Houdou et al., 1991). Peroxisomes are also important for oligodendrocyte maintenance of white matter tracts (Kassman et al., 2007).

1.2 *PEX* genes are necessary for peroxisome biogenesis

Thirty-four proteins called peroxins (encoded by *PEX* genes) are responsible for the major aspects of peroxisome biogenesis: *de novo* formation from the endoplasmic reticulum (ER), import of matrix and membrane proteins, and division (Distel et al., 1996). In mammals, short- and medium-chain

fatty acids undergo beta oxidation in the mitochondria, while peroxisomes are responsible for beta-oxidation of very long chain fatty acids (Eaton et al., 1996; Wanders et al., 2001). In contrast, beta-oxidation in yeast occurs exclusively in the peroxisome (Hiltunen et al., 2003). Hence, *PEX* genes were originally identified in *S. cerevisiae* by replica plating mutant strains onto plates containing oleic acid as the sole carbon source (Erdmann et al., 1989). The *PEX* genes encode both cytosolic and peroxisomal membrane proteins whose basic functions are given in Table 1-1. Many of these aspects of peroxisome dynamics are discussed in detail later in this thesis. The essential contribution of the *PEX* genes to maintenance of the peroxisome was an important refinement on de Duve's classical definition of this organelle.

Table 1-1. Phylogenetic origin, function and subcellular localization of the proteins encoded by *PEX* genes[a]

Protein	Domain	Origin	Function	Compartment
Pex1	AAA-ATPase	eukaryotic/archaea	matrix protein import	cytosol
Pex2	Zn-RING	eukaryotic	matrix protein import	membrane
Pex3	-	eukaryotic	membrane assembly	membrane
Pex4	ubiquitin ligase	eukaryotic	matrix protein import	cytosol and membrane
Pex5	TRR region	eukaryotic/archaea	matrix protein import	cytosol and membrane
Pex6	AAA-ATPase	eukaryotic/archaea	matrix protein import	cytosol
Pex7	WD40	eukaryotic/archaea	matrix protein import	cytosol
Pex8	Coiled-coil domain, leu-zipper	eukaryotic	matrix protein import	membrane
Pex10	RING-finger	eukaryotic	matrix protein import	membrane
Pex11	-	eukaryotic	proliferation	membrane
Pex12	Zn-RING	eukaryotic	matrix protein import	membrane
Pex13	SH3	eukaryotic	docking of receptors	membrane
Pex14	phosphorylation	eukaryotic/archaea	docking of receptors	cytosol and membrane
Pex15	-	eukaryotic	membrane assembly	membrane

Pex16	-	eukaryotic	membrane assembly	membrane
Pex17	-	eukaryotic	docking of receptors	membrane
Pex18	-	eukaryotic	PTS targeting	cytosol and membrane
Pex19	CAAX-box, farynesylation	eukaryotic	membrane assembly	cytosol and membrane
Pex20	-	eukaryotic	PTS targeting	cytosol and membrane
Pex21	-	eukaryotic	PTS targeting	cytosol and membrane
Pex22	-	eukaryotic	matrix protein import	membrane
Pex23	dysferlin	eukaryotic	proliferation	membrane
Pex24	-	eukaryotic	membrane assembly	membrane
Pex25	-	eukaryotic	proliferation	membrane
Pex26	-	eukaryotic	protein import and recruitment	membrane
Pex27	-	eukaryotic	proliferation	membrane
Pex28	-	eukaryotic	proliferation	membrane
Pex29	-	eukaryotic	proliferation	membrane
Pex30	dysferlin	eukaryotic	proliferation	membrane
Pex31	dysferlin	eukaryotic	proliferation	membrane
Pex32	dysferlin	eukaryotic	proliferation	membrane
Pex33	-	unknown	docking of receptors	unknown
Pex34	-	unknown	proliferation	membrane

[a]Adapted from Gabaldón et al., 2006; Schlüter et al., 2006; Platta and Erdmann, 2007.

1.3 Peroxisome biogenesis disorders

The detrimental effect of abnormal peroxisome biogenesis on human health is clearly underscored by the peroxisome biogenesis disorders (PBDs), a heterogenous spectrum of autosomal recessive diseases. These disorders cannot be cured; to date, the only clinical interventions are palliative in nature. PBDs include Zellweger syndrome (ZS), rhizomelic chondrodysplasia punctata (RCDP), neonatal adrenoleukodystrophy (NALD) and infantile Refsum disease (IRD), which are classified based on the severity of the clinical phenotype (Waterham and Ebberink, 2012). In addition to these four disorders, there are also a number of single-enzyme deficiencies that share many characteristics with the more severe disorders. Clinical presentation of PBDs includes profound

neurological abnormalities, muscular hypotonia, cataracts, cardiac defects, dysmorphic features, and growth and mental retardation (Bowen et al., 1964; Smith et al., 1965; Wilson et al., 1986, reviewed in Kelley, 1983; Weller et al., 2003; Steinberg et al., 2006). RCDP has a distinct presentation from the other PBDs, including proximal shortening of the long bones (rhizomelia) and epiphyseal calcification/stippling (chondrodysplasia punctata) (Wardinsky et al., 1990). Reflecting the ubiquitous nature of peroxisomes, PBDs typically affect nearly every single organ system. ZS has the most severe clinical presentation, and some children with this condition have been initially thought to have Down syndrome, Prader–Willi syndrome or spinal muscular atrophy due to the severity of muscular hypotonia (Steinberg et al., 2006). ZS is typically fatal within the first year of life (Wilson et al., 1986). Outcomes vary for the other PBDs, with some children surviving into young adulthood. PBDs are considered to be rare genetic conditions. ZS has an estimated incidence of approximately 1:50,000 in North America and approximately 1:500,000 in Japan (Steinberg et al., 2004).

At least 14 *PEX* genes functionally complement all of the peroxisome biogenesis defects that are observed in the PBDs (Table 1-2). These defects encompass defects in peroxisome biogenesis as well as matrix protein import. Mutations in *PEX1* are the most common cause of ZS. Assignment of PBDs to any single complementation group is complicated because different complementation groups share common clinical phenotypes, with any single mutation typically causing functional loss of the organelle (Steinberg et al., 2006).

The poor outcome associated with PDBs, combined with a lack of effective therapeutic interventions, makes genetic testing for carriers in families of affected patients and prenatal genetic screening invaluable (Waterham and Ebberink, 2012). This thesis is relevant to this topic because the first PBD patient with a mutation in the *PEX11* gene was recently reported (Ebberink et al., 2012).

Table 1-2. Complementation groups and *PEX* gene defects in peroxisome biogenesis disorders[a]

Gene	Locus	CG-Dutch[b]	CG-Japan[b]	CG-KKI[b]	Clinical Phenotype(s)[c]	Proportion of ZSS
PEX1	7q21.2	2	E	1	ZS, NALD, IRD	70%
PEX2	8q21.1	5	F	10	ZS, IRD	3%
PEX3	6q24.2	-	G	12	ZS	<1%
PEX5	12p13.31	4	-	2	ZS, NALD	<2%
PEX6	6p21.1	3	C	4, 6	ZS, NALD, IRD	10%
PEX7	6q21-q22.2	1	R	11	RCDP	-
PEX10	1p36.32	-	B	7	ZS, NALD	3%
PEX11β	1q21.1	-	-	-	-	-
PEX12	17q12	-	-	3	ZS, NALD, IRD	5%
PEX13	2p14-p16	-	H	13	ZS, NALD	<1%
PEX14	1p36.22	-	K	-	ZS	<1%
PEX16	11p11.2	-	D	9	ZS	<1%
PEX19	1q22	-	J	14	ZS	<1%
PEX26	22q11.21	-	A	8	ZS, NALD, IRD	5%

[a]Adapted from Steinberg et al., 2006; Waterham and Ebberink, 2012.
[b]CG, complementation group; Dutch, University of Amsterdam; Japan, Gifu University School of Medicine; KKI, Kennedy Krieger Institute.
[c]Estimates of CG frequency derived from KKI data.

1.4 Peroxisome matrix protein import

Peroxisomes do not contain DNA or a protein synthesis machinery, and as a result they must post-translationally import all matrix and membrane proteins. Sorting of matrix proteins to peroxisomes depends on the presence of two major types of peroxisomal targeting signals (PTSs). The PTS1 is a C-terminal tripeptide and is the signal utilized by most peroxisomal proteins (Rucktäschel et al., 2011). Initially discovered as a C-terminal SKL in the firefly luciferase protein (Gould et al., 1989), the PTS1 can now be thought of as the consensus sequence (S/A/C)-(K/R/H)-(L/M) (reviewed in Erdmann, 2013). The influence of amino acids upstream of the tripeptide prompted reassignment of the PTS1 as a 12-amino acid dodecamer (Neuberger et al., 2003a; Brocard and Hartig, 2006). The four amino acids immediately upstream of the C-terminal tripeptide interact with the surface of Pex5p, and there are specific properties of amino acids favored at certain positions for Metazoa and Fungi (Neuberger et al., 2003a). The remaining residues have unstructured requirements, but there is preference towards polar, hydrophilic, and solvent accessible and flexible residues (Neuberger et al., 2003a). The PTS2 is a N-terminal nonapeptide that was originally described for rat thiolase

(Swinkels et al., 1991). The PTS2 fits the consensus sequence is: (RK)-(L/V/I/Q)-XX-(L/V/I/H/Q)-(L/S/G/A/K)-X-(H/Q)-(L/A/F) (reviewed in Erdmann, 2013). PTS2 usage varies considerably between organisms. Only two known *Saccharomyces cerevisiae* proteins contain a PTS2 sequence: 3-ketoacyl thiolase (Pot1p) and NAD+-dependent glycerol-3-phosphate dehydrogenase (Gpd1p). Similarly, only a few proteins in mammals are targeted to peroxisomes using the PTS2. Conversely, about one third of plant peroxisomal proteins have a PTS2 (Reumann et al., 2009).

Strikingly, some peroxisomal matrix proteins are imported despite the fact that they contain neither a PTS1 nor a PTS2 (van der Klei and Veenhuis, 2006). For example, peroxisomal acyl-CoA oxidase in *Y. lipolytica* has five subunits (Aox1-5), all lacking a PTS1 or PTS2 (Titorenko et al., 2002). It is thought that conformational epitopes contain peroxisomal sorting information, dependent upon formation of Aox oligomers (Subramani, 2002; Titorenko et al., 2002). Similarly, *S. cerevisiae* acyl-CoA oxidase also lacks a PTS1 or PTS2, but requires Pex5p for its import (Klein et al., 2002). Other proteins lacking a PTS can piggyback on a PTS-containing protein to be imported into the peroxisome (Glover et al., 1994). When a mutant, PTS2-lacking form of Pot1p was expressed in *S. cerevisiae*, heterodimers formed with endogenous, full-length Pot1p that could be successfully imported into the peroxisome (Glover et al., 1994). It is fascinating that peroxisomes can import fully folded proteins, oligomeric protein complexes and even gold particles 9 nm in diameter (Meinecke et al., 2010). This observation is thought to be due to Pex14p associating with cargo-bound Pex5p and forming a highly dynamic, expandable pore referred to as the peroxisomal importomer (Meinecke et al., 2010).

PTS1- and PTS2-containing proteins are recognized in the cytosol by their respective receptors, Pex5p and Pex7p. Unlike Pex5p, the action of Pex7p is dependent on several different species-specific co-receptors: Pex18p and Pex21p in *S. cerevisiae* (Purdue et al., 1998), Pex20p in *Yarrowia lipolytica* and several other yeast species (Titorenko et al., 1998; Sichting et al., 2003; Otzen et al., 2005; Leon et al., 2007), and a longer splice variant of Pex5p in humans and plants (Otera et al., 2000; Dobt et al., 2001; Hayashi et al., 2005; Woodward et al., 2005). Cargo is recognized in the

cytosol by these receptors, the receptor-cargo complex docks at the peroxisomal membrane, and cargo translocates into the peroxisomal matrix. Receptors are recycled back to the cytosol for subsequent rounds of import via ubiquitination by the RING complex (Pex2p, Pex10p and Pex12p) and the ubiquitin-conjugating enzyme Pex4p (Platta et al., 2007; Williams et al., 2008; Platta et al., 2009). Receptors are exported by the two AAA-ATPases Pex1p and Pex6p (Miyata et al., 2005; Platta et al., 2005).

Variants on the PTS1 consensus sequence listed above have been described in plants. Lingner and colleagues identified a more relaxed PTS1 consensus motif in plants, with residues permitted as follows: (S/A/P/C/F/V/G/T/L/K/I)-(K/R/S/N/L/M/H/G/E/T/F/P/Q/C/Y)-(L/M/I/V/Y/F) (Lingner et al., 2011). Some of these residues were designated as high abundance: (S/A)-(K/R)-(L/M/I). At least two high abundance residues must be combined with one low-abundance residue to constitute a functional PTS1 in plants, *i.e.* X-(K/R)-(LMI), (S/A)-X-(L/M/I) or (S/A)-(K/R)-X (Lingner et al., 2011). Chowdhary and colleagues expanded on this definition by identifying two additional low-abundance residues that could be allowed: Q at position one and D at position two (Chowdhary et al., 2012). Additionally, the authors found that PTS1-containing proteins in plants also contained 4-5 selected amino acids at positions -4 to -12 before the PTS1 tripeptide. These highly diverse residues were predicted to have high targeting enhancing properties, and include Pro, basic residues (Lys, Arg, His), hydroxylated (Ser, Thr), hydrophobic (Ala, Val), and even acidic residues (Asp, Glu) (Chowdhary et al., 2012). There are likely many other examples of proteins in other eukaryotes containing aberrant PTS motifs that are successfully imported into the peroxisomal matrix.

1.5 Peroxisomal membrane protein import

Like peroxisomal matrix proteins, peroxisomal membrane proteins (PMPs) are also post-translationally acquired by the peroxisome. There are 2 general models for this mechanism. In the first model, PMPs are inserted directly into the peroxisomal membrane from the cytosol. PMPs contain a poorly defined hydrophobic sequence known as the membrane peroxisomal targeting

sequence (mPTS) (Rottensteiner et al., 2004; Halbach et al., 2005). The mPTS is the binding site for Pex19p, and in turn, this complex binds Pex3p on the peroxisomal membrane. In a poorly understood process, insertion of the membrane protein in its correct orientation into the peroxisomal membrane then occurs (Fang et al., 2004; Heilard and Erdmann, 2005). In mammalian cells, Pex16p acts as the docking site for the Pex19p/cargo complex in a Pex3p-independent manner (Matsuzaki and Fujiki, 2008). However, the *S. cerevisiae* genome does not encode Pex16p, and Pex16p in *Y. lipolytica* has a different role, acting as a negative regulator of peroxisome division (Eitzen et al., 1997).

In the second model, PMPs are first targeted to the ER before trafficking to peroxisomes. van der Zand and colleagues demonstrated that 16 *S. cerevisiae* PMPs that differ in membrane topology and function (including Pex11p, Pex25p and Pex27p) first traffic to the ER in wild-type cells as well as in *pex3*Δ and *pex19*Δ cells (van der Zand et al., 2010). How PMPs are targeted to the ER remains an open question. Mast and colleagues reviewed mechanisms to achieve peroxisomal membrane protein insertion (Mast et al., 2010). The SRP pathway can mediate cotranslational insertion of PMPs into the ER lumen in a Sec61p-dependent manner (Rapoport, 2007). Most PMPs enter the ER via the Sec61 translocon, which has been shown to be the limiting factor in this process (van der Zand et al., 2010). However, no signal peptides or signal anchor sequences were predicted for PMPs (van der Zand et al., 2010). Cytosolic HSP70/HSP40 chaperones can maintain correct membrane protein conformation until they are directed to the ER translocon (also in a Sec61p-dependent manner), or to the mitochondrial TOM complex (Antonenkov and Hiltunen, 2006). Finally, C-terminally anchored proteins depend on the guided entry of tail-anchored proteins (Get) pathway for entry into the ER lumen (Schuldiner et al., 2008). Pex15p is a known exception of a PMP that uses the Get pathway rather than the Sec61p translocon for its insertion into the ER (Schuldiner et al., 2008). All of these mechanisms are dependent on trafficking pathways between the ER and mitochondria, and peroxisomes. In this

model, the mPTS may act as a sorting signal for PMPs to exit the ER, and Pex3p and Pex19p are required for budding of pre-peroxisomal compartments from the ER (van der Zand et al., 2010).

Presently, it is not known how PMPs are transported from the ER to peroxisomes. Vesicle-mediated transport from the ER to peroxisomes has been described (Titorenko et al., 1998), but this process is not affected by COPI and COPII inhibitors (Salomons et al., 1997; Voorn-Brouwer et al., 2001). Components of the ER-associated secretory pathway (Sec20p, Sec39p, and Dsl1p) have been identified as being involved in peroxisome biogenesis (Perry et al., 2009).

1.6 *De novo* synthesis from the ER

S. cerevisiae cells lacking Pex3p or Pex19p, as well as mammalians cells lacking either of these two proteins or Pex16p, lack peroxisomal remnants. This phenotype is not universal to all eukaryotes, *e.g. Y. lipolytica pex19*Δ cells still contain morphologically identifiable peroxisomes that are defective in matrix protein import (Lambkin and Rachubinski, 2001). However, upon reintroduction of the *PEX* gene, peroxisomes can form *de novo*. Since peroxisomes cannot form from the fission of pre-existing peroxisomes in this case, it is now well established that they arise from another membrane. There are several lines of evidence supporting the endoplasmic reticulum (ER) as this membrane source. Hoepfner and colleagues observed that when *PEX3* was reintroduced into *pex3*Δ cells lacking peroxisomes, fluorescently tagged Pex3p was visible in discrete foci at the perinuclear ER (Hoepfner et al., 2005). Similarly, Tam and colleagues showed the N-terminal 46 amino acids of Pex3p were sufficient to target Pex3p to a subdomain of the ER and to initiate the formation of a preperoxisomal vesicle (Tam et al., 2005). Pex2p and Pex16p have been shown to be *N*-linked glycosylated, a protein modification that is specific to the ER (Titorenko and Rachubinski, 1998).

The extent of the contribution of the ER towards the peroxisome population in wild-type cells is not fully understood. In *S. cerevisiae, de novo* peroxisome formation from the ER can occur in the case of catastrophic loss (Motley and Hettema, 2007). This process is relatively inefficient, and thus the primary role of the ER in wild-type cells appears to be providing proteins and lipids to

sustain multiple rounds of growth and division (Motley and Hettema, 2007). For instance, in *S. cerevisiae*, ER-derived structures containing Pex3p fuse with pre-existing peroxisomes rather than maturing themselves (Motley and Hettema, 2007). A consensus on this question has not been reached for mammalian cells; studies show that either *de novo* formation of peroxisomes (Kim et al., 2006) or fission of existing peroxisomes to be the dominant pathway (Huybrechts et al., 2009; Delille et al., 2010).

1.7 Regulation of the peroxisome population

Given the importance of correct peroxisome assembly, there must be highly conserved mechanisms to maintain peroxisome number through continuous rounds of growth and cell division. Peroxisomes can form through multiple pathways, and the precise mechanisms behind the coordination of these pathways remains unknown. One pathway, termed constitutive division, relates to the fact that peroxisome number is tightly linked to the cell cycle through the need to replicate the cell's organelle population during cell division. Like most organelles in *S. cerevisiae*, peroxisomes are directionally inherited via actin tracks; this inheritance is mediated by the class V myosin motor Myo2p and its peroxisome specific receptor, Inp2p (Fagarasanu et al., 2006). This process must be accompanied by the retention of a subpopulation of organelles in the mother cell, as well as the prevention of diffusion of newly inherited peroxisomes back to the mother cell (Fagarasanu et al., 2005).

A second pathway, termed proliferative division, refers to the growth and division of existing peroxisomes in the response of cells to specific environmental or metabolic cues, such as yeast growing on a non-fermentable carbon source for which peroxisomes are essential (Veenhuis et al., 1987). This leads to an induction and upregulation of the expression of genes encoding peroxisomal proteins and the rapid expansion of the peroxisomal compartment in both size and number of peroxisomes (for more details on this process, see Section 1.8.4).

One unresolved aspect of peroxisome division is whether it is a symmetric or asymmetric process (or both) (reviewed in Mast et al., 2010). Morphological evidence for asymmetrical

peroxisome division includes electron micrographs showing tubular-like extensions from peroxisomes in *S. cerevisiae* (Erdmann and Blobel, 1995). A similar phenomenon has also been observed in the yeast *Hansenula polymorpha*, with a prominent tubule eminating from the peroxisome (Nagotu et al., 2008; Opaliński et al., 2011). Pex11 can oligomerize to form a tubule by elongating a portion of the peroxisome (Koch et al., 2010). Asymmetry of peroxisome division is not restricted to the peroxisome itself, as the enzyme acyl-CoA oxidase in rat peroxisomes is asymmetrically distributed between two newly divided peroxisomes (Wilcke et al., 1995). Conversely, overexpression of the *PEX11* gene in *S. cerevisiae* (Erdmann and Blobel, 1995) or the *PEX11*β gene in human cells (Koch et al., 2002), as well as deletion of the *VPS1* gene encoding a dynamin-related protein in yeast, results in peroxisomes with a "beads on a string" phenotype consistent with symmetric peroxisome division (Hoepfner et al., 2001).

Heterogeneity can also refer to various stages of maturity of peroxisomes within the peroxisome population. Erdmann and Blobel used the term "mature peroxisome" to explain the different densities of organelles observed after density gradient centrifugations, and concluded that this observation was due to oleic acid upregulating peroxisomal gene transcription (Erdmann and Blobel, 1995). Titorenko and Rachubinski demonstrated that there are six different populations of peroxisomes in *Y. lipolytica* (P1-P6) and that immature peroxisomes (P1 and P2) can fuse as a part of the peroxisomal maturation process (Titorenko et al., 2000). However, mature peroxisomes do not fuse (Motley and Hettema, 2007). This finding has been recently confirmed by van der Zand and colleagues, who demonstrated that peroxisomes in *S. cerevisiae* are formed through the fusion of two biochemically distinct preperoxisomal vesicles budding from the ER (van der Zand et al., 2012).

1.8 Peroxisome division

1.8.1 Role of the Pex11 protein family in peroxisome division

Peroxisome division is achieved by both general organelle divisional proteins, including the dynamin-

related GTPases, and peroxisome-specific divisional proteins, notably those of the Pex11 family of peroxins. In *S. cerevisiae*, the Pex11 family is composed of Pex11p, the founding member of the family, together with Pex25p and Pex27p, while in human, the family is made up of PEX11α, PEX11β and PEX11γ forms (Smith and Aitchison, 2013).

Pex11p (originally known as Pmp27 due to the protein's inferred mass of 27 kDa) was initially characterized in *S. cerevisiae* (Marshall et al., 1995). It was noted that Pmp27 shared extensive sequence similarity with a protein family in the yeast *Candida biodinii*, whose members were abundant membrane proteins that were inducible by different growth substrates (Marshall et al., 1995). The role of Pex11p in peroxisome division is well established. Deletion of the *PEX11* gene in *S. cerevisiae* results in cells with fewer, larger peroxisomes (Erdmann and Blobel, 1995), while overexpression of *PEX11* results conversely in cells with an increased number of smaller peroxisomes or elongated structures that are thought to be peroxisomes in the process of dividing (Marshall et al., 1995).

More recent work has shown that Pex11p acts to elongate peroxisomes prior to their scission and subsequent separation (Koch et al., 2010). This model proposes that Pex11p assembles on the peroxisomal membrane at specific sites, stimulating the accumulation of phospholipids. PEX11β was the first peroxisomal protein to display a non-uniform membrane distribution, concentrating at specific sites of peroxisome constriction (Schrader et al., 1998). Cells expressing PEX11β displayed tubular peroxisomes, with PEX11β localized to discrete bands on these structures (Schrader et al., 1998). PEX11β-mediated peroxisome elongation begins with the formation of a membrane protrusion, and then an early peroxisomal membrane compartment derived from a pre-existing peroxisome. Its stepwise conversion into a mature, metabolically active peroxisome is achieved via sequential import of certain peroxisomal matrix proteins, membrane proteins and membrane lipids across the growing tubules (Delille et al., 2010). When the action of PEX11β was inhibited using a fluorescent tag, constriction and division of peroxisomes in mammalian cells was blocked and elongated pre-peroxisomal structures (termed tubular

peroxisomal accumulations) were observed (Delille et al., 2010). Tagged PEX11β initially localized to the site of constriction, and was found only in the tubular structure. However, peroxisomal matrix proteins and PMPs with metabolic functions were found only in the mature globular structure, which was import-competent (Delille et al., 2010). Pex11p then recruits dynamin-related proteins (DRPs) for membrane scission (Koch et al., 2010; Smith and Aitchison, 2013). Truncation mutants of PEX11β revealed that the C-terminus of PEX11β is required for interaction and recruitment of DRPs, but the N-terminus of PEX11β is required for peroxisome proliferation as well as homodimerization (Kobayashi et al., 2007). In *H. polymorpha* cells deleted for the dynamin-related protein *DNM1*, fission-arrested tubular peroxisomes budding from the mother cell were also observed (Cepińska et al., 2011). These cells displayed a differential distribution of peroxisomal membrane proteins, with Pex8p, Pex10p, Pex14p, and Pex25p all found in the membrane protrosion, but Pex11p remaining enriched at the base of the membrane extension (Cepińska et al., 2011).

A link between peroxisome division and peroxisome inheritance machinery has also been demonstrated in *H. polymorpha*: Cepińska and colleagues showed that the peroxisomal membrane extension is due to the pulling force of the myosin motor Myo2p and its peroxisome-specific adaptor protein Inp2p (Cepińska et al., 2011). Another connection between peroxisome division and inheritance is that *H. polymorpha* cells deleted for *PEX11* display an inheritance defect, where peroxisomes fail to be retained in the mother cell and a single peroxisome is found in the bud (Krikken et al., 2009). The authors suggest that this phenotype occurs due to the mismatch of Myo2p and Inp1p in the absence of Pex11p (Krikken et al., 2009).

Pex11p has been shown to exhibit membrane remodeling activity *in vitro* (Opaliński et al., 2011). The N-terminal region of Pex11p of the fungus *Penicillium chrysogenum* contains a conserved amphipathic helix (half hydrophobic, and half polar and positively charged). When a peptide corresponding to this helical region was incubated with liposomes, extensive tubulation was observed. The same results were observed with the amphipathic helices from human, *S. cerevisiae*,

and *H. polymorpha* Pex11 proteins. Mutations in the N-terminal region of the *P. chrysogenum* Pex11p resulted in the absence of peroxisomal extensions and a decreased number of peroxisomes when expressed in *H. polymorpha* mutant strains. The authors suggest that Pex11p plays a role in peroxisome division through membrane remodelling (Opaliński et al., 2011).

How is Pex11p activity regulated? Pex11p was the first example of regulation of peroxisome population by phosphorylation-dependent regulation of a peroxin. Pex11p can be phosphorylated at Ser^{165} and/or Ser^{167}, and phosphorylated Pex11p is a positive effector of peroxisome proliferation. Constitutively dephosphorylated and phosphorylated Pex11p mutants are associated with mature and proliferating peroxisomes, respectively (Knoblach and Rachubinski, 2010). Constitutively dephosphorylated and phosphorylated Pex11p mutants also remain permanently associated with mature peroxisomes and the ER-peroxisome interface, respectively, unable to translocate between the ER and peroxisomes (Knoblach and Rachubinski, 2010). Later work in the yeast *Pichia pastoris* revealed that its Pex11p is similarly phosphorylated at Ser^{173} (Joshi et al., 2012). As with *S. cerevisiae* Pex11p, consitutively dephosphorylated and phosphorylated Pex11p mutants of *P. pastoris* displayed elongated peroxisome and hyper-divided peroxisomes, respectively. Additionally, Pex11p is known to homodimerize, with the monomer form thought to be the active form, whereas the dimer form is more abundant on mature peroxisomes (Marshall et al., 1996). Dimerization is dependent on the formation of cysteine cross-links; mutation of an N-terminal cysteine residue inhibits this process (Marshall et al., 1996). It has been proposed that increased oxidative metabolism in the peroxsiome inhibits further peroxisome division by forcing Pex11p into a dimerized, inactive form (Marshall et al., 1996).

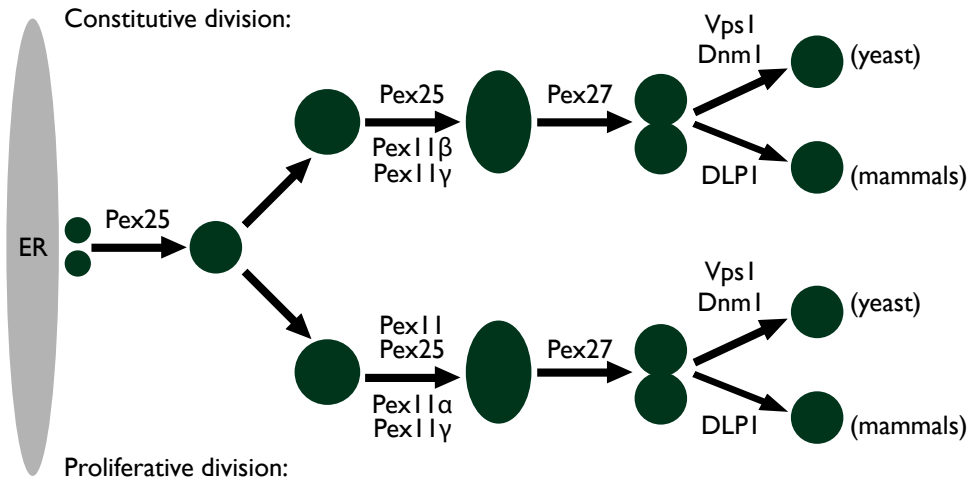
Pex11p, Pex25p and Pex27p in *S. cerevisiae* have been proposed to be a homologous family based on sequence similarity. Pex11p and Pex25p share 10.9% identity and 19.0% similarity, Pex11p and Pex27p share 9.3% identity and 18.4% similarity, and Pex25p and Pex27p share 19.5% identity and 25.9% similarity. Likewise, all three proteins share partially redundant function in peroxisome division; overexpression of *PEX11*, *PEX25* or *PEX27* leads to cells with increased

numbers of small peroxisomes, while deletion of any of these genes leads to cells with decreased numbers of enlarged peroxisomes (Smith et al., 2002; Rottensteiner et al., 2003; Tam et al., 2003). The phenotype of cells deleted for *PEX25* was somewhat heterogeneous, as some cells contained fewer, larger, peroxisomes, but others were devoid of peroxisomes (Smith et al., 2002). *pex25*Δ cells also mislocalized peroxisomal matrix proteins to the cytosol (Smith et al., 2002). Pex25p is necessary for the localization of the small GTPase Rho1p to peroxisomes. Rho1p regulates actin assembly on the peroxisomal membrane, which is necessary for peroxisome fission (Marelli et al., 2004).

In human cells, PEX11γ is thought to recruit PEX11α and PEX11β to the peroxisomal membrane to form PEX11-enriched patches, leading to peroxisome elongation (Koch et al. 2010). PEX11γ is thought to be always required for elongation of the peroxisomal membrane, whereas PEX11α supports proliferative peroxisome division and PEX11β supports constitutive peroxisome division (Koch et al. 2010). Studying the expression of the human Pex11p proteins using RT-PCR revealed that PEX11γ appears to be the limiting factor (Koch et al. 2010). PEX11α was shown to be responsive to peroxisome proliferating agents, whereas PEX11β was not (Schrader et al., 1998). Futhermore, expression of PEX11β was shown to be approximately equal across various human tissues, but differential expression of PEX11α was observed (Schrader et al., 1998). A general model summarizing the roles of the *S. cerevisiae* and *H. sapiens* Pex11 family proteins in constitutive and proliferative peroxisome division is shown in Figure 1-1.

Bioinformatic analyses have identified several putative homologues of Pex11p in Fungi, *i.e.* Pex11Bp, Pex11Cp and Pex11/25p (Kiel et al., 2006). Although the specific methodology leading to the identification of these proteins was not reported, this study raised the interesting possibility that additional members of the Pex11p family have yet to be identified.

Figure 1-1. A model for the roles of Pex11 family proteins in peroxisome division. The top and bottom rows show Pex11 family proteins that are needed for constitutive and proliferative peroxisome division, respectively, in yeast (*S. cerevisiae*) and mammals (*H. sapiens*). Under conditions of constitutive peroxisome division in *S. cerevisiae*, Pex25p acts upstream of or at the stage of peroxisome elongation, while Pex27p acts downstream of peroxisome elongation. Under conditions of proliferative peroxisome division in *S. cerevisiae*, Pex11p and Pex25p act upstream of or at the stage of peroxisome elongation, while Pex27p acts downstream of peroxisome elongation. Pex11β and Pex11α are involved in constitutive and proliferative peroxisome division in *H. sapiens*, respectively, Pex11γ always required for peroxisome division. Pex25p in *S. cerevisiae* is also required for *de novo* peroxisome formation from the ER. In both constitutive and proliferative peroxisome division, peroxisome scission is mediated by the DRPs Vps1p and Dnm1 (yeast) or DLP1 (mammals). Adapted from Tower, 2012.


Constitutive division:
ER
Pex25
Pex25
Pex11β
Pex11γ
Pex27
Vps1
Dnm1
(yeast)
DLP1
(mammals)
Pex11
Pex25
Pex11α
Pex11γ
Pex27
Vps1
Dnm1
(yeast)
DLP1
(mammals)
Proliferative division:

1.8.2 Role of other peroxins in the regulation of peroxisome size and number

Whereas PEX11α, PEX11β and PEX11γ are the only peroxins known to regulate peroxisome size and number in mammalian cells (Schrader et al., 2012), other peroxins are implicated in this role in yeast cells. *S. cerevisiae* cells deleted for *PEX28* or *PEX29* exhibit increased numbers of smaller peroxisomes that are also clustered (Vizeacoumar et al., 2003). *S. cerevisiae* Pex30, Pex31p, and Pex32p act downstream of Pex28p and Pex30p, with cells deleted for *PEX30* displaying an increased number of peroxisomes and cells deleted for *PEX31* or *PEX32* displaying enlarged peroxisomes (Vizeacoumar et al., 2004). Finally, Pex34p functions both cooperatively and independently of the other members of the *S. cerevisiae* Pex11 protein family (Pex11p, Pex25p, and Pex27p) to regulate peroxisome size and number (Tower et al., 2011). Deletion of *PEX34* results in cells with fewer, larger peroxisomes, whereas overexpression of *PEX34* results in cells with an increased number of peroxisomes (Tower et al., 2011).

1.8.3 Constriction and scission of peroxisomes

The precise mechanism of constriction of the peroxisomal membrane is still poorly understood. One possible mechanism in *Y. lipolytica* involves Pex16p, a peripheral membrane protein that has been shown to negatively regulate peroxisome division (Eitzen et al., 1997). *Y. lipolytica* peroxisomes do not grow and divide at the same time; there is a stepwise conversion of five peroxisomal precursors (termed P1-P5) into mature, division-competent peroxisomes (termed P6) (Guo et al., 2007). In immature peroxisomes, Pex16p inhibits a lipid biosynthetic pathway leading to diacylglycerol formation, a known potent inducer of membrane curvature. (Guo et al., 2007). During the peroxisomal maturation process, a decrease in PMPs and an increase in peroxisomal matrix proteins are observed (Guo et al., 2003). This increase in peroxisomal matrix proteins eventually displaces the enzyme acyl-CoA oxidase (AOX) from the peroxisomal matrix to the membrane, where it interacts with and inhibits Pex16p. Diacylglycerol synthesis proceeds and results in remodelling of the membrane bilayer and the stepwise assembly of a complex (that

includes DRPs) that can carry out peroxisome scission (Guo et al., 2007). Only mature peroxisomes can undergo division in *Y. lipolytica* when Pex16p is in complex with AOX (Guo et al., 2003). However, the *S. cerevisiae* genome does not encode a Pex16p homologue, and the mammalian Pex16p homologue has a role at the ER rather than in peroxisome division. Therefore, the molecular players in this process are likely different and remain unknown.

Dynamins form a family of large GTPases that are implicated in various vesicle scission reactions in the cell. Dynamins and DRPs are mechanochemical enzymes that use GTPase-dependent conformational changes to drive fission (Praefcke and McMahon, 2004). Peroxisomal constriction and fission are distinct processes, each requiring a separate machinery, with constriction acting later in the cycle of peroxisome division (Hoepfner et al., 2001). Again, the precise mechanism of peroxiome fission is not completely understood. Thinning of the peroxisomal tubules is likely required for the DRP ring to be effectively assembled around the peroxisome and promote scission, but this is still speculative (Yan et al., 2005; Schrader, 2006). DRPs are actually capable of both constriction and fission, but in the absence of *VPS1*, a beads-on-a-string phenotype occurs, which implies that constriction is still taking place (Hoepfner et al., 2001). The *S. cerevisiae* genome encodes three DRPs (Vps1p, Dnm1p and Mgm1p), with Vps1p being the only one actually essential for peroxisome division, while DLP1 is responsible in mammalian cells (Fagarasanu et al., 2007). Many of the proteins required for peroxisomal scission are shared with mitochondria (reviewed in Scharder et al., 2012). In yeast, the receptor Fis1p recruits Dnm1p to both peroxisomal and mitochondrial membranes (Kuravi et al., 2006; Motley et al., 2008). In mammals, the receptor Mff recruits DLP1 to both peroxisomal and mitochondrial membranes (Otera et al., 2010), whereas Fis1p may play a regulatory role (Schrader et al., 2012).

1.8.4 Transcriptional regulation of peroxisome proliferation in yeast

S. cerevisiae and other yeasts are able to utilize a variety of carbon sources. When glucose is present, enzymes for the breakdown of non-fermentable carbon sources are normally turned off via a gene regulation pathway known as glucose repression (Gancedo, 1998). Upon deletion of

glucose, a number of genes are expressed through glucose de-repression (Schuller, 2003). Oleate induction is accompanied by a drastic expansion of the peroxisomal compartment, and β-oxidation genes are induced several fold by the presence of oleic acid (Veenhuis et al., 1987).

In *S. cerevisiae*, the oleic-acid-specific mechanism affecting the most dramatic increase in transcription of genes encoding peroxisomal functions is based on the oleate response element (ORE). Upon binding a fatty acid ligand, the transcription factors Pip2p and Oaf1p form a heterodimer and bind OREs, activating the transcription of peroxisomal genes with OREs (Einerhand et al., 1993; Filipits et al., 1993). An additional transcription factor, Adr1p, binds to upstream activating sequence 1 (UAS1) promoter sites under conditions of glucose de-repression and oleate induction (Gurvitz et al., 2001). UAS1 sites are often found in close proximity to, or overlapping with, OREs (Gurvitz et al., 2001).

The ORE is a palindrome consisting of two CGG triplets spaced by 15-18 nucleotides, with at least one half-site containing a TNA triplet, making the consensus sequence: 5'-$CGGN_3TNAN_{9-12}CCG$. Predictably, there have been a number of genes published that contain deviations from this consensus sequences. The *ANT1* (adenine nucleotide transporter) ORE has only 14 intervening nucleotides rather than 15-18, but is still functional (Rottensteiner et al., 2002). Another example is the *PEX25* gene, which contains the aberrant ORE site 5'-$CGGN_3TNGN_{9-12}CCG$ missing the conserved TNA triplet (Rottensteiner et al., 2003). A purine at the third position may be suitable for function. Thus, with these two changes, the consensus sequence could be readjusted to: 5'-$CGGN_3TN^{A}/_{R}N_{8-12}CCG$.

A number of *PEX* genes in *S. cerevisiae* have OREs (Table 3-3). Vizeacoumar and colleagues identified putative OREs in the promoter regions of *PEX30, PEX31* and *PEX32* (Vizeacoumar et al., 2003). The levels of Pex30p and Pex32p, but not Pex31p, are induced by incubation in oleic acid-containing medium. However, it was not tested whether these ORE-like motifs are actually functional. Likewise, the levels of Pex5p, Pex11p, Pex14p and Pex25p are oleic acid-inducible, but not that of Pex13p (reviewed in Rottensteiner et al., 2003).

Table 1-3. Oleate response elements (OREs) in *S. cerevisiae PEX* genes

Gene	ORE	UAS1	Reference
PEX5	$CGGN_{10}TTAACTCCG$	$CTCCGAN_4CTCCGAN_{20}GCGGAG$	(Rottensteiner et al., 2003)
PEX7	$CAGN_{10}TNAN_3CCG$	$CTCCAGN_{15}GCCGAG$ $CTCCTCN_{26}AAGGTG$	(Rottensteiner et al., 2003)
PEX11	$CGGN_3TNAN_4GGAGN_3CCG$	$TTCCATN_{19}N_{11}TCGGAG$	(Rottensteiner et al., 2003)
PEX13	$CGGN_{12}TNAN_3CGG$	$CGCCGGN_{17}CAGAAGTTGGTG$	(Rottensteiner et al., 2003)
PEX14	$CGGN_3TNAN_7CCG$	-	(Rottensteiner et al., 2003)
PEX25	$CGGN_3TAGN_5GAAN_3CCG$	-	(Rottensteiner et al., 2003)
PEX30	$CGGN_{12}TNAN_3CCG$	Not reported	(Vizeacoumar et al., 2003)
PEX31	$CGGN_{13}TNAN_8CCG$	Not reported	(Vizeacoumar et al., 2003)
PEX32	$CAGN_2TNAN_{13}CCG$	Not reported	(Vizeacoumar et al., 2003)

Another system for transcriptional regulation of peroxisome proliferation was identified in the fungus *Aspergillus nidulans*. FarA and FarB are both Zn_2-Cys_6 binuclear cluster-domain-containing proteins, like the transcription factors Pip2p and Oaf1p in *S. cerevisiae*. FarA and FarB were isolated when screening for mutants unable to use fatty acids as carbon sources but still capable of growth on acetate (Hynes et al., 2006). Deletion of *farA* resulted in failure to utilize all fatty acids as carbon sources. Deletion of *farB* resulted in a loss of utilization of short chain fatty acids but did not affect long-chain fatty acid utlization (lauric acid, C_{12} and longer) (Hynes et al., 2006). Two homologues that could be classified as FarA and FarB on the basis of alignment were predicted for the ascomycete Fungi *Aspergillus fumigatus*, *Chaetomium globosum*, *Fusarium graminearium*, *Magnaporthe grisea*, *Nectria hematococca*, *Neurospora crassa* and *Stagonospora nodorum* (Hynes et al., 2006). A single Far homologue was identified in several Saccharomycotina yeast genomes: *Candida albicans*, *Debaryomyces hansenii* and *Y. lipolytica* (Hynes et al., 2006). The authors note that FarA and FarB homologues were not detected in higher ascomycete yeasts such as *Ashbya gossypii* and *S. cerevisiae*, and infer that these proteins are only present in lower Ascomycota.

In *A. nidulans*, the transcription factors FarA and FarB bind long-chain and medium-chain fatty acids, respectively, and induce the transcription of genes involved in fatty acid catabolism and peroxisomal proliferation. FarA and FarB bind a conserved CCTCGG motif in the promoter region of target genes (within 1 kilobase of the start codon and present in one or more copies). This conserved CCTCGG motif has been identified in genes implicated in fatty acid metabolism and encoding some peroxins, including Pex11p, in several yeasts (Table 3-4) (Hynes et al., 2006).

Table 1-4. CCTCGG motifs in the 5'-upstream regions of *PEX11* genes from selected Fungi and of selected genes of *Y. lipolytica*

Genome	Gene (Accession)	CCTCGG site(s)
Aspergillus nidulans	*PEX11* (EAA65086.1)	-280[a]
Candida albicans	*PEX11* (XP_712077.1)	-308[a]
Debaryomyces hansenii	*PEX11* (XP_456578.1)	-183[a]
Magnaporthe grisea	*PEX11* (AAX07686.1)	-224, -364, -624[a]
Neurospora crassa	*PEX11* (NCU04802)	-471[a]
Yarrowia lipolytica	*PEX11* (XP_501425.1)	-115[a]
	ICL (isocitrate lyase) (X72848)	-291, -908[a]
	Malate synthase (XP_503993)	-663[a]
	carnitine acetyl transferase (cytoplasmic) (XP_505698)	-697[a]
	carnitine acetyl transferase (peroxisomal) (XP_500716)	-138, -661, -994[a]
	ACOX2 (acyl-CoA oxidase 2) (AJ001300)	-120, -638[a]
	ACOX3 (acyl-CoA oxidase 3) (AJ001301)	-70, -583, -809[a]
	ACOX4 (acyl-CoA oxidase 4) (AJ001302)	-120[a]
	ACOX5 (acyl-CoA oxidase 5) (AJ001303)	-339, -620[a]
	enoyl-CoA dehydrogenase (mitochondrial) (XP_500719)	-64[a]
	short chain acyl-CoA dehydrogenase (mitochondrial) (XP_502873)	-109[a]
	POT1 (Q05493.1)	-127[b]
	POX2 (XP_505264.1)	-120[b]
	PEX5 (Q99144.1)	-221[b]
	PAT1 (Q6L8K7.1)	-55[b]

[a]Hynes et al., 2006.
[b]Poopanitpan et al., 2010.

The yeast *Candida albicans* lacks homologues of the *S. cerevisiae* proteins Oaf1p and Pip2p. Rather, the *C. albicans* genome encodes a single homologue of the *A. nidulans* transcription factors FarA and FarB designated CTF1 (Ramirez and Lorenz, 2009). *ctf1*Δ/Δ cells failed to utilize oleic acid

and other long-chain fatty acids as carbon sources. Northern blots demonstrated that RNAs coding for *C. albicans* peroxisomal proteins and enzymes of β-oxidation are induced in the presence of fatty acids, and this induction is dependent on CTF1.

Three homologues of the *S. cerevisiae* transcription factors Oaf1p and Pip2p were proposed to be present in *Y. lipolytica* (Poopanitpan et al., 2010). Surprisingly, individual deletion strains did not display a growth defect on oleic acid-containing medium, in stark contrast to strains deleted for *OAF1* and *PIP2* in *S. cerevisiae*. The *Y. lipolytica* Por1p (Primary Oleate Regulator 1) is a homologue of FarA in *A. nidulans* and Ctf1p in *C. albicans*. *por1*Δ cells fail to utilize fatty-acid containing medium as a carbon source, and northern blots revealed decreased RNA levels for enzymes involved in β-oxidation in *por1*Δ cells upon induction with oleic acid in comparison to cells of a wild-type strain (Poopanitpan et al., 2010). A conserved FarA binding site was identified in some *Y. lipolytica* genes involved in fatty acid utilization (Table 1-4), although the functional significance of these particular elements in *Y. lipolytica* remains to be elucidated. In summary, these data show that *Y. lipolytica* has mechanisms of transcriptional regulation of fatty acid utilization distinct from those used by *S. cerevisiae*, and that Por1p regulates the transcription of genes involved in fatty acid utilization in *Y. lipolytica* in response to the presence of fatty acids.

1.8.5 Transcriptional regulation of peroxisome proliferation in other eukaryotes

This section briefly summarizes some of the different systems controlling peroxisome proliferation in other eukaryotes. In mammals, the nuclear receptor peroxisome proliferator activated receptor α (PPARα) (Issemann et al., 1990; Pyper et al. 2010; Rakhshandehroo et al., 2010) belongs to a family of steroid/thyroid/retinoid receptors and is responsible for changing the expression of peroxisome genes, including those involved in β-oxidation of fatty acids, and the *PEX11*α gene (Shimizu et al., 2004). Constitutive expression of peroxisomal genes is not dependent on PPARα because mice deleted for PPARα still contain normal peroxisomes (Lee et al., 1995; Gonzalez, 1997). Ligand binding causes a conformational change in PPARα, permitting its interaction with retinoid X

receptor-α (RXRα). The PPARα/RXRα heterodimer can in turn bind PPARα-responsive elements (PPREs) (reviewed in Schrader et al., 2012). Recently, another transcriptional coactivator was described in human and mouse, PGC-1α (peroxisome proliferator activated receptor γ coactivator-1α), that is independent of the PPARα pathway (Bagattin et al., 2010). However, the transcription factor and response element involved in this pathway are yet to be defined.

Curiously, secondary structure predictions and alignments have revealed high similarity between *S. cerevisiae* Pex11p and the ligand-binding domain of nuclear receptors, in particular the PPAR subfamily (Barnett et al., 2000). Nuclear receptors consist of a DNA-binding domain and a ligand-binding domain; the ligand-binding domain precursor is unknown, but Pex11p is proposed in this paper due to this sequence similarity with the PPAR family (Barnett et al., 2000). These authors discussed the fact that nuclear receptors, an abundant class of transcriptional regulators, are absent in unicellular organisms, and suggested that since a ligand-binding domain was identified in yeast, the multidomain nuclear receptor family arose from the combination of pre-existing protein modules. The functional significance of this work remains undetermined.

There are multiple inducers of peroxisome proliferation in *A. thaliana*, and this process is poorly understood (reviewed in Schrader et al., 2012). One potent known inducer of peroxisome proliferation in *A. thaliana* is light exposure. Upon exposure of the plant to far-red light, the photoreceptor phytochrome A (phyA) translocates from the cytoplasm to the nucleus, where it interacts with various transcription factors to activate light-inducible genes (Wang and Deng, 2003). For example, the transcription factor HYH binds a region of the *PEX11B* promoter known as the light responsible element (LRE), which is also present in other genes that are light-inducible. Pex11B initiates peroxisome elongation, thus initiating light-induced peroxisome proliferation (Desai and Hu, 2008).

1.9 Introduction to eukaryotic diversity

The number of sequenced eukaryotes has magnified exponentially over the last decade. The overwhelming amount of publically available sequence data and the relatively low cost of sequencing make genome projects and phylogenetic analyses easier and more common than ever before. This thesis aims to delve into the evolution of peroxisomes and peroxisomal proteins, for which a broader understanding of eukaryotic life outside of typical model systems is required. This section is intended as a very brief overview of eukaryotic diversity to describe the taxonomic placement of familiar model organisms and parasites and for reference for the genome projects included in this thesis (Figure 1-2). One consensus is that six eukaryotic supergroups are recognized, which is the convention kept in this thesis (Walker et al., 2011).

Most common cell biological model systems come from the supergroup Opisthokonta, which can be broadly divided into the Holozoa and Holomycetes (reviewed in Walker et al., 2011). The Holozoa is comprised of Metazoa (better known as animals, for which numerous genomes are available), choanoflagellates (*e.g. Nematostella vectensis*) and several related protists. The Holomycetes are comprised of Fungi and basally divergent protist lineages.

In addition to the massive increase in eukaryotic genome sequencing, there have been many initiatives to increase the number of fungal genomes sequenced as well. This is due to the tremendous impact this lineage has on ecosystem functioning and human health. For instance, the Joint Genome Institute (JGI) has a 1000 Fungal Genomes Project (1000.fungalgenomes.org/home/). Most fungal genomes are from the dikarya clade, which is comprised of the phyla Ascomycota and Basidiomycota (James et al., 2006). The Ascomycota is comprised of three subphyla: Saccharomycotina, Pezizomycotina and Taphrinomycotina (James et al., 2006). The Saccharomycotina is often referred to as the 'true yeasts'. The defining morphological feature of this subphylum is the ascus, a sac-like structure in which ascospores are formed. The Pezizomycotina consists of nine classes and is largely composed of filamentous Fungi and fruit-producing Fungi (James et al., 2006). The Taphrinomycotina is the earliest diverging clade with many species that

Figure 1-2. Evolutionary relationships between selected lineages from the six eukaryotic supergroups. An unrooted tree illustrates the relative evolutionary relationships between selected lineages of eukaryotes from which the genomes used in comparative genomic analyses in this thesis are taken. Some lineages have been omitted for simplicity. Adapted from Walker et al., 2011.

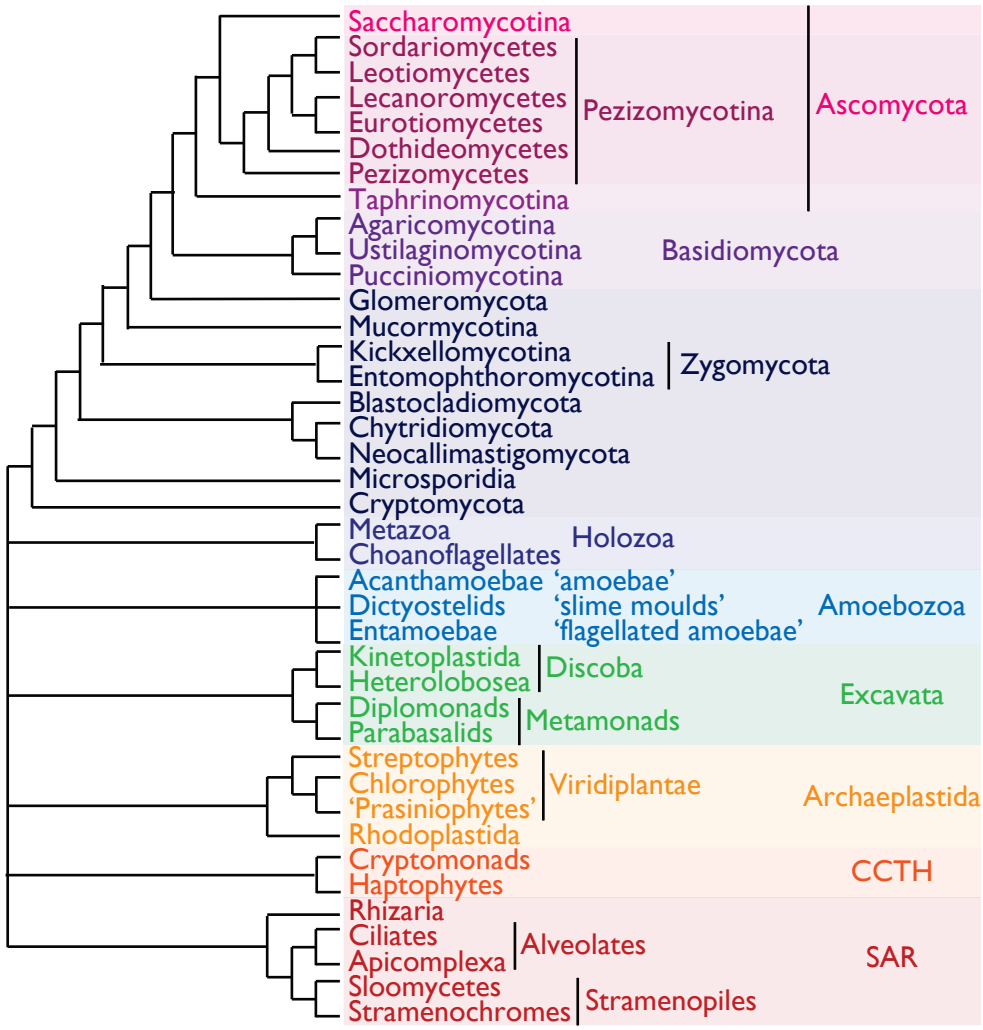

Saccharomycotina
Sordariomycetes
Leotiomycetes
Lecanoromycetes
Eurotiomycetes
Dothideomycetes
Pezizomycetes
Pezizomycotina
Ascomycota
Taphrinomycotina
Agaricomycotina
Ustilaginomycotina
Pucciniomycotina
Basidiomycota
Glomeromycota
Mucormycotina
Kickxellomycotina
Entomophthoromycotina
Zygomycota
Blastocladiomycota
Chytridiomycota
Neocallimastigomycota
Microsporidia
Cryptomycota
Metazoa
Choanoflagellates
Holozoa
Acanthamoebae
Dictyostelids
Entamoebae
'amoebae'
'slime moulds'
'flagellated amoebae'
Amoebozoa
Kinetoplastida
Heterolobosea
Discoba
Diplomonads
Parabasalids
Metamonads
Excavata
Streptophytes
Chlorophytes
'Prasiniophytes'
Viridiplantae
Archaeplastida
Rhodoplastida
Cryptomonads
Haptophytes
CCTH
Rhizaria
Ciliates
Apicomplexa
Alveolates
Sloomycetes
Stramenochromes
Stramenopiles
SAR

exhibit yeast-like and dimorphic morphology. The Basidiomycota are comprised of three subphyla: Agaricomycotina, Ustilaginomycotina, and Pucciniomycotina (James et al., 2006). Basidiomyctes include mushrooms and rusts and are characterized by the formation of club-shaped cells (basidia) on which basidiospores are formed.

The Amoebozoa is comprised of amoebae and amoeboid flagellates, and is the supergroup most closely related to the Opisthokonta (Walker et al., 2011; Adl et al., 2012). Although the taxonomy for this supergroup is presently undefined, for easy reference, the Amoebozoa can be divided into three general groups (reviewed in Walker et al., 2011): amoebae that include *Acanthamoeba castellanii,* which can cause amoebic keratitis; flagellated amoebae including *Entamoeba histolytica*, the causative agent of amoebic dysentery; and slime moulds such as *Dictyostelium discoideum*.

The Excavata is an assemblage of protists that are characterized by a longitudinal feeding groove, which has been secondarily lost in many taxa, that is used for the capture and ingestion of small food particles collected from a current generated by the beating of posteriorly directed cilia. Many excavates contain non-aerobic mitochondrial alternatives (Walker et al., 2011; Adl et al., 2012). The Excavata can be broadly divided into two major divisions: the Metamonada, generally lacking typical mitochondria, and the Discoba, generally having typical mitochondria. The Metamonada includes the parabasalids, which contain an organelle called the hydrogenosome that functions in place of the mitochondrion and include the human parasite *Trichomonas vaginalis*, as well as the diplomonads, which contain mitosomes, an organelle homologous to mitochondria, and include *Giardia lamblia*, the causative agent of beaver fever. Notable Discoba lineages include the Heterolobosea, which lacks visible Golgi stacks and includes the genus *Naegleria* (see Section 1.9.3). Another prominent lineage is the Euglenozoa, a diverse group of flagellates, notably the euglenids, that contain secondary, green algal derived chloroplasts and include *Euglena gracilis*. The Kinetoplastida is characterized by the presence of the kinetoplastid, a mitochondrial genome arranged in circles, and includes the bodonids (see Section 1.9.2) and trypanosomatidae, which

contain various medically relevant parasites including *Trypanosoma brucei* (African Sleeping Sickness), *Trypanosoma cruzi* (Chagas's disease), and *Leishmania major* (Leishmaniasis).

The Archaeplastida is comprised of three phyla, the Viridiplantae, Rhodoplastida and Glaucophyta, that are unified by the presence of primary plastids believed to have arisen from a single endosymbiotic event with a cyanobacterium (reviewed in Walker et al., 2011). The Viridiplantae is comprised of green plants and algae, including the chlorophyceae *Chlamydomonas reinhardtii* and *Volvox carteri*, the mamiellophytes *Osterococcus tauri* and *Micromonas* sp., and the embryophytes, which are land and vascular plants, *e.g. A. thaliana*. There are very few rhodoplastid, or red algal, genomes available (*e.g. Cyanidioschizon merolae*). Glaucophytes contain blue-green chloroplasts and are tentatively more ancestral than primary plastid-containing taxa.

The SAR clade is comprised of stramenopiles, alveolates and Rhizaria. Stramenopiles are a diverse assemblage of photosynthetic and non-photosynthetic unicellular or mutlicellular organisms. Stramenopiles include the sloomycete *Phytophthora ramorum* (sudden oak death), the slopalinids (*e.g. Blastocystis hominis*, see Section 1.9.1), and the stramenochromes (brown algae), including *Thalassiosira pseudonana*, *Phaeodactylum tricornutum* and *Ectocarpus siliculosus*. The alveolates include the ciliates *Tetrahymena thermophila* and *Paramecium tetraurelia*, along with the Apicomplexa, which contains some of the most prominent human parasites: *Plasmodium falciparum* (the causative agent of malaria), *Toxoplasma gondii* (the causative agent of toxoplasmiosis) and *Cryptosporidium parvum* (an intestinal parasite). The Rhizaria is a lineage of eukaryotes with fine root-like, reticulate or filose pseudopodia. There are several major divisions of the Rhizaria, including the chlorarachniophytes (*e.g. Bigelowiella natans*) that contain a secondary, green-algal-derived chloroplast that retains a highly reduced nucleomorph (the remnant nucleus of the algal endosymbiont) (Archibald, 2007).

The CCTH clade comprised of cryotomonads, centrohelids, telonemids and haptophytes is an assemblage of free-living, heterotrophic, mixotrophic and autotrophic organisms. Support for this supergroup is tentative at best, and many studies have concluded that this assemblage is not even a supergroup (Adl et al., 2012). Complete genome sequences are extremely limited, and include the

haptophyte *Emiliania huxleyi*, which contains a red-algal-derived chloroplast and is completely encased in calcium carbonate scales, and the cryptomonad *Guillardia theta*. Cryptomonads have a secondary, red-algal-derived chloroplast that retains a highly reduced nucleomorph (Archibald, 2007).

1.9.1 Introduction to Blastocystis hominis

This section provides brief background information on the eukaryotes that were the focus of international genome projects discussed in this thesis.

Blastocystis hominis is the only stramenopile known to cause infections in humans (Arisue et al., 2002). Frequently found in the intestinal tract of humans and animals, *Blastocystis* is the most frequent protozoan reported in human fecal samples (Tan, 2008; Denoeud et al., 2011). *Blastocystis* has a worldwide distribution with a prevalence of infection of between 30% and 60% in some developing countries. Classification of *Blastocystis* is extremely challenging due to the extensive genetic diversity it exhibits; ten distinct subtypes (ST1 to ST10) have been identified based on molecular analysis of the small subunit RNA gene (Tan, 2008). Symptoms attributed to *Blastocystis* can include gastrointestinal disease accompanied by persistent rashes (Souppart et al., 2010), irritable bowel syndrome (Windsor, 2007; Boorom et al., 2008) and inflammatory bowel disease (al-Tawil et al., 2006), but are not always observed. Therefore, the designation of *Blastocystis* as a pathogen *per se*, *i.e.* as the direct cause of enteric symptoms, is still dubious.

Blastocystis is a member of the SAR clade, for which the characterization of peroxisomes is limited (Gabaldon, 2010). There is ultrastructural and/or biochemical evidence for peroxisomes in the stramenopiles *Phytophthora* (Philippi et al., 1975) and *Phaedactylum tricornutum* (Gonzalez et al., 2011), and peroxisomes have been predicted to be present in *Thalassiosira pseudonana* (Armbrust et al., 2004). Peroxisomes are also present in the ciliates *T. thermophila* and *Paramecium* sp. (Muller, 1973; Stelly et al., 1975). Apicomplexans are the first known eukaryotic lineage to be devoid of peroxisomes (*i.e.* lacking *PEX* genes) in the presence of classical mitochondria (Schlüter et al., 2006). Like nearly all other parasitic organisms (Gabaldon, 2010), peroxisomes have not been described

ultrastructurally for *Blastocystis*. Accordingly, we undertook a comparative genomics survey to identify putative peroxisomal proteins in *Blastocystis*, with the goals of understanding the major protein families in a parasitic genome, trying to predict whether or not peroxisomes are present in *Blastocystis*, and increasing our knowledge about peroxisomes in the less extensively studied SAR clade.

1.9.2 Introduction to Bodo saltans

Bodo saltans is a free-living, non-parasitic kinetoplastid that is found worldwide. It has the characteristic features of the Kinetoplastida lineage, *i.e.* a kinetoplast (mitochondrial DNA at the base of the flagellum) and glycosomal enzymes subcompartmentalized in the divergent peroxisome known as the glycosome. Preliminary results of the *Bodo* genome project have been published (Jackson et al., 2009). *Bodo* is an invaluable outgroup to the kinetoplastida, containing the deadly parasites *Trypanosoma* and *Leishmania*. It has been well established that members of the Kinetoplastida, including *B. saltans*, *Trypanosoma* and *Leishmania*, contain divergent peroxisomes known as glycosomes (Opperdoes et al., 1977). The completed genome of *B. saltans* is potentially very interesting for studying glycosome evolution. More specifically, are there differences in the glycosomal *PEX* gene complement and other glycosomal proteins between the non-parasitic *B. saltans* and the parasitic trypanosomatids? The *B. saltans* genome also addresses the question of how trypanosomatid parasites acquired parasitic strategies and how typical kinetoplastid features, such as the glycosome, might have evolved and been modified for parasitism.

1.9.3 Introduction to Naegleria fowleri

It is not known if typical peroxisomes are present in the Excavata outside of the predominantly parasitic kinetoplastids. A single metamonad organism, *Giardia lamblia*, has been shown to be devoid of peroxisomes (de Souza et al., 2004) and *PEX* genes (Schlüter et al., 2006). By and large, previous studies of peroxisomal proteins in this supergroup have been limited by the number of completed genome projects. One such genome of particular interest was the *Naegleria fowleri* genome

(Herman et al., 2013). *N. fowleri*, better known as the "brain-eating amoeba", is free living in warm freshwater but is also an opportunistic pathogen, causing fatal primary amoebic meningoencephalitis in humans. (Visvesvara et al., 2007). *N. fowleri* is thermotolerant, growing in temperatures up to 45°C, and infections occur worldwide. Infection occurs in swimmers or people with fresh water contact and can occur in immunocompetant people. Infection occurs when the amoeba enters the nose and passes through the cribiform plate to reach the olfactory bulb (Martinez and Visvesvara, 1997). Clinical presentation includes severe headaches, fever, nausea, vomiting, stiff neck, neurological abnormalities and onset of coma (Carter, 1972; Martinez and Visvesvara, 1997). Since these symptoms are identical to those of bacterial meningitis, this is probably one of the reasons why *N. fowleri* infection is likely underreported (Visvesvara et al., 2007). Onset of symptoms is extremely rapid, and death occurs within about ten days without prompt diagnosis and intervention. The mortality rate is approximately 99% (Visvesvara and Stehr-Green, 1990).

There are over 30 known species of *Naegleria*, but only *N. fowleri* causes disease in humans (Visvesvara et al., 2007). The genome of *N. gruberi*, which is non-pathogenic and non-thermotolerant, has been completed (Fritz-Laylin et al., 2010). Full analysis of the genome of *N. fowleri* will provide insight as to why *N. fowleri* is deadly while its close relative is harmless. More specifically, *N. fowleri* may encode genes obtained by HGT, novel paralogues of known gene families and novel genes with no known homolgoues, all of which may be potentially involved in pathogenesis. From the perspective of peroxisome evolution, peroxisomes in the Excavata have not been described outside of the Kinetoplastida (organisms with divergent glycosomes). In the course of *N. gruberi* genome annotation, homologues of peroxins and lipid-metabolizing enzymes were noted. PTSs were also noted for a number of metabolic enzymes; the authors conclude that the organelle is present based on these observations (Fritz-Laylin et al., 2010). However, there was no ultrastructural evidence for a peroxisome in this project (*i.e.* peroxisomes were not noted in electron micrographs of *N. gruberi*) (Fritz-Laylin et al., 2010). Thus, the *Naegleria* genomes are an

important sampling point for peroxisome evolution, as these species may have canonical or modified peroxisomes, or no peroxisomes or *PEX* genes at all.

1.10 Diverse peroxisomes for diverse eukaryotes

Some eukaryotic peroxisomes have such highly specialized functions that they are known by distinctive names (reviewed in Gabaldón, 2010). Peroxisomes in trypanosomes are known as glycosomes because they compartmentalize glycolysis enzymes (Opperdoes et al., 1977), and some plant peroxisomes are called glyoxysomes because they contain glyoxylate cycle enzymes (Hayashi et al., 2002). Filamentous fungi have modified peroxisomes called Woronin bodies involved in the maintenance of cellular integrity by sealing the septal pore in response to wounding (Liu et al., 2008).

1.11 Peroxisome evolution

The peroxisome is an example of a eukaryotic organelle that does not have a clear autogenous or endosymbiotic origin (for an extensive discussion, see de Duve, 2007). Peroxisomes may have initially been a respiratory organelle, acquired as the result of the appearance of molecular oxygen in the Earth's atmosphere that would have been toxic to anaerobic organisms (de Duve, 2007). Peroxisomes were likely present in the last eukaryotic common ancestor (LECA) based on the fact that they are in at least one genome from every eukaryotic supergroup (at the time surveryed, this is the most parsimonious explanation) (Schlüter et al., 2006; Gabaldón et al., 2006). Functionally, peroxisomes appear to be derived from the ER (Hoepfner et al., 2005). Shared homology between selected peroxisomal proteins and the endoplasmic reticulum-associated protein degradation (ERAD) system has also been reported (Schlüter et al., 2006; Gabaldón et al., 2006). An evolutionary model proposes that peroxisomes are derived from the ER based on the need to sequester enzymes and their toxic byproducts and free radicals (Gabaldón, 2013).

But peroxisomes do have interactions with organelles of an endosymbiotic origin: mitochondrially derived vesicles fuse with a subset of peroxisomes (Neuspiel et al., 2008). The only

known cargo for these vesicles is MAPL (mitochondrial-anchored protein ligase), and MAPL-contaning vesicles are targeted to a subpopulation of peroxisomes (Neuspiel et al., 2008). This trafficking pathway is dependent on the retromer subunit Vps35 (Braschi et al. 2010). The authors compare this process to the fusion of bacterial membrane vesicles, suggesting that mitochondrially derived vesicles may have other targets and that this is an ancient, conserved process (Andrade-Navarro et al., 2009). In wild-type *S. cerevisiae* cells, a subset of peroxisomes was found adjacent to ER-mitochondrial junctions marked by ERMES (ER-Mitochondrial Encounter Structure) (Cohen et al., 2014). Cells deleted for ERMES components displayed aberrant peroxisome morphology (Cohen et al., 2014). A subset of peroxisomes was also found adjacent to the pyruvate dehydrogenase (PDH) complex in mitochondria, which may concentrate metabolites in specific mitochondrial domains or be an efficient way to transport metabolites from one organelle to another (Cohen et al., 2014). PBDs are often characterized by mitochondrial dysfunction, which is poorly understood but ultimately amplifies cellular damage (Mohanty and McBride, 2013). Peroxisomes and mitochondria also share multiple functions, including fatty acid β-oxidation, thermogenesis, metabolism of reactive oxygen species, photorespiration and, in plants, the glyoxylate cycle (Waterham, 2006; Schrader and Yoon, 2007). A number of proteins also exhibit dual targeting to peroxisomes and mitochondria (reviewed in Schrader and Yoon, 2007). Since fatty acid β-oxidation takes place exclusively the peroxisome in yeast, it has been postulated that yeast peroxisomes are more closely tied to the ER rather than the mitochondria (Mohanty and McBride, 2013). In *pex3*Δ, *pex16*Δ or *pex19*Δ mammalian fibroblast cells, mislocalization to the mitochondria has been documented for some peroxisomal proteins (Sacksteder et al., 2000; South et al., 2000; Kim et al., 2006; Toro et al., 2009). Whether or not peroxisomes themselves are derived from mitochondria is an ongoing debate (Speijer, 2013; Gabaldón, 2014).

In Fungi, the 3′-untranslated regions of genes encoding metabolic enzymes can be activated by post-translational processing to express sequences for cryptic PTS1s (Freitag et al., 2012). This finding explains the dual localization of some enzymes and the partial peroxisomal targeting of well

established cytosolic enzymes, resulting in a peroxisomal metabolic network that was previously thought to be present only in the cytoplasm. This work is significant because it shows how easy it is for a protein to acquire a PTS1 and be targeted to peroxisomes, with this event potentially occuring many times throughout eukaryotic diversity.

1.12 Evolution of *PEX* genes

The conservation of *PEX* genes has been previously studied (Schlüter et al., 2006; Gabaldón et al., 2006). Similarly, the PeroxisomeDB initiative (peroxisomeDB.org) seeks to maintain a list of peroxins across completed eukaryotic genomes (Schlüter et al., 2010). Although informative, this resource is heavily biased towards mammalian and fungal genomes. Schlüter and colleagues completed a comparative genomic study of peroxins using mammalian and yeast genomes, as well as a few protist genomes that were available at the time of publication. They found that while some *PEX* genes are conserved and ancestral, there are also lineage-specific losses. The authors also deduced that there are four peroxins (Pex3p, Pex10p, Pex12p and Pex19p) that can be considered peroxisomal markers, because they were always present in organisms with peroxisomes and always absent from organisms for which peroxisomes had not been described. Gabaldón and colleagues took a similar approach to study the evolution of the peroxisomal proteome. Using annotated peroxisomal proteins coded in rat, yeast, plant and trypanosome genomes, minimal peroxisomal proteomes for rat, yeast, the opisthokont lineage and eukaryotes were reconstructed. The minimal eukaryotic peroxisome was comprised of the peroxins Pex1p, Pex2p, Pex4p, Pex5p, Pex10p and Pex14p, and proteins involved in fatty acid metabolism and transport (Pxa1p/Pxa2p, Fox2p and Faa2p), and catalase; more proteins were added with increased genome complexity (Schlüter et al., 2006; Gabaldón et al., 2006).

Both of these studies also investigated the origin of peroxisomal proteins. The majority of peroxins and the peroxisomal membrane proteins tested had a eukaryotic origin, but strikingly, the matrix and membrane enzymes were of bacterial or archael origin (Schlüter et al., 2006; Gabaldón

et al., 2006). Pex1p, Pex5p, Pex7p and Pex14p were of prokaryotic origin, which is interesting because these peroxins are required for matrix protein targeting and import. These results indicate that peroxins cannot be excluded from having a prokaryotic origin, and that the evolutionary origin of peroxisomes themselves is still debatable. But phylogenetic analyses failed to recover statistical support for a prokaryotic origin of some *PEX* genes (Gabaldón and Capella-Gutiérrez, 2010).

Since these publications, new *PEX* genes have been identified (Managadze et al., 2010; Tower et al., 2011), and there has been a dramatic increase in the number of eukaryotic genome sequencing projects, allowing for a more complete picture of peroxisomal protein evolution.

1.13 Focus of this thesis

The objective of this thesis was to advance our understanding of peroxisome evolution. There were two specific aims:

1.13.1 Specific Aim 1. Evolution of the Pex11 protein family and functional characterization of the Pex11 protein family in Yarrowia lipolytica

In Chapter 3 of this thesis, I present the results of an evolutionary analysis of the Pex11 protein family. I completed a pan-eukaryotic comparative genomic survey and phylogenetic analyses of this protein family. Pex11p itself is highly conserved and ancestral, and has undergone numerous lineage-specific duplications, while other Pex11 protein family members are fungal-specific innovations. I expanded on this evolutionary study by functionally characterizing the *in silico* predicted Pex11 protein family members of the yeast *Y. lipolytica*, *i.e.* Pex11p, Pex11Cp and Pex11/25p. I showed that these proteins are peroxisomal integral membrane proteins. Pex11Cp and Pex11/25p have conserved a role in the regulation of peroxisome size and number characteristic of Pex11 protein family members. Unexpectedly, deletion of *PEX11* in *Y. lipolytica* produced cells that lack morphologically identifiable peroxisomes and mislocalize peroxisomal matrix and membrane proteins. This has not been reported so far for any other eukaryotic cell deleted for the *PEX11* gene. My results demonstrate an unprecedented role for Pex11p in peroxisome assembly.

1.13.2 Specific Aim 2. Peroxisomal proteins in extant eukaryotic genomes

I have contributed to three international genome projects by annotating peroxisomal proteins. In Chapter 4 of this thesis, I present my results of comparative genomic surveys of peroxins and/or other peroxisomal proteins in the intenstinal parasite *Blastocystis hominis*; *Bodo saltans*, a close relative of the parasitic trypanosomatids that contains a divergent peroxisome known as the glycosome; and *Naegleria fowleri*, perhaps better known as the 'brain-eating amoeba'. I showed that *Blastocystis* lacks peroxisomal proteins, consistent with parasites that generally lack the peroxisome organelle. I demonstrated that *B. saltans* and *N. fowleri* encode a relatively complete *PEX* gene complement, but with certain losses. I also wrote a program to predict PTS-containing proteins, thereby contributing an additional approach to the *in silico* prediction of peroxisomal proteins, and used it during some of my analyses.

Chapter 2: Materials and Methods

2.1 Materials for molecular and cellular biology

Table 2-1. Chemicals and reagents

Reagent	Source
1-butanol	Fisher
2-mercaptoethanol	BioShop
2-(N-Morpholino)ethanesulfonic acid (MES)	Sigma-Aldrich
2-propanol	Fisher
2,4,6-tri-(dimethylaminomethyl) phenol (DMP-30)	Marivac
5-(N-2,3-dihydroxypropylacetamido)-2,4,6-tri-iodo-N-N'-bis(2,3-dihydroxypropyl)isophthalamide (Nycodenz)	BioLynx/Axis-Shield POC
acetone	Fisher
acrylamide	Roche
adenosine 5′-triphosphate (ATP), disodium salt hydrate	Sigma-Aldrich
agar	Difco
agarose, UltraPure	Invitrogen
albumin, bovine serum (BSA)	Roche
ammonium persulfate	BDH
ammonium sulfate ($(NH4)_2SO_4$)	BDH
ampicillin	Sigma-Aldrich
anhydrous ethyl alcohol	Commercial Alcohols
boric acid	EM Science
Brij 35	EM Science
bromophenol blue	BDH
chloroform	Fisher
complete, Mini, protease inhibitor cocktail tablets	Roche
complete, Mini, EDTA-free protease inhibitor cocktail tablets	Roche
complete supplement mixture (CSM)	BIO 101
complete supplement mixture, leucine dropout (CSM-LEU)	BIO 101
complete supplement mixture, uracil dropout (CSM-URA)	BIO 101
Coomassie Brilliant Blue R-250	ICN
D-(+)-glucose	EM Science
DDSA, specially distilled	Marivac
deoxyribonucleotide triphosphate mixture (dNTPs)	Invitrogen
dimethyl sulfoxide (DMSO)	Sigma-Aldrich
dithiothreitol (DTT)	Fisher
DNase I	Roche
ethidium bromide	Sigma-Aldrich
ethylenedinitrilo-tetraacetic acid (EDTA)	EM Science
glacial acetic acid, 17.4 M	Fisher
glass beads, acid washed	Sigma-Aldrich
glycerol	EM Science
glycine	Roche
isoamyl alcohol	Fisher
isopropyl β-D-thiogalactopyranoside (IPTG)	Roche
hydrochloric acid	BDH
hygromycin	Sigma-Aldrich
lanolin	Alfa Aesar
lithium acetate	Sigma-Aldrich
lysozyme	Sigma-Aldrich

L-histidine	Sigma-Aldrich
L-leucine	Sigma-Aldrich
L-lysine	Sigma-Aldrich
magnesium chloride	EM Science
magnesium sulfate ($MgSO_4$)	Sigma-Aldrich
methanol	Fisher
methyl nadic anhydride (MNA)	Marivac
neosourthricin (NAT)	HKI
N,N,N',N'-tetramethylethylenediamine (TEMED)	EM Science
oleic acid	Fisher
paraffin	Fisher
penicillin	Invitrogen
peptone	Difco
phenol, buffer saturated	Invitrogen
poly L-lysine	Sigma-Aldrich
polyethylene glycol, M.W. 3350 (PEG)	Sigma-Aldrich
Ponceau S	Sigma-Aldrich
potassium chloride	BDH
potassium permanganate ($KMnO_4$)	BDH
potassium phosphate, dibasic (K_2HPO_4)	EM Science
potassium phosphate, monobasic (KH_2PO_4)	EM Science
propylene oxide	Marivac
salmon sperm DNA, sonicated	Sigma-Aldrich
skim milk	Carnation
sodium acetate	EM Science
sodium azide	Sigma-Aldrich
sodium carbonate (Na_2CO_3)	BDH
sodium chloride	EM Science
sodium deoxycholate	Sigma-Aldrich
sodium dodecyl sulfate (SDS)	Bio-Rad
sodium hydroxide (NaOH)	BDH
sodium periodate	Sigma-Aldrich
sorbitol	EM Science
sucrose	EM Science
TAAB 812 resin	Marivac
trichloroacetic acid (TCA)	EM Science
tris(hydroxymethyl)aminomethane (Tris)	Roche
Triton-X-100	VWR
tryptone	Difco
Tween 20	Sigma-Aldrich
Tween 40	Sigma-Aldrich
uracil	Sigma-Aldrich
vaseline	Vaseline
xylene cyanol FF	Sigma-Aldrich
yeast extract	Difco
yeast nitrogen base without amino acids (YNB)	Difco

Table 2-2. Enzymes

Reagent	Source
CIP (calf intestinal alkaline phosphatase)	NEB
DNA Polymerase I, Large (Klenow) Fragment	NEB
Easy-A high-fidelity polymerase	Stratagene
pJET1.2/blunt vector system	Thermo Scientific
restriction endonucleases	NEB
RNase A (ribonuclease A), bovine pancreas	Sigma-Aldrich
T4 DNA ligase	NEB
Taq DNA polymerase	NEB
Zymolyase 100T	ICN

Table 2-3. Molecular size standards

Molecular size standard	Source
1 kb DNA ladder (500-10,000 bp)	NEB
prestained protein marker, broad range (6-175 kDa)	NEB

Table 2-4. Kit systems for molecular biology

Kit system	Source
BigDye Terminator Cycle Sequencing Ready Reaction Kit	Applied Biosystems
DC Protein Assay	Bio-Rad
QIAprep Spin Miniprep Kit	Qiagen
QIAquick Gel Extraction Kit	Qiagen
QIAquick PCR Purification Kit	Qiagen

Table 2-5. Plasmids

Plasmid	Description	Source
pINA443	expression vector for *Y. lipolytica*	Barth and Gaillardin, 1996
pINA443-*GFP-HDEL*	*GFP-HDEL* expression vector for visualizing a fusion of GFP and HDEL in *Y. lipolytica*	this study
pINA445	expression vector for *Y. lipolytica*	Nuttley et al., 2003
pINA445-*GFP-HDEL*	*GFP-HDEL* expression vector for visualizing a fusion of GFP and HDEL in *Y. lipolytica*	this study
pJET1.2/blunt	multipurpose cloning vector for amplifying PCR-generated DNA in *E. coli*	Thermo Scientific
pMAL-c2	expression vector for purification of MBP fusion proteins in *E. coli*	NEB
pMAL-c2-*PEX2*	expression protein for protein purification of a fusion protein consisting of MBP and *Y. lipolytica* Pex2p	Eitzen et al., 1996

pTC3	expression vector for *Y. lipolytica* for exogenous gene expression regulated by the promoter and terminator of the *Y. lipolytica POT1* gene; source of *URA3* for targeted gene deletion in *Y. lipolytica*	Lin et al., 1999
pTC3-*PEX11-mCherry*	*PEX11-mCherry* expression vector for visualizing a fusion of Pex11p and mCherry in *Y. lipolytica*	this study
pTC3-*PEX11C-mCherry*	*PEX11C-mCherry* expression vector for visualizing a fusion of Pex11Cp and mCherry in *Y. lipolytica*	this study
pTC3-*PEX11/25-mCherry*	*PEX11/25-mCherry* expression vector for visualizing a fusion of Pex11/25p and mCherry in *Y. lipolytica*	this study
pTC3-NAT	expression vector for *Y. lipolytica* for exogenous gene expression regulated by the promoter and terminator of the *Y. lipolytica POT1* gene	this study
pTC3-NAT-*PEX11C-mCherry*	*PEX11C-mCherry* expression vector for visualizing a fusion of Pex11Cp and mCherry in *Y. lipolytica*	this study
pUB4	expression vector for *Y. lipolytica*	Kerscher et al., 2001
pUB4-*mRFP-SKL*	*mRFP-SKL* expression vector for visualizing a fusion of mRFP and SKL in *Y. lipolytica*	Chang et al., 2009
pUB4-*PEX11*	*PEX11* expression vector	this study
pUB4-*PEX11C*	*PEX11C* expression vector	this study
pUB4-*PEX11/25*	*PEX11/25* expression vector	this study
pUB4-*PEX3B-mCherry*	*PEX3B-mCherry* expression vector for visualizing a fusion of Pex3Bp and mCherry in *Y. lipolytica*	Chang et al., 2009
pUB4-*PEX11-mCherry*	*PEX11-mCherry* expression vector for visualizing a fusion of Pex11p and mCherry in *Y. lipolytica*	this study
pUB4-*PEX11C-mCherry*	*PEX11C-mCherry* expression vector for visualizing a fusion of Pex11Cp and mCherry in *Y. lipolytica*	this study
pUB4-*PEX11/25-mCherry*	*PEX11/25-mCherry* expression vector for visualizing a fusion of Pex11/25p and mCherry in *Y. lipolytica*	this study
pUB4-*PEX14-mCherry*	*PEX14-mCherry* expression vector for visualizing a fusion of Pex14p and mCherry in *Y. lipolytica*	this study
pUB4-*POT1-mRFP*	*POT1-mRFP* expression vector for visualizing a fusion of Pot1p and mRFP in *Y. lipolytica*	Chang et al., 2013

Table 2-6. Primary antibodies

Specificity	Type	Name	Dilution	Source
DsRed	rabbit	-	1:2,000	Takara Bio, Inc.
Y. lipolytica AOX1p	rabbit	-	1:4,000	Titorenko et al., 2000
Y. lipolytica AOX3p	rabbit	-	1:4,000	Titorenko et al., 2000
Y. lipolytica AOX5p	rabbit	-	1:4,000	Titorenko et al., 2000
Y. lipolytica Pex2p, affinity-purified	guinea pig	Pay5-NN	1:1,000	Eitzen et al., 1996; this study
Y. lipolytica Pex3Bp	guinea pig	V12-final	1:2,000	Chang et al., 2009
Y. lipolytica Pex24p	rabbit	L257-3°	1:1,000	Tam and Rachubinski, 2002
Y. lipolytica Pot1p	guinea pig	Notch	1:10,000	Eitzen et al., 1996
Y. lipolytica ICLp	rabbit	E405-1°	1:2,000	Eitzen et al., 1996
C-terminal SKL	rabbit	16	1:2,000	Aitchison et al., 1992
S. cerevisiae Sdh2p	rabbit	-	1:2,000	Dibrov et al., 1998

Table 2-7. Secondary antibodies

Specificity	Type	Dilution	Source
horseradish peroxidase (HRP)-conjugated anti-rabbit IgG	donkey	1:20,000	GE Healthcare
HRP-conjugated anti-guinea pig IgG	goat	1:10,000	Sigma-Aldrich

Table 2-8. Oligonucleotides

Name	Sequence	Application
AA0834-PEX19-Pr1	TAGAATTCATGTCACACGAAGAAGATCTTG	Amplify *PEX19*, checking oligo for *PEX19* deletion, fwd
AA0875-PEX19-Pr3	TATCTAGATCACTCGGGCATGTTCTCGGGC	Amplify *PEX19*, checking oligo for *PEX19* deletion, rev
AA0896-A16	AATGAGAGCGCTGTATTC	Amplify *PEX3*, checking oligo for *PEX3* deletion, fwd
AA0897-A15	TTGCCTCCTCATCTCTTC	Amplify *PEX3*, checking oligo for *PEX3* deletion, rev
0284QC-mRFP-FWD2	CGCCCTCGCCCTCGATC	*mCherry* 5'-end sequencing oligo
0720-YLPEX16-OFR-Forward	ATGACGGACAAGCTGGTCAA	Amplify *PEX16*, checking oligo for *PEX16* deletion, fwd
0721-YLPEX16-OFR-Reverse	TTAGAGAGTAGAGGCGGTGA	Amplify *PEX16*, checking oligo for *PEX16* deletion, rev
0775-RP-GFP+200bp5'out	CGGGAAAAGCATTGAACACCA	*GFP*+ 5'-end sequencing oligo
0776-RP-GFP+520bp3'out	AATTCGCCACAACATTGAAGAT	*GFP*+ 3'-end sequencing oligo
0847-JC-Pr7	ACGTCTCTGACTACGAGAACT	Checking oligo for deletion with *URA3*, fwd
0848-JC-Pr8	AGTTCTCGTAGTCAGAGACGT	Checking oligo for deletion with *URA3*, rev
1007-RP-mCherry-3'out	GAGGCTGAAGCTGAAGGAC	*mCherry* 3'-end sequencing oligo
1224-JC-Pr P-THI	ATTATCGATAACCTACCGGTTGTTGCTCTC	Amplify pTC3 thiolase promoter with ClaI cut site
1225-JC-Pr T-THI	ATTATCGATATCTACGACCTGGGAAACATG	Amplify pTC3 thiolase promoter with ClaI cut site
1708-JC-Pr P-TC3	CCGAAAGTTGCAACTACC	pTC3 sequencing oligo, fwd

1709-JC-Pr T-TC3	ACTCGTACACTCGTACTC	pTC3 sequencing oligo, rev
2653-JC-Pex11-Pr1	CGTACATCCTTCGATTCCGTA	*PEX11* deletion
2654-JC-Pex11-Pr2	ATACTCGTCGAC**CGCAAGAGACGTATGCTC AAT**	*PEX11* deletion
2655-JC-Pex11-Pr3	GCCTTTGTCGAC**TTCGTCACCACCGCAAAT AAG**	*PEX11* deletion
2656-JC-Pex11-Pr4	AAGGAGTAAGCTGAGTGCGTA	*PEX11* deletion
2657-JC-Pex11-Pr5	ACGTCTCTTGCG**GTCGACGAGTATCTGTCT GAC**	*PEX11* deletion
2658-JC-Pex11-Pr6	GGTGGTGACGAA**GTCGACAAAGGCCTGTT TCTC**	*PEX11* deletion
2659-JC-Pex11-Pr7	CCATCACGTGACATACACAGA	*PEX11* deletion
2660-JC-Pex11-Pr8	TCACGATTACATCATCTGCGG	*PEX11* deletion
2661-JC-Pex11/25-Pr1	CAGGTTCTGAAAGCGTTCTTG	*PEX11/25* deletion
2662-JC-Pex11/25-Pr2	ATACTCGTCGAC**TTTGAGGAGTCACAGAAA CCG**	*PEX11/25* deletion
2663-JC-Pex11/25-Pr3	GCCTTTGTCGAC**ACAGGTTGGATCATGCTC AAC**	*PEX11/25* deletion
2664-JC-Pex11/25-Pr4	AATTGCTGTCAAAGGCGAGAGC	*PEX11/25* deletion
2665-JC-Pex11/25-Pr5	TGACTCCTCAAA**GTCGACGAGTATCTGTCT GAC**	*PEX11/25* deletion
2666-JC-Pex11/25-Pr6	GATCCAACCTGT**GTCGACAAAGGCCTGTTT CTC**	*PEX11/25* deletion
2667-JC-Pex11/25-Pr7	CTGACTAGAACCAGAAGTGCT	*PEX11/25* deletion
2668-JC-Pex11/25-Pr8	TTGCCTCATCAGCTTCGTTTC	*PEX11/25* deletion
2669-JC-Pex11C-Pr1	AGTTTGCTGCTTCGGATGTGA	*PEX11C* deletion
2670-JC-Pex11C-Pr2	ATACTCGTCGAC**TTGGAAGTGGAAGAAGG CTCT**	*PEX11C* deletion
2671-JC-Pex11C-Pr3	GCCTTTGTCGAC**TGTGTGTCTGTGTACTCG TAG**	*PEX11C* deletion
2672-JC-Pex11C-Pr4	TGAGAGGTTACCAATGGTCAC	*PEX11C* deletion
2673-JC-Pex11C-Pr5	TTCCACTTCCAA**GTCGACGAGTATCTGTCT GAC**	*PEX11C* deletion
2674-JC-Pex11C-Pr6	CACAGACACACA**GTCGACAAAGGCCTGTT TCTC**	*PEX11C* deletion
2675-JC-Pex11C-Pr7	AGTCACAGCTCTATGTGCTCA	*PEX11C* deletion
2676-JC-Pex11C-Pr8	TACCTCCAGAAGACTCCTTGA	*PEX11C* deletion

2736-JC-Pr7 in LEU	AACGAGGCGTTCGGTCTGTA	Checking oligo for deletion with *LEU2*, fwd
2738-JC-Pr8 in LEU	GCAGACAGAATGGTGGCAAT	Checking oligo for deletion with *LEU2*, rev
3256-PrPex11-1	ATTGAATTCATGTCCGTTTGCCTCGCCC	*PEX11* with 5' EcoR1 cut site
3257-PrPex11-2	ATTGAATTCCTAAGCAGTAGCGGCCCAG	*PEX11* with 3' EcoR1 cut site
3258-PrPex11-3	TATATCTCCTTCAGCAGTAGCGGCCCAGGC	*PEX11-mCherry* construction
3259-PrPex11-4	GCCGCTACTGCTGAAGGAGATATACATGGCGGA	*PEX11-mCherry* construction
3260-PrPex11-5	ATTGAATTCTTACTTGTACAGCTCGTCCATG	*mCherry* with 3' EcoR1 cut site
3278-JC-Pex11C-1	ATTGAATTCATGTCGCTAGCATACACAGT	*PEX11C* with 5' EcoR1 cut site
3279-JC-Pex11C-2	TATATCTCCTTCCTGCATCTTCTTCTCGACAG	*PEX11C-mCherry* construction
3280-JC-Pex11C-3	AAGAAGATGCAGGAAGGAGATATACATGGCGGA	*PEX11C-mCherry* construction
3281-JC-Pex11/25-1	ATTGAATTCATGGTGTCTAAGCTGGTTGG	*PEX11/25* with 5' EcoR1 cut site
3282-JC-Pex11/25-2	TATATCTCCTTCGCAGGTCTTGCATCGCTCC	*PEX11/25-mCherry* construction
3283-JC-Pex11/25-3	TGCAAGACCTGCGAAGGAGATATACATGGCGGA	*PEX11/25-mCherry* construction
4993-MK1-YlPEX11-F	ATTGAATTCATGTCCGTTTGCCTCGCC	*PEX11* with 5' EcoR1 cut site
4994-MK2-YlPEX11C-F	ATTGAATTCATGTCGCTAGCATACACAGT	*PEX11C* with 5' EcoR1 cut site
4995-MK3-YlPEX1125-F	ATTGAATTCATGGTGTCTAAGCTGGTTGG	*PEX11/25* with 5' EcoR1 cut site
5099-MK7-YlPEX11C-R	ATTGAATTCCTACTGCATCTTCTTCTCGA	*PEX11C* with 3' EcoR1 cut site
5100-MK8-YlPEX1125-R	ATTGAATTCCTAGCAGGTCTTGCATCGC	*PEX11/25* with 3' EcoR1 cut site
5513-MK37-pTC3-THI-AclI-F	ATTAACGTTAACCTACCGGTTGTTGCTCTC	Amplify pTC3 thiolase promoter with AclI cut site
5514-MK38-pTC3-THI-AclI-R	ATTAACGTTATCTACGACCTGGGAAACATG	Amplify pTC3 thiolase terminator with AclI cut site
5422-MK25-pBiP-ClaI-F	ATTATCGATCCCTGTTCACTCCGTGTAG	*Y. lipolytica BiP* promoter with 5' ClaI cut site
5423-MK26-BiPt-ClaI-R	ATTATCGATCAGGTTCAGCTCCGAAGC	*Y. lipolytica BiP* terminator with 3' ClaI cut site

5430-MK33-pBiP-ss-GFP-HDEL-R	AGTTCTTCTCCTTTGCTAGCATCAGCCTGAACGCCGGC	*GFP-HDEL* construction, rev
5431-MK34-pBiP-ss-GFP-HDEL-F	GCCGGCGTTCAGGCTGATGCTAGCAAAGGAGAAGAACT	*GFP-HDEL* construction, fwd
5432-MK35-BiPt-ss-GFP-HDEL-R	AAACCTT**TTAAAGCTCATCGTGTTTTTT**GTAGAGCTCATCCATGC	*GFP-HDEL* construction, rev
5433-MK36-BiPt-ss-GFP-HDEL-F	GCATGGATGAGCTCTACAAACACGATGAGCTTTAA**AAGGTTT**	*GFP-HDEL* construction, fwd
pJET1.2 forward sequencing oligo	CGACTCACTATAGGGAGAGCGGC	pJET1.2 5'-end sequencing oligo, fwd
pJET1.2 reverse sequencing oligo	AGAACATCGATTTTCCATGGCAG	pJET1.2 3'-end sequencing oligo, rev

Cut sites for restriction endonucleases are underlined.
Fluorescence sequences are colored.
Primers for overlapping PCR are bolded.

Table 2-9. Common solutions

Solution	Composition	Reference
1 X PBS	137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 1.5 mM K_2HPO_4, pH 7.3	Pringle et al., 1991
1 X TBST	20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.05% (w/v) Tween 20	Huynh et al., 1985
1 × transfer buffer	20 mM Tris-HCl, 150 mM glycine, 20% (v/v) methanol	Towbin et al., 1979; Burnette, 1981
5 × SDS-PAGE running buffer	0.25 M Tris-HCl, pH 8.8, 2 M glycine, 0.5% SDS	Ausubel et al., 1989
10 × TBE	0.89 M Tris-borate, 0.89 M boric acid, 0.02 M EDTA	Maniatis et al., 1982
2 × sample buffer	20% (v/v) glycerol, 167 mM Tris-HCl, pH 6.8, 2% SDS, 0.005% bromophenol blue	Ausubel et al., 1989
6 × DNA loading dye	0.25% bromophenol blue, 0.25% xylene cyanol FF, 30% (v/v) glycerol	Maniatis et al., 1982
breakage buffer, yeast	2% (v/v) Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0	Maniatis et al., 1982
Ponceau stain	0.1% Ponceau S, 1% TCA	Szilard, 2000
TE	10 mM Tris-HCl, pH 7.0-8.0 (as needed), 1 mM EDTA	Maniatis et al., 1982

2.2 Microorganisms and culture conditions

2.2.1 Bacterial strains and culture conditions

The bacterial strains and culture medium used in this study are listed in Table 2-10 and Table 2-11, respectively. Cells of *E. coli* strain DH5α were grown in culture tubes in a rotary shaker at 200 rpm at 37°C. Culture volumes were approximately 20% of flask volumes. For short term storage, bacterial cultures were maintained on agar plates at 4°C for approximately 1 week.

Table 2-10. Bacterial strains

Strain	Genotype	Source
DH5α	F^-, Φ80d*lacZ*ΔM15, Δ(*lacZYA-argF*), U169, *rec*A1, *end*A1, *hsd*R17 (r_K^-, m_K^+), *pho*A, *sup*E44, λ^-, *thi*-1, *gyr*A96, *rel*A1	Invitrogen

Table 2-11. Bacterial culture medium

Medium	Composition	Reference
LB[a]	1% tryptone, 0.5% yeast extract, 1% NaCl	Maniatis et al., 1982

[a]For solid medium, agar was added to 2%.

2.2.2 Yeast strains and culture conditions

The *Y. lipolytica* strains used in this study are listed in Table 2-12. The yeast culture media used in this study are listed in Table 2-13. Yeasts were grown on agar plates at 28°C or in liquid cultures at 30°C, unless otherwise indicated. Cultures of 10 mL or less were grown in 16 × 150-mm glass tubes in a rotating wheel. Cultures greater than 10 mL were grown in flasks in a rotary shaker at 200 rpm. Culture volumes were approximately 20% of flask volumes. For short term storage, yeast cultures were maintained on agar plates at 4°C for 1-6 months. For long term storage, glycerol was added to liquid yeast cultures at a 30% (v/v) final concentration, and the suspension was frozen and stored at -80°C.

Table 2-12. *Yarrowia lipolytica* strains

Strain	Genotype	Reference
E34	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2)*	Chang *et al.*, 2009
E122	*MATα, ura3-302, leu2-270, lys8-11*	C. Gaillardin
*pex11*Δ	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pex11::URA3*	this study
*pex11C*Δ	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pex11C::URA3*	this study
pex11/25Δ	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pex11/25::URA3*	this study
E34-pTC3-*PEX11-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pTC3-PEX11-mCherry (URA3)*	this study
E34-pTC3-*PEX11C-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pTC3-PEX11C-mCherry (URA3)*	this study
E34-pTC3-*PEX11/25-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pTC3-PEX11/25-mCherry (URA3)*	this study
E34-pTC3	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pTC3 (URA3)*	this study
E122-pUB4-*mRFP-SKL*	*MATα, ura3-302, leu2-270, lys8-11, pUB4-mRFP-SKL (HygBR)*	this study
*pex11*Δ-pUB4-*mRFP-SKL*	*MATα, ura3-302, leu2-270, lys8-11, pex11::URA3, pUB4-mRFP-SKL (HygBR)*	this study
E34-pUB4-*mRFP-SKL*	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pUB4-mRFP-SKL (HygBR)*	this study
*pex11*Δ-pUB4-*mRFP-SKL*	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pex11::URA3, pUB4-mRFP-SKL (HygBR)*	this study
*pex11C*Δ-pUB4-*mRFP-SKL*	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pex11C::URA3, pUB4-mRFP-SKL (HygBR)*	this study
*pex11/25*Δ-pUB4-*mRFP-SKL*	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pex11/25::URA3, pUB4-mRFP-SKL (HygBR)*	this study
*pex11*Δ-pUB4-*PEX11*	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pex11::URA3, pUB4-PEX11 (HygBR)*	this study
*pex11*Δ-pUB4-*PEX11C*	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pex11::URA3, pUB4-PEX11C (HygBR)*	this study
*pex11*Δ-pUB4-*PEX11/25*	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pex11::URA3, pUB4-PEX11/25 (HygBR)*	this study

*pex11*Δ-pUB4	*MATα, ura3-302, leu2-270, lys8-11, pot1::POT1-GFP (LEU2), pex11::URA3, pUB4 (HygBR)*	this study
E122-pUB4-*PEX11-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pUB4-PEX11-mCherry (HygBR)*	this study
*pex3*Δ-pUB4-*PEX11-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pex3::URA3, pUB4-PEX11-mCherry (HygBR)*	this study
*pex16*Δ-pUB4-*PEX11-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pex16::LEU2, pUB4-PEX11-mCherry (HygBR)*	this study
*pex19*Δ-pUB4-*PEX11-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pex19::URA3, pUB4-PEX11-mCherry (HygBR)*	this study
E122-pTC-*NAT-PEX1C-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pTC3-NAT-PEX11C-mCherry (NATR)*	this study
*pex3*Δ-pTC3-*NAT-PEX11C-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pex3::URA3, pTC3-NAT-PEX11C-mCherry (NATR)*	this study
*pex16*Δ-pTC3-*NAT-PEX11C-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pex16::LEU2, pTC3-NAT-PEX11C-mCherry (NATR)*	this study
*pex19*Δ-pTC3-*NAT-PEX11C-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pex19::URA3, pTC3-NAT-PEX11C-mCherry (NATR)*	this study
E122-pUB4-*PEX11/25-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pUB4-PEX11/25-mCherry (HygBR)*	this study
*pex3*Δ-pUB4-*PEX11/25-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pex3::URA3, pUB4-PEX11/25-mCherry (HygBR)*	this study
*pex16*Δ-pUB4-*PEX11/25-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pex16::LEU2, pUB4-PEX11/25-mCherry (HygBR)*	this study
*pex19*Δ-pUB4-*PEX11/25-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pex19::URA3, pUB4-PEX11/25-mCherry (HygBR)*	this study
E122-pUB4-*PEX3B-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pUB4-PEX3B-mCherry (HygBR)*	this study
*pex11*Δ-pUB4-*PEX3B-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pex11::URA3, pUB4-PEX3B-mCherry (HygBR)*	this study
E122-pUB4-*PEX14-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pUB4-PEX14-mCherry (HygBR)*	this study
*pex11*Δ-pUB4-*PEX14-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pex11::URA3, pUB4-PEX14-mCherry (HygBR)*	this study
*pex11*Δ-pTC3-*NAT-PEX11C-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pex11::URA3, pTC3-NAT-PEX11C-mCherry (NATR)*	this study
*pex11*Δ-pUB4-*PEX11/25-mCherry*	*MATα, ura3-302, leu2-270, lys8-11, pex11::URA3, pUB4-PEX11/25-mCherry (HygBR)*	this study
E122-pINA445-*GFP-HDEL*	*MATα, ura3-302, leu2-270, lys8-11, pINA445-GFP-HDEL (LEU2)*	this study
*pex11*Δ-pINA445-*GFP-HDEL*	*MATα, ura3-302, leu2-270, lys8-11, pex11::URA3, pINA445-GFP-HDEL (LEU2)*	this study
*pex3*Δ-pINA445-*GFP-HDEL*	*MATα, ura3-302, leu2-270, lys8-11, pex3::URA3, pINA445-GFP-HDEL (LEU2)*	this study

*pex16*Δ-pINA443-*GFP-HDEL*	*MATα, ura3-302, leu2-270, lys8-11, pex16::LEU2, pINA443-GFP-HDEL (URA3)*	this study
*pex19*Δ-pINA445-*GFP-HDEL*	*MATα, ura3-302, leu2-270, lys8-11, pex19::URA3, pINA445-GFP-HDEL (LEU2)*	this study

[a]A gift from C. Gaillardin (Thiverval-Grignon, France).

Table 2-13. Yeast culture media

Medium	Composition	Reference
non-fluorescent medium	6.61 mM KH_2PO_4, 1.32 mM K_2HPO_4, 4.06 mM $MgSO_4 \cdot 7H_20$, 26.64 mM $(NH_4)SO_4$, 1 × CSM, 2% glucose	Fagarasanu et al., 2009
non-fluorescent, non-fermentative medium	6.61 mM KH_2PO_4, 1.32 mM K_2HPO_4, 4.06 mM $MgSO_4 \cdot 7H_20$, 26.64 mM $(NH_4)SO_4$, 1 × CSM, 0.6 M sorbitol, 1% glycerol	Mast, 2013
SM[a]	0.67% YNB, 2% glucose, 1× CSM without leucine or uracil as required	Tam et al., 2005
YEPA	1% yeast extract, 2% peptone, 2% sodium acetate	Brade, 1992
YEPD[a]	1% yeast extract, 2% peptone, 2% glucose, supplemented with 125 mg hygromycin/mL or 100 mg nourseothricin/mL as required	Rose et al., 1988
YNO	1.34% yeast nitrogen base without amino acids, 0.05% (w/v) Tween 40, 0.2% (w/v) oleic acid	Nuttley et al., 2003
YPBO[a]	0.3% yeast extract, 0.5% peptone, 0.5% K_2HPO_4, 0.5% KH_2PO_4, 0.2% (w/v) Tween 40 or 1% (v/v) Brij 35, 1% (v/v) oleic acid	Kamiryo et al., 1982

[a]For solid medium, agar was added to 2%.

2.3 DNA manipulation and analysis

All reactions were carried out in 1.5 mL microcentrifuge tubes, and microcentrifugation was performed in an Eppendorf microcentrifuge at 16,000 × g, unless otherwise indicated.

2.3.1 Amplification of DNA by the polymerase chain reaction (PCR)

PCR was used to amplify specific DNA sequences from chromosomal or plasmid DNA and sometimes to introduce modifications in the amplified DNA sequence including recognition sites for restriction endonucleases or intracellular targeting sequences for encoded proteins. Primer design, reaction conditions and cycling conditions followed standard protocols (Innis and Gelfand, 1990; Saiki, 1990). A reaction usually contained 0.1 to 0.5 µg of yeast genomic DNA or 0.1 to 0.2 µg of plasmid DNA to act as a template for the reaction. Each reaction also contained 25 pmol each of a

forward and reverse primer, 0.25 mM of each dNTP, and 1 µL (5 U) of Easy-A high-fidelity polymerase in the supplied reaction buffer (Stratagene). Alternatively, when fidelity of the reaction was not critical, 1 µL (5 U) of Taq polymerase in the supplied reaction buffer (NEB) were used as recommended by the manufacturer. Reactions were performed in 0.6-ml microcentrifuge tubes in a Robocycler 40 with a Hot Top attachment (Stratagene) or in 0.2 mL microfuge tubes in a 2720 Thermocycler (Applied Biosystems).

2.3.2 Digestion of DNA by restriction endonucleases

Approximately 1 to 2 µg of plasmid DNA or purified DNA were digested by restriction endonucleases for 1.5 h to 3 h according to the manufacturer's instructions (NEB). Digestion was immediately terminated by agarose gel electrophoresis of the DNA fragments.

2.3.3 Dephosphorylation of 5'-ends

Following linearization by digestion with a single restriction endonuclease, plasmid DNA was subjected to dephosphorylation at its 5'-ends to prevent intramolecular ligation. Reactions were mixed with 1 unit of CIP per 1 pmol of DNA ends (approximately 1 µg of a 3 kb plasmid) and incubated at 37°C for 1 h. The dephosphorylation reaction was terminated by agarose gel electrophoresis of the plasmid DNA.

2.3.4 Blunting 5'- and 3'-ends with DNA Polymerase I, Large (Klenow) Fragment

Following digestion by restriction endonucleases, the ends of plasmid DNA and purified DNA were sometimes blunted by 3'-overhang removal and fill in of 5'-overhangs. DNA fragments were solubilized in 1 x NEBuffer 1-4 supplemented with 33 µM each dNTP, and 1 unit of Klenow fragment per µg of DNA was added. The reaction was incubated at 25°C for 15 min and terminated by agarose gel electrophoresis of the DNA fragments.

2.3.5 Separation of DNA fragments by agarose gel electrophoresis

DNA fragments in solution were mixed with 0.2 volume of 6 × DNA loading dye (Table 2-4) and separated by electrophoresis in 0.8% agarose gels in 1 × TBE (Table 2-4) containing 0.5 µg of ethidium bromide/mL. Gels were subjected to electrophoresis at 10 V/cm in 1 × TBE, and DNA fragments were subsequently visualized on an ultraviolet transilluminator (Photodyne, Model 3-3006).

2.3.6 Purification of DNA fragments from agarose gel

When purification of a DNA fragment was required, *e.g.* after digestion of DNA by restriction endonuclease, it was separated by electrophoresis in an agarose gel, and the desired band was excised with a razor blade. DNA was purified from the excised agarose slice using the QIAquick Gel Extraction Kit according to the manufacturer's instructions (Qiagen). This method is based on the dissolution of agarose gel and subsequent selective adsorption of DNA to a silica membrane in the presence of a high concentration of chaotropic salts, followed by washing and elution of DNA in the presence of low salt. DNA was eluted in 30 to 50 µL of the supplied elution buffer.

2.3.7 Purification of DNA from solution

Contaminants (oligonucleotides, salts, enzymes, *etc.*) were removed from a DNA solution using the QIAquick PCR Purification Kit according to the manufacturer's instructions (Qiagen). The principle of this method is the selective adsorption of DNA to a silica membrane in the presence of a high concentration of chaotropic salts, followed by washing and elution of DNA in the presence of low salt. DNA was eluted in 30 to 50 µL of the supplied elution buffer.

2.3.8 Ligation of DNA fragments

DNA fragments digested with restriction endonuclease (Section 2.3.2) and purified from agarose gel (Section 2.3.6) were ligated using 1-2.5 µL (~400-1000 U) of T4 DNA ligase in the buffer supplied by the manufacturer (NEB). Reactions were typically conducted in a volume of 10-25 µL, with the molar ratio of plasmid to insert varying between 1:1 and 1:10, and incubated at room temperature

for 1.5 h or overnight at 16°C. Occasionally, PCR products after purification by agarose gel electrophoresis were ligated with the vector pJET1.2/blunt using the CloneJET PCR Cloning Kit according to the manufacturer's instructions (Thermo Scientific).

2.3.9 DNA sequencing

DNA sequencing was performed using the BigDye Terminator v1.1/3.1 Cycle Sequencing Ready Reaction Kit as described by the manufacturer (Applied Biosystems). This method involves the random incorporation of fluorescent dideoxy terminators during the elongation of DNA sequences with a modified version of Taq DNA polymerase (Sanger et al., 1977). Essentially, a reaction contained 1 ng of plasmid DNA, 3.2 pmol of primer, 3 µL of Terminator Ready Reaction Mix, and 4 µL of the supplied 5 × buffer in a total volume of 20 µL. The reaction was subjected to cycle sequencing using the Robocycler 40 with a Hot Top attachment under the following conditions: 1 cycle at 96°C for 2 min; 25 cycles at 96°C for 46 sec, 50°C for 51 sec and 60°C for 4 min 10 sec; 1 cycle at 6°C and hold. Reaction products were precipitated by addition of 80 µL of 75% (v/v) 2-propanol for 30 min at room temperature, pelleted by microcentrifugation at 16,000 × g for 20 min, washed with 250 µL of 75% 2-propanol, dried in a rotary vacuum dessicator, solubilized in 15 µL of Template Suppression Reagent, briefly vortexed, and finally separated by capillary electrophoresis. Fluorescence was detected and recorded by an ABI 310 Genetic Analyzer (Applied Biosystems).

2.4 Introduction of DNA into microorganisms

2.4.1 Chemical transformation of E. coli

Plasmid DNA was introduced into Subcloning Efficiency, chemically competent *E. coli* DH5α cells, as recommended by the supplier (Invitrogen). 5 µL of ligation reaction (Section 2.3.8) or 0.5 µl (~0.25 µg) of plasmid DNA was added to 25 µL of cells. The mixture was incubated on ice for 30 min, subjected to a 30 sec heat shock at 37 °C, and incubated on ice for 2 min. 975 µL of LB medium (Table 2-11) were added, and the cells were incubated in a rotary shaker at 200 rpm for 60 min at

37°C. Cells were spread onto LB agar plates (Table 2-11) containing 100 mg ampicillin/mL and incubated overnight at 37°C.

2.4.2 *Chemical transformation of* Y. lipolytica

Autonomously replicating plasmid DNA was introduced into yeast as previously described (Gietz and Woods, 2002). Essentially, yeast cells were grown overnight in 5 mL of YEPD and subcultured the following morning in 5 mL of fresh YEPD. After 4-5 h, or until the culture reached an OD_{600} of ~1.0, cells were harvested by centrifugation for 3 min at 2,500 × g. Cells were resuspended in 1 ml of 0.1 M lithium acetate and vortexed. Cells were harvested by microcentrifugation, and the following components were added in order to the cell pellet: 10 µL of 5 mg sheared single-stranded salmon sperm DNA/mL in 40 µL of water, 2.5 µL (~10 pg) plasmid DNA, 500 µL of lithium acetate/polyethylene glycol (PEG) solution (8 mL 50% PEG, 1 mL 1M lithium acetate, 50 µL 0.5 M EDTA, pH 8.0, 0.95 mL water), and 53 µL of DMSO. The cell pellet was resuspended in the mixture by use of a pipet, and the resuspension was mixed by inversion following the addition of DMSO. After a 15 min incubation at room temperature, the cells were subjected to a 15 min heat shock at 42°C. Cells were collected by microcentrifugation and washed by gentle resuspension in 1 mL of water before being harvested by microcentrifugation and resuspended in 80 µL of water. The cells were plated onto SM agar plates or onto YEPD agar plates supplemented with the appropriate antibiotic (Table 2-12). Plates were incubated at 30°C for 2-3 days for colony formation.

2.4.3 *Electroporation of* Y. lipolytica

Y. lipolytica cells were made electrocompetent as previously described (Ausubel et al., 1989). Essentially, yeast cells were grown overnight in 5 mL of YEPA and subcultured the following morning in 10 mL of fresh YEPA. After 4-5 h, or until the culture reached an OD_{600} of ~1.0, cells were harvested by centrifugation for 3 min at 2,500 × g. The pellet was resuspended in 9 mL of TE buffer, pH 7.5 (Table 2-9) containing 1 mL of 1 M lithium acetate, and incubated for 30 min at 30°C with gentle agitation. DTT was added to a final concentration of 20 mM, and the incubation was

continued for another 15 min. Cells were harvested by centrifugation at 2,500 × g, washed successively with 5 mL each of room-temperature water, ice-cold water, and ice-cold 1 M sorbitol. After washing, cells were resuspended in a minimal volume of ice-cold 1 M sorbitol. 40 µL of cells were mixed with 2 µL of plasmid DNA, placed between the bosses of an ice-cold microelectroporation chamber (Bio-Rad), and subjected to an electrical pulse of 250 V (amplified to ~1.6 kV) at a capacitance of 2 µF and a resistance of 4 kΩ using a Bio-Rad MicroPulser. Cells were immediately resuspended in 100 µL of ice-cold 1 M sorbitol and plated onto selective SM agar plates or onto YEPD agar plates supplemented with appropriate antibiotic (Table 2-12). Plates were incubated at 30°C for 2-3 days for colony formation.

2.5 Isolation of DNA from microorganisms

2.5.1 Isolation of plasmid DNA from bacteria

Single bacterial colonies were inoculated into either 2 mL of LB containing ampicillin and incubated for 8 h at 200 rpm at 37°C, or into 3 mL of LB containing ampicillin and incubated overnight at 200 rpm at 37°C. Cells were harvested by microcentrifugation, and plasmid DNA was isolated using a QIAprep Spin Miniprep Kit according to the manufacturer's instructions (Qiagen). This method is based on the alkaline lysis of bacterial cells, followed by adsorption of DNA onto silica in the presence of high salt and elution of DNA in low salt buffer. Plasmid DNA was eluted in 30-50 µL of the supplied elution buffer.

2.5.2 Isolation of chromosomal DNA from yeast

Yeast genomic DNA was prepared as previously described (Ausubel et al., 1989). Cells were grown overnight in 10 mL of YEPD, harvested by centrifugation for 3 min at 2,000 rpm at room temperature, and washed once with 10 mL of water. Cells were resuspended in 200 µL of breakage buffer (Table 2-9) and transferred to a 1.5 mL microcentrifuge tube. ~200 µL of glass beads, 100 µL of phenol, and 100 µL of chloroform/isoamyl alcohol (24:1) were successively added, and the mixture was vortexed for 4 min to lyse the cells. 200 µL of TE (Table 2-9) were added,

and the mixture was vortexed and then subjected to microcentrifugation at 16,000 × g for 5 min at room temperature. The upper aqueous phase was removed and re-extracted by successively adding 100 µL of phenol and 100 µL of chloroform/isoamyl alcohol (24:1) followed by vortexing for 4 min. The organic and aqueous phases were separated by centrifugation at 16,000 × g for 10 min at room temperature, and the upper aqueous phase was removed. If necessary, additional extractions of the aqueous phase with an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) were performed to improve the purity of the DNA. 1 mL of absolute ethanol was added to the aqueous phase and mixed by inversion to precipitate the DNA. The DNA was pelleted by centrifugation at 16,000 × g for 5 min at room temperature. The pellet was washed once with 1 ml of 70% (v/v) ethanol and subjected to centrifugation at 16,000 × g for 5 min at room temperature. The pellet was air dried and solubilized in 50 µl of TE 8.0 containing 50 µg RNase A/mL. DNA was incubated at 37°C for 3 h to allow for digestion of RNA, with occasional mixing by inversion.

2.6 Protein manipulation and analysis

2.6.1 Preparation of yeast whole cell lysates

Denatured yeast whole cell lysates were prepared by harvesting ~2.0 OD_{600} of yeast cells from an overnight 10 mL culture by centrifugation at 16,000 × g for 1 min. The pellet was resuspended in 240 µL of alkali solution containing reducing agent (1.85 M NaOH, 7.4% (v/v) 2-mercaptoethanol) and incubated on ice for 5 min. 240 µL of 50% TCA were added and mixed by vortexing, and incubated on ice for an additional 5 min. The solution was subjected to microcentrifugation at 16,000 × g for 10 min at 4°C. The pellet was washed once with ice-cold water and then resuspended successively in 50 µL of Magic A (1 M unbuffered Tris, 13% SDS) and 50 µL of Magic B (30% (v/v) glycerol, 200 mM DTT, 0.25% (w/v) bromophenol blue). The mixture was incubated for 15 min at 65°C and then subjected to microcentrifugation at 16,000 × g for 1 min. The supernatant was collected and resolved by SDS-PAGE.

2.6.2 Precipitation of proteins

Proteins were precipitated from solution by addition of 100% TCA to a final concentration of 10%, vortexing and incubation on ice overnight. Precipitates were collected by microcentrifugation at 16,000 × g for 20 min at 4°C. The pellet was resuspended in 1 mL of ice-cold 80% acetone, incubated on ice for 20 min, and subjected to microcentrifugation at 16,000 × g for 20 min at 4°C.. The pellet was then dried in a rotary vacuum dessicator and dissolved in an appropriate volume of 2 × sample buffer (Table 2-9).

2.6.3 Determination of protein concentration

The protein concentration of a sample was determined by the method of Lowry (Lowry et al., 1951), as simplified by Peterson (Peterson, 1979). This is a colorimetric assay based on the reaction between protein and an alkaline copper tartrate solution, followed by reduction of Folin reagent. Typically, 100 µL of sample, *e.g.* 20KgP subcellular fraction, was centrifuged at 20,000 × g for 5 min at 4°C, and proteins were precipitated by addition of 240 µL of alkali solution containing reducing agent (1.85 M NaOH, 7.4% (v/v) 2-mercaptoethanol) and incubation on ice for 5 min. Aliquots of the lysate were taken in duplicate (usually one of 10 µL and a second of 20 µL) and made up to 1 mL with water. 100 µL of 0.15% sodium deoxycholate were added, and the mixture was incubated for 10 min at room temperature. 100 µL of 72% TCA were added, and the solution was vortexed and centrifuged at 16,000 × g for 15 min at room temperature. The pellet was dissolved in 400 µL of water, and 100 µL of the sample were subsequently used in the protein assay.

Protein concentration of a sample was sometimes determined using the *DC* Protein Assay kit according to the manufacturer's instructions (Bio-Rad). Protein standards containing 0.0, 0.2, 0.4, 0.6, 0.8, 1.0, and 1.5 mg BSA/mL in 100 µL of water were prepared. Samples were incubated for 15 min at room temperature, and absorbance was measured at 750 nm using a Beckman DU640 spectrophotometer. Absorbance values were plotted against the different amounts of BSA to generate a standard curve. Absorbance of a protein sample was measured as for the BSA standards,

and the protein concentration was estimated by comparing the absorbance value with the standard curve.

2.6.4 Separation of proteins by electrophoresis

Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) as previously described (Ausubel et al., 1989). Protein samples were mixed with an equal volume of 2 × sample buffer containing 10 mM 2-mercaptoethanol, denatured by boiling for 10 min or incubating at 70°C for 20 min, and separated by electrophoresis on discontinuous slab gels. Stacking gels contained 3% acrylamide (30:0.8 acrylamide:N,N'-methylene-bis-acrylamide), 60 mM Tris-HCl, pH 6.8, 0.1% SDS, 0.1% (v/v) TEMED, 0.1% ammonium persulfate. Resolving gels contained 10% acrylamide (30:0.8 acrylamide:N,N'-methylene-bis-acrylamide), 370 mM Tris-HCl, pH 8.8, 0.1% SDS, 0.1% (v/v) TEMED, 0.043% ammonium persulfate. Electrophoresis was conducted in 1 × SDS-PAGE running buffer (Table 2-4) at 100-200 V using a Bio-Rad Mini Protean III vertical gel system.

2.6.5 Detection of proteins by gel staining

Proteins in polyacrylamide gels were visualized by staining with 0.1% Coomassie Brilliant Blue R-250, 10% (v/v) acetic acid, 35% (v/v) methanol for 1 h with gentle agitation. Unbound dye was removed by multiple washes in 10% (v/v) acetic acid, 35% (v/v) methanol.

2.6.6 Detection of proteins by immunoblotting

Proteins separated by SDS-PAGE were transferred to nitrocellulose membrane (Bio-Rad) in 1 × transfer buffer (Table 2-9) using a Trans-Blot tank transfer system with plate electrodes (Bio-Rad) at 100 V for 1 h at 4°C. Proteins transferred to nitrocellulose were visualized by staining in Ponceau S (Table 2-9) for 1-5 min. The nitrocellulose was destained in water, washed with 1 x TBST (Table 2-9) for 10 min with gentle agitation, and incubated in blocking solution (5% skim milk in 1 x TBST) twice for 5 min each and then for 1 h with gentle agitation. Proteins of interest were detected by incubation with primary antibody in blocking solution for 1-3 h at room temperature, or overnight at 4°C, with gentle agitation. Following incubation with primary antibody, unbound antibodies were

removed by washing the nitrocellulose twice in blocking solution, once in 1 x TBST, and twice in blocking solution (all 5 min each). The nitrocellulose was incubated with an appropriate HRP-labeled secondary antibody in blocking solution for 1 h. Following incubation with secondary antibody, unbound antibodies were removed by washing the nitrocellulose twice in blocking solution and three times in 1 x TBST (all 5 min each). Antigen-antibody complexes were detected using an ECL Western Blotting Detection Kit according to the manufacturer's instructions (GE Healthcare or Thermo Scientific) and exposing the nitrocellulose to X-Omat BT film (Kodak).

Used nitrocellulose could be reblotted using a Re-Blot Western Blot Recycling Kit according to the manufacturer's instructions (Chemicon). The nitrocellulose was incubated with 1 × Antibody Stripping Solution at room temperature for 10 to 30 min with gentle agitation, washed with 1 × TBST for 10 min and blocking solution twice for 5 min each, and blotted as described above.

2.7 Affinity purification of polyclonal antibodies

Affinity purification of antibodies against Pex2p using purified inclusion bodies was performed by Elena Savidov, Department of Cell Biology, University of Alberta. Production and purification of fusion proteins were done using the pMAL Protein Fusion and Purification System according to the manufacturer's instructions (NEB). This method is based on the induction of fusion protein synthesis by the addition of IPTG. The plasmid pMAL-c2-*PEX2* (Eitzen et al., 1996) was transformed into DH5α cells, and a single bacterial colony was inoculated into 100 mL of LB containing ampicillin and and incubated overnight at 200 rpm at 37°C. The following morning, 1 mL of culture was subcultured into 1 L of LB containing ampicillin, and bacteria were grown to an OD_{600} ~0.7. 1 mL of 1 M IPTG was added to the culture to a final concentration of 1 mM, and the culture was incubated for 3 h at room temperature at 200 rpm. The culture was cooled on ice for 15 min, cells were harvested by centrifugation at 2,000 × g for 30 min at 4°C, and the cell pellet was stored at -80°C for later purification of inclusion bodies.

To confirm the expression of a fusion protein upon IPTG induction, 1 mL samples of culture taken before and after IPTG induction were harvested by microcentrifugation at 16,000 × g for 1 min at 4°C, and the pellet was washed twice with 1 X PBS. Pellets were mixed with an equal volume of 2 × sample buffer containing 10 mM 2-mercaptoethanol, and proteins were denatured by boiling for 10 min, resolved by SDS-PAGE, and stained with Coomassie Brilliant Blue R-250 to confirm expression of the correctly sized MBP fusion protein. Pex2p has a predicted molecular mass of 41.72 kDa, and its fusion with MBP has a molecular mass of approximately 80 kDa.

For large-scale production of an MBP-fusion protein, the *E. coli* pellet from 1 L of culture was resuspended in 13 mL of solution buffer, pH 8.0 (50 mM Tris-HCl, 25% sucrose, 1 mM sodium EDTA, 0.1% sodium azide, 10 mM DTT). Cells were lysed on ice by sonication with a Branson Sonifer 250 set to a duty of 50% and an output control of 4-5, (approximately 30 pulses). 100 µL of lysozyme (50 mg/mL in water), 250 µL of DNase I (1 mg/mL in 50% glycerol, 75 mM sodium chloride), and 50 µL of 0.5 M magnesium chloride were added to the lysate and mixed by vortexing. 12.5 mL of lysis buffer, pH 8.0 (50 mM Tris-HCl, 1% Triton X-100, 1% sodium deoxycholate, 100 mM sodium chloride, 0.1% sodium azide, 10 mM DTT) were added to the lysate and mixed by vortexing, and the lysate was incubated for 1 h at room temperature. 350 µL of 0.5 M sodium EDTA were added, and the lysate was flash frozen in liquid nitrogen and thawed at room temperature for 30 min. 200 µL of 0.5 M sodium chloride were added, followed by incubation of the lysate at room temperature for 1 h and centrifugation at 11,000 × g for 20 min at 4°C. The pellet was resuspended in 10 mL of washing buffer with Triton X-100, pH 8.0 (50 mM Tris-HCl, 1% Triton X-100, 1% sodium deoxycholate, 100 mM sodium chloride, 0.1% sodium azide, 1mM DTT), lysed on ice by sonication with a Branson Sonifer 250 set to a duty of 50 % and an output control of 4-5 (approximately 30 pulses) and subjected to centrifugation at 11,000 × g for 20 min at 4°C. The pellet was resuspended in 10 mL of washing buffer, pH 8.0 (50 mM Tris-HCl, 1% sodium deoxycholate, 100 mM sodium chloride, 0.1% sodium azide, 1mM DTT), followed by sonication with a Branson Sonifer 250 set to a duty of 50 % and an output control of 4-5 (approximately 30

pulses) and centrifugation at 11,000 × g for 20 min at 4°C. The pellet containing purified inclusion bodies was resuspended in an appropriate volume of 2 x sample buffer and resolved by SDS-PAGE, and proteins were transferred to nitrocellulose membrane and strained with Ponceau S to confirm the correct size of MBP-fusion protein.

Strips of nitrocellulose membrane containing the separated MBP-fusion protein were excised and incubated in 3% BSA in TBST for 1 h at room temperature with gentle agitation. Strips were washed twice with 1 X TBST for 5 min each and incubated with 2 mL of antiserum specific to the Pex2p protein (guinea pig Pay5-NN, final bleed) diluted with 10 mL of 1 X TBST for 3 h at 4°C with gentle agitation. The strip was washed several times with 1 X TBST to remove unbound antibodies, and all traces of washing buffer were aspirated. Affinity-purified antibody was eluted by three sequential incubations of the strip with 1 mL glycine buffer (0.1 M glycine-HCl, pH 2.3) for 3 min each, each time collecting the eluant in a tube containing 250 µL of 2 M Tris, pH 8.5, to neutralize the glycine buffer. The eluted antibody was concentrated on an Amicon Ultra-50K filter device (Millipore), the elution buffer was replaced with 50 mM Tris, pH 7.5, and the filter device was subjected to centrifugation at 3000 rpm. Glycerol was added to the affinity purified solution to a final volume of 25%, and the affinity purified antibody was stored at -20°C. The purity and optimal dilution of the affinity purified antibody were assessed by immuoblotting whole cell lysates of wild-type and *PEX2* deletion strains. Affinity purifed Pex2p antibody was preferentially used at a dilution of 1:1,000.

2.8 Subcellular fractionation

2.8.1 Peroxisome isolation from Y. lipolytica

Cells were grown in glucose-containing medium and transferred to oleic acid-containing medium for 8 h. Subcellular fractionation and isolation of peroxisomes were performed essentially as described (Smith et al., 2002; Tam et al., 2003). Cells were harvested by centrifugation at 6000 × g in a Beckman JA10 rotor at 23°C, washed once with water, harvested again by ultracentrifugation, resuspended in 10 mM DTT, 100 mM Tris-HCl, pH 9.4, at a concentration of 10 mL per g of wet

cells, and incubated for 30 min at 70 rpm at 30°C. Cells were harvested by centrifugation at 2,000 × g in a Beckman JS13.1 rotor for 8 min at 23°C and washed once with Zymolase buffer (50 mM potassium phosphate, pH 7.5, 1.2 M sorbitol, 1 mM EDTA). Pelleted cells were resuspended in Zymolase buffer containing 1 mg of Zymolase 100T/mL at a concentration of 8 mL of Zymolyase buffer per g of wet cells and incubated for 30 min to 1.5 h at 30°C at 70 rpm to convert cells to spheroplasts. Spheroplasts were harvested by centrifugation at 1,850 × g in a Beckman JS13.1 rotor for 8 min at 23°C, washed with Zymolase buffer and harvested by centrifugation at 20,000 × g for 8 min at 4°C. Spheroplasts were resuspended in 20 mL of ice-cold buffer H (0.6 M sorbitol, 2.5 mM MES, pH 6.0, 1 mM EDTA) containing 1 × complete protease inhibitor cocktail (Roche), incubated on ice for 15 min and then transferred to a homogenization mortar and disrupted by 15-45 strokes of a Teflon pestle driven at 1,000 rpm by a stirrer motor (Model 4376-00, Cole-Parmer). Cell debris, unbroken cells, and nuclei were pelleted by centrifugation at 1,000 × g in a Beckman JS13.1 rotor for 7 min at 4°C. The postnuclear supernatant (PNS) was subjected to four additional centrifugations at 1,000 × g in a Beckman JS13.1 rotor for 7 min at 4°C. The PNS was subjected to centrifugation at 20,000 × *g* for 35 min at 4°C to yield a pellet (20KgP) fraction enriched for peroxisomes and a supernatant (20KgS) fraction enriched for cytosol. The 20KgP fraction was resuspended in buffer H containing 11% Nycodenz and 1× complete protease inhibitor cocktail, and overlaid onto a 30-mL discontinuous gradient consisting of 17%, 25%, 35% and 50% (w/v) Nycodenz, all in buffer H containing 1× complete protease inhibitor cocktail. Organelles were separated by centrifugation at 100,000 × *g* for 90 min at 4°C in a VTi50 rotor (Beckman). 2-mL fractions were collected from the bottom of the gradient. For select experiments, the 20KgS fraction was subjected to ultracentrifugation at 200,000 × *g* for 1 h at 4°C in a SW55Ti rotor (Beckman) to yield a pellet (200KgP) fraction enriched for small vesicles and a cytosolic supernatant fraction (200KgS). Equivalent proportions of each fraction were separated by SDS-PAGE and analyzed by immunoblotting.

2.8.2 Extraction and subfractionation of peroxisomes

Extraction of peroxisomes was performed essentially as described (Tam, 2005). Peroxisomes in the 20KgP fraction (approximately 100 µg of protein) were lysed by incubation in 10 volumes of ice-cold Ti8 buffer (10 mM Tris-HCl, pH 8.0) containing 2× complete protease inhibitor cocktail on ice for 2 h with occasional vortexing and 5 freeze-thaw cycles from -80°C to 4°C. Organelles were separated by ultracentrifugation at 200,000 × *g* for 1 h at 4°C in a TLA120.2 rotor (Beckman) to yield a pellet (Ti8P) fraction enriched for membrane proteins and a supernatant (Ti8S) fraction enriched for matrix proteins. The Ti8P fraction was resuspended in ice-cold Ti8 buffer to a final protein concentration of approximately 1.0 mg/mL, incubated with 10 volumes of ice-cold 0.1 M Na_2CO_3, pH 11.3, on ice for 45 min with occasional vortexing, and subjected to ultracentrifugation at 200,000 × *g* for 1 h at 4°C in a TLA120.2 rotor to yield a pellet (CO_3P) fraction enriched for integral membrane proteins and a supernatant (CO_3S) fraction enriched for peripheral membrane proteins. Proteins in the Ti8S and CO_3S fractions were precipitated by addition of TCA. Equivalent proportions of each fraction were separated by SDS-PAGE and analyzed by immunoblotting.

2.9 Electron microscopy

Y. lipolytica cells were processed for electron microscopy as previously described (Eitzen et al., 1997). All microcentrifugations were done at 16,000 × *g* for 1 min, and all incubations were done in 1.5-mL microcentrifuge tubes at room temperature with agitation, unless otherwise indicated. Cells were grown in YEPD and then transferred to YPBO for 8 h, harvested by centrifugation at 1,500 × *g* in an Allegra X15-R centrifuge (Beckman), and washed once with 1% Brij 35. Approximately 100 µL of cells were fixed in 1 mL of 3% $KMnO_4$ for 20 min, harvested by centrifugation, and washed twice with water. Cells were incubated in 1 mL of 1% sodium periodate for 10 min, harvested by centrifugation, and washed once with water, and then incubated in 1 mL of 1% NH_4Cl for 10 min, harvested by centrifugation, and washed once with water. Cells were next subjected to serial dehydration in 60%, 80%, 95% and 100% ethanol, and in propylene oxide. Each

incubation was for 5 min, followed by centrifugation and removal of the supernatant. Incubation in 100% ethanol was repeated twice and incubation in propylene oxide was repeated three times. Cells were harvested by centrifugation and incubated in 1 mL of a 1:1 mixture of propylene oxide and resin (10 g TAAB 812 resin, 7.448 g DDSA, 3.32 g MNA, 1% DMP-30; proportions suggested by the manufacturers (Canemco and Marivac)) for 30 min with agitation, followed by a 30 min period with opened microcentrifuge caps. Cells were harvested by centrifugation and resuspended in 1 mL of resin without DMP-30. Cells were then incubated for 1 h with agitation and 3 h in the fume hood with open microcentrifuge caps. Cells were harvested by centrifugation for 8 min, resuspended in 1 mL of resin, and incubated overnight in a dark room with constant agitation. Cells were harvested by centrifugation for 8 min, resuspended in resin with 2% DMP-30, and incubated for 2 h with agitation. Small portions of cells were transferred to embedding capsules (EMS) containing resin with 2% DMP-30. Embedding capsules were placed in an oven at 60°C to allow the resin to polymerize. 80-nm ultra-thin sections were cut with an Ultra-Cut E Microtome (Reichert-Jung), stained with 1% lead citrate, and visualized on a Phillips 410 electron microscope. Images were captured with a digital camera (Soft Imaging System).

Morphometric analysis of electron micrographs was performed as previously described (Tam et al., 2003). For each strain, electron micrographs of approximately 100 randomly selected cells were captured, and areas of individual cells and peroxisomes, as well as the number of peroxisomes, were determined using iTEM software (Olympus). To determine the average peroxisome area, the total peroxisome area was calculated and divided by the total number of peroxisomes counted. Numerical density of peroxisomes (number of peroxisomes per cubic micrometer of cell volume) was calculated as previously described (Weibel and Bolender, 1973). Essentially, the total number of peroxisomes was counted and reported as the number of peroxisomes per cell area assayed (NA). The peroxisome volume density (VV) was calculated as: total peroxisome area/total cell area assayed. Numerical density of peroxisomes was determined using the values of NA and VV.

2.10 Microscopy

2.10.1 3D confocal microscopy of living yeast

Cells were cultured as described in figure legends. For the images in Figures 3-7 and 3-11, 1-mL aliquots were prepared for imaging by two washes in 1 mL of water. Cells were resuspended in 20-200 µL of water, and 2 µL were placed onto a glass #1.5 coverslip for imaging. Slides were essentially prepared as previously described (Fagarasanu et al., 2009) for the images in Figure 3-8 and Figures 3-12 to 3-15. 1-mL aliquots were prepared for imaging by two washes in 1 mL of nonfluroscent medium, and cells were resuspended in 20-200 µL of nonfluroscent medium. 150 µl of hot 1% agarose in nonfluorescent medium were used to prepare a thin agarose pad on a slide with two 18-mm square wells (Cel-line Brand). 2 µl of culture were placed onto the slide, and the coverslip was sealed with Valap (1:1:1 mixture of vaseline, lanolin and paraffin). Cells were incubated for 15 min at room temperature prior to image aquisition.

Images were captured at room temperature with a Plan-Apochromat 63×/1.4 NA oil immersion DIC objective on an Axiovert 200 M inverted microscope equipped with a side-mounted LSM 510 META confocal scanner (Carl Zeiss). For the images in Figures 3-7, 3-11 and 3-12, stacks of 21 optical sections spaced 0.25 µm apart were captured. For the images in Figure 3-8 and Figures 3-13 to 3-15, a piezoelectric actuator was used to drive continuous objective movement, allowing for the rapid collection of *z*-stacks (Hammond and Glick, 2000). Stacks of 55 optical sections spaced 0.125 µm apart were captured. GFP was excited using a 488-nm laser, and its emission was collected using a 505-nm long-pass filter (Semrock). mCherry was excited using a 543-nm laser, and its emission was collected using a 610-nm long-pass filter (Semrock). For the colocalization experiment presented in Figure 3-7, GFP was excited using a 488-nm laser and its emission was collected using a 514/30 band-pass filter (Semrock), and mCherry was excited using a 543-nm laser and its emission was collected using a 600 long-pass filter (Semrock). For the colocalization experiment presented in Figure 3-11, GFP was excited using a 488-nm laser and its

emission was collected using a 514/30 nm band-pass filter, and mRFP was excited with a 543-nm laser and its emission collected with a 629/53 nm band-pass filter (Semrock).

2.10.2 Deconvolution and image processing

Images were deconvolved using algorithms provided by Huygens Professional Software (Scientific Volume Imaging BV). 3D data sets were deconvolved using an iterative Classic Maximum Likelihood Estimation confocal algorithm with an experimentally derived point spread function from latex recording beads. Imaris software (Bitplane) was used to create maximum intensity projections of the deconvolved 3D datasets. Transmission images of yeast cells were processed by applying a Gaussian filter in Huygens, and blue color was applied to the transmission images using Imaris software. Internal structures in the transmission images were removed in Photoshop to allow better visualization of the data in the fluorescent channels. Figures were assembled using Photoshop and Illustrator (Adobe).

2.10.3 Quantification of peroxisome number

For the experiment presented in Figure 3-8C, cells were grown in either YEPD or YPBO to an OD_{600} of approximately 1.0, and then prepared for imaging (Section 2.10.1). Following image acqusition and processing (Sections 2.10.1, 2.10.2), statistics on peroxisome number were calculated using the Spots algorithm in Imaris software (Bitplane).

2.11 Comparative genomic survey

2.11.1 Comparative genomic survey of the Pex11 protein family

The genomes surveyed in this study are listed in Table 2-14. This list was chosen to include genomes from all six eukaryotic supergroups and to maximize sampling from all sequenced fungal lineages, while still remaining computationally tractable.

To identify putative Pex11p, Pex11/25p, Pex25p and Pex27p homologues, *Homo sapiens, Saccharomyces cerevisiae* and/or *Yarrowia lipolytica* protein sequences (Appendix 1, 4-6) were used as queries for searches by pHMMer (Eddy, 2009) against locally hosted genomes. Initial queries

were proteins that have been functionally characterized in established model systems (for Pex11p, Pex25p and Pex27p) or were the single known homologue (for Pex11/25p). Candidate homologues with an expect (E-) value of less than or equal to 0.05 were subjected to reciprocal pHMMer searches against the locally hosted query genome, as well as reciprocal pHMMer searches against the locally hosted non-redundant (NR) genome. Thus, the retrieval of the original query (or an identically named orthologue) as the top reciprocal pHMMer hit with an E-value less than or equal to 0.05, as well as the retrieval of a named homologue as the top reciprocal pHMMer hit in the NR database, were the criteria for identification of putative homologues.

Verified Pex11p, Pex11/25p, Pex25p and Pex27p homologues were aligned using MUSCLE 3.6 (Edgar, 2004). Hidden Markov models (HMM) (Eddy, 1998) were built from the alignments and used by HMMer 3.0 (Eddy, 2009) for searches against locally hosted genomes. Candidate homologues with E-values less than or equal to 0.05 were subjected to reciprocal pHMMer searches as described above. In cases where homologues were not identified by HMMer searches, validated homologues from the same eukaryotic supergroup were used as queries for pHMMer searches, followed by reciprocal pHMMer searches as described above. Experimentally characterized Pex11p homologues from the literature were also added to the comparative genomics survey and used as queries to identify homologues in the same supergroup as described above. Newly validated homologues were added to the HMM iteratively, and HMMer searches were repeated until no additional putative homologues were identified. For genomes without available protein sequences, nHMMer (Wheeler and Eddy, 2013) was used to search locally hosted nucleotide databases with protein queries following the same procedure as described above.

To identify putative Pex11Bp and Pex11Cp homologues, protein sequences for all Pex11Bp and Pex11Cp homologues identified by Kiel and colleagues (Kiel et al., 2006) (Appendix 2-3) were aligned using MUSCLE 3.6. HMMs were built from the alignments and used by HMMer 3.0 to search locally hosted genomes. Candidate homologues were subjected to reciprocal pHMMer searches as described above. Newly validated homologues were added to the HMM

iteratively, and HMMer searches were repeated until no additional putative homologue was identified. For genomes without available protein sequences, the nucleotide sequences corresponding to the Pex11Bp and Pex11Cp query protein sequences were used to build a HMM using nHMMer. nHMMer was then used to search locally hosted nucleotide databases following the same procedure as described above.

Table 2-14. Genomes for the comparative genomic survey of the Pex11 protein family

Genome	Taxonomy Info	Source[a]	Reference
Fungi			
Acremonium alcalophilum	Ascomycota; Sodariomycetes	JGI	JGI[b]
Agaricus bisporus	Basidiomycota; Agaricomycotina	JGI	Morin et al., 2013
Allomyces macrogynus	Blastocladiomycota	Broad	Broad Institute[c]
Ascoidea rubescens	Ascomycota; Saccharomycotina; Saccharomycetales insertae serdis	JGI	JGI[b]
Ashbya gossypii	Ascomycota; Saccharomycotina; Saccharomycetaceae	NCBI	Dietrich et al., 2004
Aspergillus fumigatus	Ascomycota; Eurotiomycetes	Broad	Nierman et al., 2005; Broad Institute[d]
Aspergillus nidulans	Ascomycota; Eurotiomycetes	NCBI	Galagan et al., 2005
Aureobasidium pullulans	Ascomycota; Dothideomycetes	JGI	JGI[b]
Batrachochytrium dendrobatidis	Chytridiomycota	Broad	Broad Institute[e]
Botryotinia fuckeliana	Ascomycota; Leotiomycetes	NCBI	Broad Institute[f]
Candida albicans	Ascomycota; Saccharomycotina; mitosporic Saccharomycetales	NCBI	Jones et al., 2004
Candida glabrata	Ascomycota; Saccharomycotina; Saccharomycetaceae	NCBI	Dujon et al., 2004
Catenaria anguillulae	Blastocladiomycota	JGI	JGI[b]
Cladonia grayi	Ascomycota; Lecanoromycetes	JGI	JGI[b]
Coccidioides immitis	Ascomycota; Eurotiomycetes	Broad	Broad Institute[g]
Cochliobolus heterostrophus	Ascomycota; Dothideomycetes	JGI	Ohm et al., 2012
Coemansia reversa	Zygomycota; Kickxellomycotina	JGI	JGI[b]
Conidiobolus coronatus	Zygomycota; Entomophthoromycotina	JGI	JGI[b]
Coprinopsis cinerea	Basidiomycota; Agaricomycotina	Broad	Broad Institute[h]
Cryphonectria parasitica	Ascomycota; Sordariomycetes	JGI	JGI[b]

Cryptococcus neoformans	Basidiomycota; Agaricomycotina	Broad	Broad Institute[i]
Dacryopinax sp.	Basidiomycota; Agaricomycotina	JGI	Floudas et al., 2012
Debaryomyces hansenii	Ascomycota; Saccharomycotina; Debaryomycetaceae	NCBI	Dujon et al., 2004
Dekkera bruxellensis	Ascomycota; Saccharomycotina; Saccharomycetaceae	JGI	Piškur et al., 2012
Encephalitozoon cuniculi	Microsporidia	NCBI	Katinka et al., 2001
Gibberella zeae	Ascomycota; Sordariomycetes	NCBI	Broad Institute[j]
Gonapodya prolifera	Chytridiomycota	JGI	JGI[b]
Hanseniaspora valbyensis	Ascomyocta; Saccharomycotina; Saccharomycetaceae	JGI	JGI[b]
Hansenula polymorpha	Ascomycota; Saccharomycotina; Saccharomycetales insertae serdis	JGI	JGI[b]
Hysterium pulicare	Ascomycota; Dothideomycetes	JGI	Ohm et al., 2012
Kluyveromyces lactis	Ascomycota; Saccharomycotina; Sacccharomycetaceae	NCBI	Dujon et al., 2004
Kluyveromyces waltii	Ascomycota; Saccharomycotina; Saccharomycetaceae	JGI	Manollis et al., 2004
Laccaria bicolor	Basidiomycota; Agaricomycotina	NCBI	Martin et al., 2008
Lipomyces starkeyi	Ascomycota; Saccharomycotina; Lipomycetaceae	JGI	JGI[b]
Lodderomyces elongisporus	Ascomycota; Saccharomycotina; Debaryomycetaceae	Broad	Broad Institute[k]
Magnaporthe oryzae	Ascomycota; Sordariomycetes	Broad	Broad Institute[l]
Malassezia globosa	Basidiomycota; Ustilaginomycotina	JGI	JGI[b]
Metschnikowia bicuspidata	Ascomycota; Saccharomycotina; Metschnikowiaceae	JGI	JGI[b]
Mixia osmundae	Basidiomycota; Pucciniomycotina	JGI	Toome et al., 2013
Mucor circinelloides	Zygomycota; Mucormycotina	JGI	JGI[b]
Myceliophthora thermophila	Ascomycota; Sordariomycetes	JGI	Berka et al., 2011
Mycosphaerella graminicola	Ascomycota; Dothideomycetes	JGI	Goodwin et al., 2011; Ohm et al., 2012
Neurospora crassa	Ascomycota; Sordariomycetes	JGI	Galagan et al., 2003
Pachysolen tannophilus	Ascomycota; Saccharomycotina; Saccharomycetaceae	JGI	JGI[b]

Penicillium chrysogenum	Ascomycota; Eurotiomycetes	NCBI	van den Berg et al., 2008
Phaenerochaete chrysosporium	Basidiomycota; Agaricomycotina	JGI	Martinez et al., 2004
Phaeosphaeria nodorum	Ascomycota; Dothideomycetes	Broad	Broad Institute[m]
Phycomyces blakesleeanus	Zygomycota; Mucormycotina	JGI	JGI[b]
Pichia membranifaciens	Ascomycota; Saccharomycotina; Pichiaceae	JGI	JGI[b]
Pichia pastoris	Ascomycota; Saccharomycotina; Saccharomycetaceae	NCBI	De Schutter et al., 2009
Pichia stipitis	Ascomycota; Saccharomycotina; Debaryomycetaceae	JGI	Jeffries et al., 2007
Piromyces sp.	Neocallimastigomycota	JGI	JGI[b]
Puccinia graminis	Basidiomycota; Pucciniomycotina	Broad	Broad Institute[n]
Rhizophagus irregularis	Glomeromycota	JGI	JGI[b]
Rhizopus oryzae	Zygomycota; Mucormycotina	Broad	Ma et al., 2009
Rhodotorula graminis	Basidiomycota; Pucciniomycotina	JGI	JGI[b]
Rhytidhysteron rufulum	Ascomycota; Dothideomycetes	JGI	Ohm et al., 2012
Rozella allomycis	Cryptomycota	JGI	James et al., 2013
Saccharomyces bayanus	Ascomycota; Saccharomycotina; Saccharomycetaceae	SGD	Kellis et al., 2003
Saccharomyces castellii	Ascomycota; Saccharomycotina; Saccharomycetaceae	SGD	Cliften et al., 2003
Saccharomyces cerevisiae	Ascomycota; Saccharomycotina; Saccharomycetaceae	SGD	Goffeau et al., 1996; Cherry et al., 1997; Liti et al., 2009
Saccharomyces kluyveri	Ascomycota; Saccharomycotina; Saccharomycetaceae	SGD	The Genolevures Consortium et al., 2009
Saccharomyces mikatae	Ascomycota; Saccharomycotina; Saccharomycetaceae	SGD	Kellis et al., 2003
Saccharomyces paradoxus	Ascomycota; Saccharomycotina; Saccharomycetaceae	SGD	Kellis et al., 2003
Saitoella complicata	Ascomycota; Taphrinomycotina	JGI	JGI[b]
Schizophyllum commune	Basidiomycota; Agaricomycotina	JGI	Ohm et al., 2010
Schizosaccharomyces pombe	Ascomycota; Taphrinomycotina	Broad	Broad Institute[o]

Sclerotinia sclerotiorum	Ascomycota; Leotiomycetes	Broad	Broad Institute[p]
Spathaspora passalidarum	Ascomycota; Saccharomycotina; Debaryomycetaceae	JGI	Wohlbach et al., 2011
Sphaerobolus stellatus	Basidiomycota; Agaricomycotina	JGI	JGI[b]
Spizellomyces punctatus	Chytridiomycota	JGI	Broad Institute[c]
Thielavia terrestris	Ascomycota; Sordariomycetes	JGI	Berka et al., 2011
Tremella mesenterica	Basidiomycota; Agaricomycotina	JGI	Floudas et al., 2012
Trichoderma reesei	Ascomycota; Sordariomycetes	JGI	Martinez et al., 2008
Tuber melanosporum	Ascomycota; Pezizomycetes	JGI	Martin et al., 2010
Ustilago maydis	Basidiomycota; Ustilaginomycotina	Broad	Kämper et al., 2006
Vanderwaltozyma polyspora	Ascomycota; Saccharomycotina; Saccharomycetaceae	JGI	Scannell et al., 2007
Wickerhamomyces anomalus	Ascomycota; Saccharomycotina; Phaffomycetaceae	JGI	JGI[b]
Xanthoria parietina	Ascomycota; Lecanoromycetes	JGI	JGI[b]
Xylona heveae	Ascomycota; Xylonomycetes	JGI	JGI[b]
Yarrowia lipolytica	Ascomycota; Saccharomycotina; Dipodascaceae	NCBI	Dujon et al., 2004
Zygosaccharomyces rouxii	Ascomycota; Saccharomycotina; Saccharomycetaceae	JGI	The Genolevures Consortium et al., 2009
Holozoa			
Caenorhabditis elegans	Metazoa, Nematoda	WormBase	*C. elegans* Sequencing Consortium, 1998
Canis familiaris	Metazoa, Chordata, Craniata	NCBI	Lindblad-Toh et al., 2005
Capsaspora owczarzaki	Fungi/Metazoa insertae serdis, Filasterea	Broad	Broad Institute[c]
Danio rerio	Metazoa, Chordata, Craniata	NCBI	Howe et al., 2013
Drosophila melanogaster	Metazoa, Arthopoda	NCBI	Adams et al., 2000
Homo sapiens	Metazoa, Chordata, Craniata	NCBI	International Human Genome Sequencing

			Consortium, 2001; Venter et al., 2001
Monosiga brevicollis	Fungi/ Metazoa insertae sedis, Choanoflagellida	JGI	King et al., 2008; Broad Institute[c]
Mus musculus	Metazoa, Chordata, Craniata	NCBI	Mouse Genome Sequencing Consortium, 2002
Nematostella vectensis	Metazoa, Cnidaria	JGI	Putnam et al., 2007
Salpingoeca rosetta	Fungi/ Metazoa insertae sedis, Choanoflagellida	Broad	Broad Institute[c]
Sphaeroforma arctica	Fungi/ Metazoa insertae sedis, Ichthyosporea	Broad	Broad Institute[c]
Thecamonas trahens	Apusomonadidae	Broad	Broad Institute[c]
Amoebozoa			
Acanthamoeba castellanii	Acanthamoebae	BCM-HGSC	Clarke et al., 2013
Dictyostelium discoideum	Dictyostelids	dictyBase	Eichinger et al., 2005
Entamoeba histolytica	Archamoebae, Entamoebae	AmoebaDB	Loftus et al., 2005
Archaeplastida			
Arabidopsis thaliana	Viridiplantae, Streptophyta, Embryophytes	NCBI	*Arabidopsis* Genome Initiative, 2000
Chlamydomonas reinhardtii	Viridiplantae, Chlorophyta, Chlorophyceae, Chlamydomonadales	Phytozome	Merchant et al., 2007
Cyanidioschyzon merolae	Rhodoplastids, Cyanidiophytes	*C. merolae* genome project	Matsuzaki et al., 2004
Micromonas sp.	Viridiplantae, Chlorophyta, Mamiellales	JGI	Worden et al., 2009
Oryza sativa	Viridiplantae, Streptophyta, Embryophytes	NCBI	Goff et al., 2002

Ostreococcus tauri	Viridiplantae, Chlorophyta, Mamiellales	JGI	Palenik et al., 2007
Populus trichocarpa	Viridiplantae, Streptophyta, Embryophytes	Phytozome	Tuskan et al., 2006
Volvox carteri	Viridiplantae, Chlorophyta, Chlorophyceae, Chlamydomonadales	Phytozome	Prochnik et al., 2010
SAR			
Bigelowiella natans	Rhizaria, Cercoza, Filosa, Chlorarachniophytes	JGI	Curtis et al., 2012
Cryptosporidium parvum	Alveolates, Apicomplexa	CryptoDB	Abrahamsen et al., 2004
Ectocarpus siliculosus	Alveolates, Ciliates	ORCAE	Cock et al., 2010
Phytophthora ramourum	Stramenopiles	JGI	Tyler et al., 2006
Plasmodium falciparum	Alveolates, Apicomplexa	PlasmoDB	Gardner et al., 2002
Tetrahymena thermophila	Alveolates, Ciliates	TGD	Eisen et al., 2006
Thalassiosira pseudonana	Stramenopiles	JGI	Armbrust et al., 2004
Theileria parva	Alveolates, Apicomplexa	NCBI	Gardner et al., 2005
Toxoplasma gondii	Alveolates, Apicomplexa	ToxoDB	Kissinger et al., 2003
CCTH			
Emiliania huxleyi	Haptophytes	JGI	Read et al., 2013
Guillardia theta	Cryptophytes	JGI	Curtis et al., 2012
Excavata			
Bodo saltans	Discoba, Euglenozoa, Kinetoplastida, Bodonidae	-	In preparation[q]
Crithidia fasciculata	Discoba, Euglenozoa, Kinetoplastida, Trypanosomatidae	TriTrypDB	Aslett et al., 2009
Giardia sp.	Fornicata, Diplomonadida	GiardiaDB	Aurora et al., 2009
Leishmania major	Discoba, Euglenozoa, Kinetoplastida, Trypanosomatidae	TriTrypDB	Ivens et al., 2005
Naegleria gruberi	Discoba, Heterolobosea	JGI	Fritz-Laylin et al., 2010

Trichomonas vaginalis	Metamonads, Parabasalia	TrichDB	Carlton et al., 2007
Trypanosoma brucei	Discoba, Euglenozoa, Kinetoplastida, Trypanosomatidae	TriTrypDB	Berriman et al., 2005
Trypanosoma cruzi	Discoba, Euglenozoa, Kinetoplastida, Trypanosomatidae	TriTrypDB	El-Sayed et al., 2005

[a]SGD (yeastgenome.org); NCBI (nchi.nlm.nih.gov); JGI (genome.jgi.doe.gov/programs/fungi/index.jsf); Broad Institute (broadinstitute.org); WormBase (wormbase.org); BCM-HGSC (hgsc.bcm.edu/microbiome/acanthamoeba-castellani-neff); dictyBase (dictybase.org); AmoebaDB (amoebadb.org); Phytozome (phytozome.net); *C. merolae* genome project (merolae.biol.s.u-tokyo.ac.jp); CryptoDB (cryptodb.org); ORCAE (bioinformatics.psb.ugent.be/orcae/overview/Ectsi); PlasmoDB (plasmodb.org); TGD (ciliate.org); ToxoDB (toxodb.org); GiardiaDB (giardiadb.org); TriTrypDB (tritrypdb.org)

[b]Unpublished sequence data produced by the US Department of Energy Joint Genome Institute (http://www.jgi.doe.gov/) in collaboration with the user community.
[c]The Origins of Multicellularity Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[d]Aspergillus Comparative Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[e]The Batrachochytrium dendrobatidis Genome Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[f]The Botrytis cinerea Genome Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[g]Coccidioides Group Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[h]Coprinopsis cinerea Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[i]Cryptococcus neoformans var. grubii H99 Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[j]The Fusarium Comparative Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[k]Candida Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[l]Magnaporthe Comparative Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[m]Stagonospora nodorum Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[n]Puccinia Group Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[o]Schizosaccharomyces Group Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[p]Sclerotinia sclerotiorum Sequencing Project, Broad Institute of Harvard and MIT (broadinstitute.org)
[q]Unpublished data (*Bodo saltans* Genome Project, Andrew Jackson, University of Liverpool)

2.11.2 Comparative genomic survey of peroxins in Blastocystis hominis

H. sapiens, *S. cerevisiae*, and/or *N. crassa* protein sequences (Appendix 7) were used as initial queries for BLASTp and tBLASTn (Altschul et al., 1997) searches against locally hosted *Blastocystis*

protein sequences and scaffolds, respectively, to identify putative peroxins in *Blastocystis*. Candidate homologues with E-values less than or equal to 0.05 were subjected to reciprocal BLASTp (protein sequences) or BLASTx (nucleotide sequences) searches against the locally hosted query genome, as well as the locally hosted NR genome. The retrieval of the original query, or an equivalently named orthologue, as the top reciprocal BLASTp hit with an E-value less than or equal to 0.05 was the criterion for identification of a putative homologue.

H. sapiens, *S. cerevisiae*, and/or *N. crassa* protein sequences were used as queries for pHMMer (Eddy, 2009) searches against locally hosted query genomes to identify putative peroxins in *Phytophthora ramorum* and *Thalassiosira pseudonana*, and to confirm previously predicted peroxins in *Arabidopsis thaliana* (Nito et al., 2007) and *Phaeodactylum tricornutum* (Gonzalez et al., 2011). Candidate homologues with E-values less than or equal to 0.05 were subjected to reciprocal pHMMer searches against the locally hosted genome and NR database. Thus, the retrieval of the original query as the top reciprocal pHMMer hit with an E-value less than or equal to 0.05 was the criterion for identification of a putative homologue.

Newly identified or verified *A. thaliana*, *P. ramorum*, *T. pseudonana* and *P. tricornutum* peroxin sequences (Appendices 8-11) were subsequently used as queries for BLASTp and and tBLASTn searches against locally hosted *Blastocystis* protein sequences and scaffolds, respectively. Candidate homologues were subjected to reciprocal BLASTp or BLASTx searches according to the criteria described above.

2.11.3 Comparative genomic survey of peroxins and glycosomal proteins in Bodo saltans

To identify putative peroxin proteins in *N. gruberi*, *H. sapiens*, *S. cerevisiae*, *N. crassa* and/or *A. thaliana* protein sequence (Appendices 7-8) were used as queries for pHMMer searches against locally hosted *N. gruberi* protein sequences. Candidate homologues with E-values less than or equal to 0.05 were subjected to reciprocal pHMMer searches against the locally hosted query genome, as well as the locally hosted non-redundant (NR) genome. The retrieval of the original query (or an

equivalently named orthologue) as the top reciprocal pHMMer hit with an E-value less than or equal to 0.05 was the criteria for identification of a putative homologue.

Putative peroxin proteins in *Bodo saltans* were identified by Dr. Fred Mast, Seattle Biomedical Research Institute, Seattle, WA. To identify putative peroxin proteins in *Bodo saltans*, *H. sapiens*, *A. thaliana* and/or *T. brucei* protein sequences, were used as queries for pHMMer searches against locally hosted *B. saltans* protein sequences. Candidate homologues with E-values less than or equal to 0.05 were subjected to reciprocal pHMMer searches against the locally hosted query genome, as well as the locally hosted non-redundant (NR) genome. The retrieval of the original query (or an equivalently named orthologue) as the top reciprocal pHMMer hit with an E-value less than or equal to 0.05 was the criteria for identification of a putative homologue.

To identity putative proteins comprising the *B. saltans* peroxisomal proteome, *T. brucei* queries for 201 high-confidence glycosomal proteins (Supplementary tables, Guther et al., 2014) were used as queries for pHMMer searches against locally hosted *B. saltans* protein sequences. Candidate homologues with E-values less than or equal to 0.05 were subjected to reciprocal pHMMer searches against the locally hosted query genome. The retrieval of the original query (or an equivalently named orthologue) as the top reciprocal pHMMer hit with an E-value less than or equal to 0.05 was the criteria for identification of a putative homologue.

To confirm previously predicted glycolysis enzymes in *N. gruberi*, protein sequences were subjected to reciprocal pHMMer searches against locally hosted *T. brucei* protein sequences. The retrieval of an equivalently named orthologue as the top reciprocal pHMMer hit with an E-value less than or equal to 0.05 was the criteria for identification of a positive homologue.

PTS-containing proteins in *B. saltans, N. gruberi*, *N. fowleri*, and other trypanosomatid genomes were identified using the program PTS Finder (Section 4.1). Trypanosomatid genomes were obtained from TriTrypDB (tritrypDB.org). PTS-containing proteins in *B. saltans* were subjected to reciprocal pHMMer searches against the *T. brucei* genome, with the top hit recorded.

2.11.4 Comparative genomic survey of peroxins in Naegleria fowleri

H. sapiens, *S. cerevisiae*, and/or *N. crassa* protein sequences (Appendix 7), as well as *N. gruberi* protein sequences identified in Section 2.11.3 (Appendix 13A), were used as initial queries for pHMMer searches against locally hosted *N. fowleri* protein sequences to identify putative peroxins in *N. fowleri*. Candidate homologues with E-values less than or equal to 0.05 were subjected to reciprocal BLASTp searches against the locally hosted query genome, as well as the locally hosted NR genome. The retrieval of the original query, or an equivalently named orthologue, as the top reciprocal BLASTp hit with an E-value less than or equal to 0.05 was the criterion for identification of a putative homologue. *N. gruberi* protein sequences encoding glycolysis enzymes (Section 2.11.3) were used as queries for pHMMer searches against locally hosted *N. fowleri* protein sequences, followed by reciprocal pHMMer searches as described above.

2.12 Alignment

Protein sequences were aligned using MUSCLE 3.6 (Edgar, 2006) and then analyzed using ZORRO (Wu et al., 2012). The alignments were masked and trimmed in Mesquite 2.75 (www.mesquiteproject.org) so as to retain only unambiguously homologous positions for analysis. Alignments are available on request. 19 taxa and 189 positions were used for the analysis whose results are presented in Figure 3-2. 84 taxa and 214 positions were used for the analysis whose results are presented in Figure 3-4. 17 taxa and 304 positions were used for the analysis whose results are presented in Figure 3-5.

2.12.1 Use of ZORRO for masking multiple sequence alignments

ZORRO is a probabilistic masking program that assigns confidence scores to each alignment position (Wu et al., 2012). ZORRO uses a version of the pair hidden Markov model (Durbin et al., 1998) to calculate the posterior probability of two residues being aligned in all possible alignnments. Whereas the output of a HMM is a single sequence, the output of a pair HMM is a pairwise alignment (Durbin et al., 1998). ZORRO uses a weighing scheme to sum all of the posterior

probabilities in an alignment column, and assign a confidence score between from 0 to 1. ZORRO outperformed other masking programs tested in terms of both sensitivity and specificity, produced alignments that were more consistent (especially as sequences became more divergent and difficult to align) and improved the resulting phylogenetic reconstruction (particularly for longer and/or faster evolving proteins) (Wu et al., 2012). Prior to phylogenetic anlysis, alignments are typically masked and trimmed manually. However, this approach is somewhat subjective and is not practical for large-scale projects where phylogenetic analyses of thousands of genes are done automatically. The importance of multiple sequence alignments is underscored by the impact on the resulting phylogenetic tree. Alignment uncertainty and assessing the quality of each individual position in an aligment (whether manual or automated) are important questions. In Section 3.2, I describe one approach to using ZORRO to mask multiple sequence alignments prior to phylogenetic anlaysis.

2.13 Phylogenetic analysis

ProtTest version 1.3 (Abascal et al., 2005) was used to determine the optimal model of sequence evolution. The optimal models of sequence evolution for the analyses were: LG + I + G + F (Figure 3-2), LG + I + G + F (Figure 3-4) and LG + I + G + F (Figure 3-5). MrBayes version 3.2 (Ronquist and Huelsenbeck, 2003) was used for Bayesian analysis to produce posterior probability values. Analyses were run for 1,000,000 Markov chain Monte Carlo generations. Two independent runs were performed, with convergence of the results confirmed by ensuring a Splits Frequency of < 0.1. The burn-in values were obtained by removing all trees prior to a graphically determined plateau. Additionally, PhyML version 2.44 (Guindon and Gascuel, 2003) and RAxML version 8.0.0 (Stamatakis, 2006) were used for maximum likelihood analyses, with bootstrap values based on 100 pseudoreplicates of each dataset. RAxML trees were run using Cyberinfrastructure for Phylogenetic Research (CIPRES) (www.phylo.org/index.php/portal).

Chapter 3: An ancestral role in *de novo* peroxisome assembly is retained by the divisional peroxin Pex11 in *Yarrowia lipolytica*

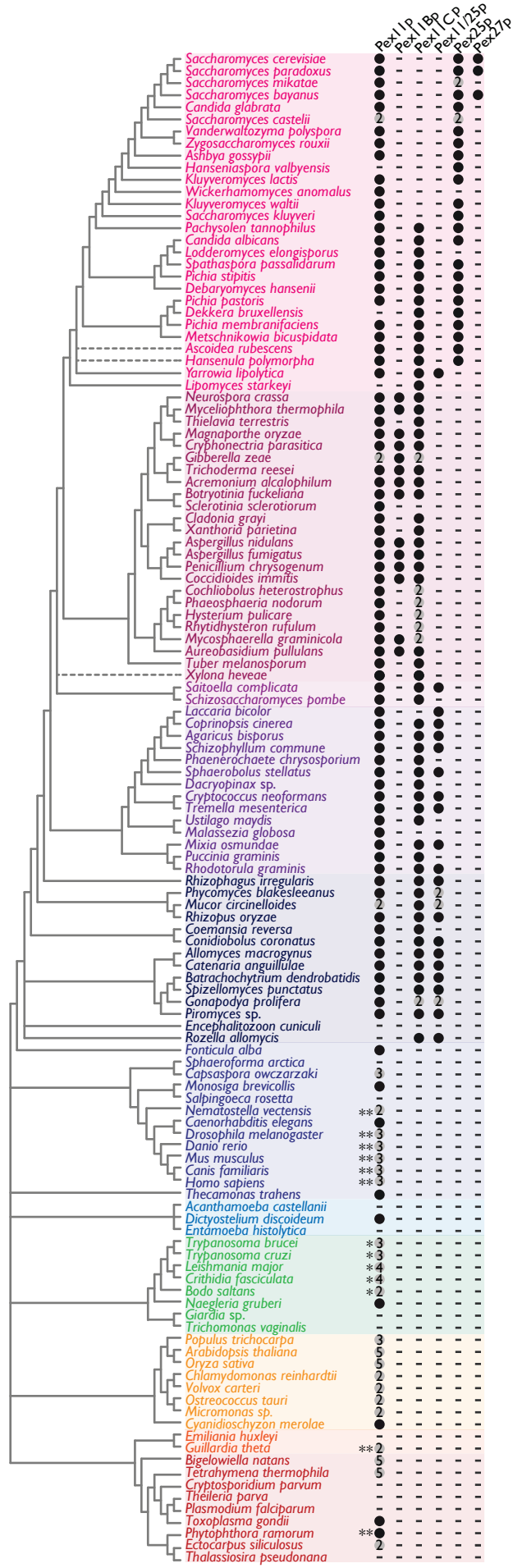
A version of this chapter has been submitted for publication as:
Klute, M.J., Chang, J., Tower, R.J., Mast, F.D., Dacks, J.B., and R.A. Rachubinski. 2014. An ancestral role in *de novo* peroxisome assembly is retained by the divisional peroxin Pex11 in the yeast *Yarrowia lipolytica*. *J. Cell Sci.*

3.1 A comparative genomic survey of the Pex11 protein family

Two previous studies of *PEX* gene evolution found that most *PEX* genes are of eukaryotic origin and that there are patterns of peroxin conservation and loss specific to lineages of the eukaryotic tree of life (Schlüter et al., 2006; Gabaldón et al., 2006). Since these studies, the number of eukaryotic genome sequencing projects has increased dramatically, and new *PEX* genes have been identified (Managadze et al., 2010; Tower et al., 2011), allowing for a more complete picture of peroxisomal protein evolution. We have therefore revisited the evolution and distribution of the Pex11 protein family by completing a comparative genomics survey using 125 genomes that span the six eukaryotic supergroups (Walker et al., 2011).

Our analysis showed that Pex11p itself was likely present at the time of the last eukaryotic common ancestor (LECA) (Figure 3-1) (Appendix 1). As previously noted, numerous genomes contain multiple paralogues of Pex11p. The human and mouse genomes both encode PEX11α, PEX11β and PEX11γ, and five copies of Pex11p are present in the plants *Arabidopsis thaliana* (Lingard and Trelease, 2006) and *Oryza sativa* (Nayidu et al., 2008). Three copies of Pex11p are present in *Drosophila melanogaster* in addition to single paralogues that were previously identified (Mast et al., 2011; Faust et al., 2012). Additional genomes encode multiple copies of Pex11p; there are two or three copies in Archaeplastida genomes (*Chlamydomonas reinhardtii, Micromonas* sp.*, Ostreococcus tauri, Populus trichocarpa, and Volvox carteri*), and five copies in both the rhizarian alga *Bigelowiella natans* and the ciliate *Tetrahymena thermophila*. Peroxisomes have not been reported in many parasites (de Souza et al., 2004; Gabaldón et al., 2006; Gabaldon, 2010), and it was therefore not unexpected that Pex11p was not identified in the parasite genomes analyzed in this study: *Encephalitozoon cuniculi, Entamoeba histolytica, Giardia* sp.*, Trichomonas vaginalis, Theileria parva* and *Plasmodium falciparum*. Surprisingly, Pex11p was identified in the parasite *Toxoplasma gondii*, although peroxisomes have not been observed in this organism (Ding et al., 2000). However, using the putative Pex11p from *T. gondii* as a query for a pHMMer search in the NR database returned Pex11p homologues from Fungi, *H. sapiens* and other Opisthokonts as top hits (rather than

Figure 3-1. Comparative genomic survey of the Pex11 protein family in Fungi and other eukaryotes. Each column represents a Pex11 family protein that has been characterized in *H. sapiens* and/or *S. cerevisiae*, or that was identified by a previous bioinformatic analysis (Kiel et al., 2006). Individual genomes from the comparative genomics survey are color-coded according to eukaryotic supergroup and are grouped according to taxanomic classification. A black circle indicates the presence of a protein in the indicated taxon based on positive reciprocal pHMMer searches, grey circles with numbers indicate multiple paralogues of a protein and dashes indicate that proteins were not identified in the indicated taxon. The evolutionary relationships between the genomes included in this study are given at left, but the rooting of the tree is arbitrary. Dotted lines indicate genomes with unknown evolutionary placement. Asterisks indicate GIM5 homologues. Double asterisks indicate Pex11p homologues that grouped with Pex11Cp in phylogenetic analysis. The evolutionary relationships between the taxa presented were determined using: http://www.ncbi.nlm.nih.gov/taxonomy and (Aime et al., 2006; Hibbett, 2006; James et al., 2006; Suh et al., 2006; Wapinski et al., 2007; Boehm et al., 2010; Schoch et al., 2010; Medina et al., 2011; Walker et al., 2011; Ebersberger et al., 2012; Mast et al., 2012).


Pex11p
Pex11Bp
Pex11Cp
Pex11i25p
Pex25p
Pex27p
Saccharomyces cerevisiae
Saccharomyces paradoxus
Saccharomyces mikatae
Saccharomyces bayanus
Candida glabrata
Saccharomyces castelii
Vanderwaltozyma polyspora
Zygosaccharomyces rouxii
Ashbya gossypii
Hanseniaspora valbyensis
Kluyveromyces lactis
Wickerhamomyces anomalus
Kluyveromyces waltii
Saccharomyces kluyveri
Pachysolen tannophilus
Candida albicans
Lodderomyces elongisporus
Spathaspora passalidarum
Pichia stipitis
Debaryomyces hansenii
Pichia pastoris
Dekkera bruxellensis
Pichia membranifaciens
Metschnikowia bicuspidata
Ascoidea rubescens
Hansenula polymorpha
Yarrowia lipolytica
Lipomyces starkeyi
Neurospora crassa
Myceliophthora thermophila
Thielavia terrestris
Magnaporthe oryzae
Cryphonectria parasitica
Gibberella zeae
Trichoderma reesei
Acremonium alcalophilum
Botryotinia fuckeliana
Sclerotinia sclerotiorum
Cladonia grayi
Xanthoria parietina
Aspergillus nidulans
Aspergillus fumigatus
Penicillium chrysogenum
Coccidioides immitis
Cochliobolus heterostrophus
Phaeosphaeria nodorum
Hysterium pulicare
Rhytidhysteron rufulum
Mycosphaerella graminicola
Aureobasidium pullulans
Tuber melanosporum
Xylona heveae
Saitoella complicata
Schizosaccharomyces pombe
Laccaria bicolor
Coprinopsis cinerea
Agaricus bisporus
Schizophyllum commune
Phanerochaete chrysosporium
Sphaerobolus stellatus
Dacryopinax sp.
Cryptococcus neoformans
Tremella mesenterica
Ustilago maydis
Malassezia globosa
Mixia osmundae
Puccinia graminis
Rhodotorula graminis
Rhizophagus irregularis
Phycomyces blakesleeanus
Mucor circinelloides
Rhizopus oryzae
Coemansia reversa
Conidiobolus coronatus
Allomyces macrogynus
Catenaria anguillulae
Batrachochytrium dendrobatidis
Spizellomyces punctatus
Gonapodya prolifera
Piromyces sp.
Encephalitozoon cuniculi
Rozella allomycis
Fonticula alba
Sphaeroforma arctica
Capsaspora owczarzaki
Monosiga brevicollis
Salpingoeca rosetta
Nematostella vectensis
Caenorhabditis elegans
Drosophila melanogaster
Danio rerio
Mus musculus
Canis familiaris
Homo sapiens
Thecamonas trahens
Acanthamoeba castellanii
Dictyostelium discoideum
Entamoeba histolytica
Trypanosoma brucei
Trypanosoma cruzi
Leishmania major
Crithidia fasciculata
Bodo saltans
Naegleria gruberi
Giardia sp.
Trichomonas vaginalis
Populus trichocarpa
Arabidopsis thaliana
Oryza sativa
Chlamydomonas reinhardtii
Volvox carteri
Ostreococcus tauri
Micromonas sp.
Cyanidioschyzon merolae
Emiliania huxleyi
Guillardia theta
Bigelowiella natans
Tetrahymena thermophila
Cryptosporidium parvum
Theileria parva
Plasmodium falciparum
Toxoplasma gondii
Phytophthora ramorum
Ectocarpus siliculosus
Thalassiosira pseudonana

homologues from genomes that are more closely related to *T. gondii*). This result indicates possible contamination or horizontal gene transfer. Failure to identify Pex11p homologues in the remaining eukaryotic genomes could be because an organism lacks peroxisomes but has never been experimentally characterized, the genomes encode highly divergent sequences that were not detected by our search algorithms, or poor genome assembly and/or coverage.

Unlike for most parasites, a divergent peroxisome called the glycosome has been described ultrastructurally for the Kinetoplastida (Opperdoes et al., 1977). It has been previously noted that *Trypanosoma brucei* encodes multiple Pex11 family proteins, *i.e.* Pex11p itself and the Pex11p-related proteins GIM5A and GIM5B (Maier et al., 2001). GIM5 and Pex11p are most abundant and second most abundant glycosomal membrane proteins in *T. brucei*, respectively, and are both essential genes for typanosome survival (Lorenz et al., 1998; Maier et al., 2001; Voncken et al., 2003). The remaining trypanosomatid genomes queried in this survey also encode multiple Pex11p homologues, with Pex11p itself or GIM5 as the top reciprocal pHMMer hits in the *T. brucei* genome (Figure 3-1). An evolutionary relationship was proposed to exist between GIM5 and trypanosome Pex11p on the basis of sequence similarity and common function (Voncken et al., 2003). *T. brucei* GIM5A and GIM5B share 13% and 14% sequence identity with *T. brucei* Pex11p (Voncken et al., 2003). Knock down of *T. brucei PEX11* expression resulted in decreased numbers of larger glycosomes, while overexpression of *PEX11* resulted in elongated glycosomal structures (Lorenz et al., 1998). Depletion of GIM5 resulted in cells with fewer, enlarged glycosomes (Maier et al., 2001; Voncken et al., 2003). Although Pex11p and GIM5 appear to share a common function in regulating glycosome size and number, there was no strong evidence for kinetoplastid Pex11p and GIM5 proteins being homologous. Reciprocal pHMMer searches into trypansome genomes retrieved only the query sequence and no additional protein. While still potentially being true Pex11p homologues, these proteins are highly divergent in sequence and frequently failed to retrieve Pex11p homologues in non-kinetoplastid genomes (data not shown). A phylogenetic analysis of the Pex11 family proteins in Excavata was completed to definitively classify these proteins, and resolved

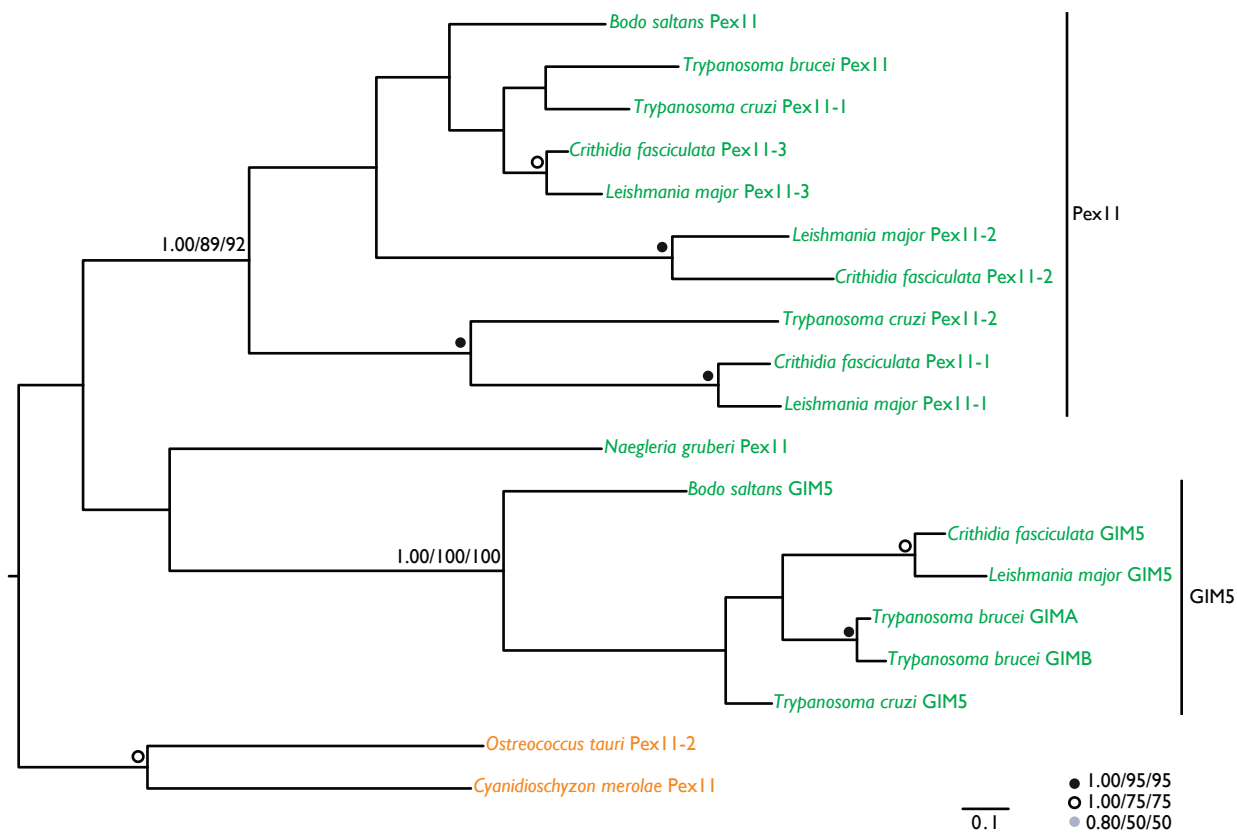
Pex11p and GIM5 into distinct clades, predating the split of kinetoplastids with their relative *Bodo saltans* (Figure 3-2). Genomes selected as outgroups throughout this chapter were those that resulted in the best tree resolution.

With few exceptions, the fungal genomes in this study encoded a single Pex11p homologue. However, paralogues of the remaining Pex11 family proteins were frequently observed in Fungi (Figure 3-1) (Appendices 2-6). Pex27p, initially characterized in *S. cerevisiae* (Tam et al., 2003), was identified in only two other *Saccharomyces* species: *S. paradoxes* and *S. bayanus* (Figure 3-1). Pex25p homologues were restricted to the Saccharomycotina, a lineage of ascomycete fungi (James et al., 2006). Pex11Bp, Pex11Cp and Pex11/25p are present in a much larger number of fungal genomes than just the filamentous Fungi, as initially reported (Kiel et al., 2006). Pex11Bp homologues were identified in members of the Pezizomycotina, which is a collective of fungal lineages that are more basal to the Saccharomycotina (James et al., 2006). Pex11Cp was identified across all lineages of Fungi studied. Pex11/25p homologues, with the exception of Pex11/25p homologues identified in *Y. lipolytica* and *Saitoella complicata* were restricted to more basal fungal lineages outside the Ascomycota. These findings indicate that all Pex11 family proteins, with the exception of Pex11p itself, appear to be present only in the Fungi, and their distribution may have been sculpted by fungal evolution.

3.2 Use of ZORRO for masking multiple sequence alignments

We developed a procedure for selecting a ZORRO score threshold corresponding to positions in multiple sequence alignments that are kept in the resulting phylogenetic analysis. A percentage threshold cutoff rather than an absolute value was used because the maximum ZORRO score and distribution of ZORRO scores from 0-10 for each alignment vary. After running ZORRO on an aligned protein sequence file, the relative frequency of HMM scores was expressed as a function of HMM scores. Such a graph demonstrates the alignment quality, and there should be one or more discernable plateaus. HMM scores below a given plateau correspond to poorly aligned columns that

Figure 3-2. Phylogenetic analysis of Pex11p and GIM5 in Excavata. Node values indicate statistical support by MrBayes/PhyML/RAxML (posterior probability/bootstrap value/bootstrap value), with statistical values for highly supported nodes replaced by symbols as indicated. Best Bayesian topology is shown rooted on the *O. tauri* and *C. merolae* Pex11p sequences as outgroups. Species names are color-coded according to Figure 3-1. Excavate Pex11p and GIM5 clades are supported, but the exact position of *N. gruberi* Pex11p cannot be resolved.


Bodo saltans Pex11
Trypanosoma brucei Pex11
Trypanosoma cruzi Pex11-1
Crithidia fasciculata Pex11-3
Leishmania major Pex11-3
Leishmania major Pex11-2
Crithidia fasciculata Pex11-2
Trypanosoma cruzi Pex11-2
Crithidia fasciculata Pex11-1
Leishmania major Pex11-1
Pex11
1.00/89/92
Naegleria gruberi Pex11
Bodo saltans GIM5
Crithidia fasciculata GIM5
Leishmania major GIM5
Trypanosoma brucei GIMA
Trypanosoma brucei GIMB
Trypanosoma cruzi GIM5
GIM5
1.00/100/100
Ostreococcus tauri Pex11-2
Cyanidioschyzon merolae Pex11
0.1
1.00/95/95
1.00/75/75
0.80/50/50

are not informative. Likewise, HMM scores above a given plateau indicate that there is no gain in informative positions with increasing the score threshold further.

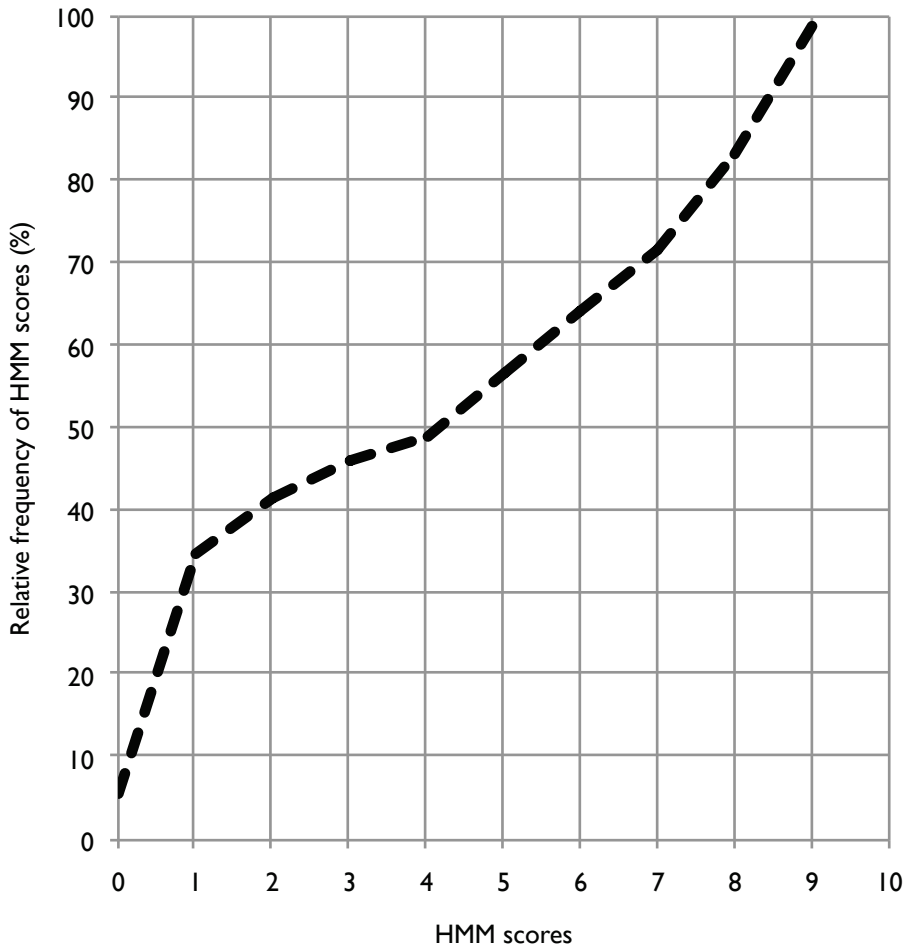
This procedure was applied to a sample alignment of Pex11p family proteins. The relative HMM scores were displayed graphically and possible plateaus of approximately 35%, 50% and 70% were observed (there is not always a single, clear plateau) (Figure 3-3A). The alignments were masked and trimmed using ZORRO score percentage thresholds of 0% (no mask) to 90%. ProtTest was used to determine the model of sequence evolution chosen for each masked and trimmed alignment, and the model of sequence evolution chosen did not typically differ with increasing masking thresholds, except when very few positions were included, typically at the high range of threshold (data not shown). Phylogenetic trees were built for each alignment as described in Section 2-13. For this experiment, consistent topological observations and supported clades were observed up to and including the 50% threshold (Figure 3-3B-K). Resolved PEX11α, PEX11β and PEX11γ clades are observed, with consistent placement of the *Drosophila* Pex11p sequences at the base of the PEX11α/PEX11β and PEX11γ clades. After the 60% threshold, support is lost for the *D. rerio* PEX11α sequence belonging to the PEX11α clade. At the 80% and 90% thresholds, support is lost for the previously observed placement of the *Drosophila* Pex11 proteins in the tree. Thus, 50% could be an appropriate threshold for automatic alignment masking for this experiment, which was one of the plateaus described above (Figure 3-3A). It is important to note that the root of the tree, although always consistent, is never supported until the 90% threshold, but this tree likely contains too few positions to be informative. This method appears to be a useful starting point for implanting ZORRO in phylogenetic tree building. Future work on applying ZORRO to alignments of interest would be very useful for using this tool further.

3.3 Evolution of the Pex11 protein family

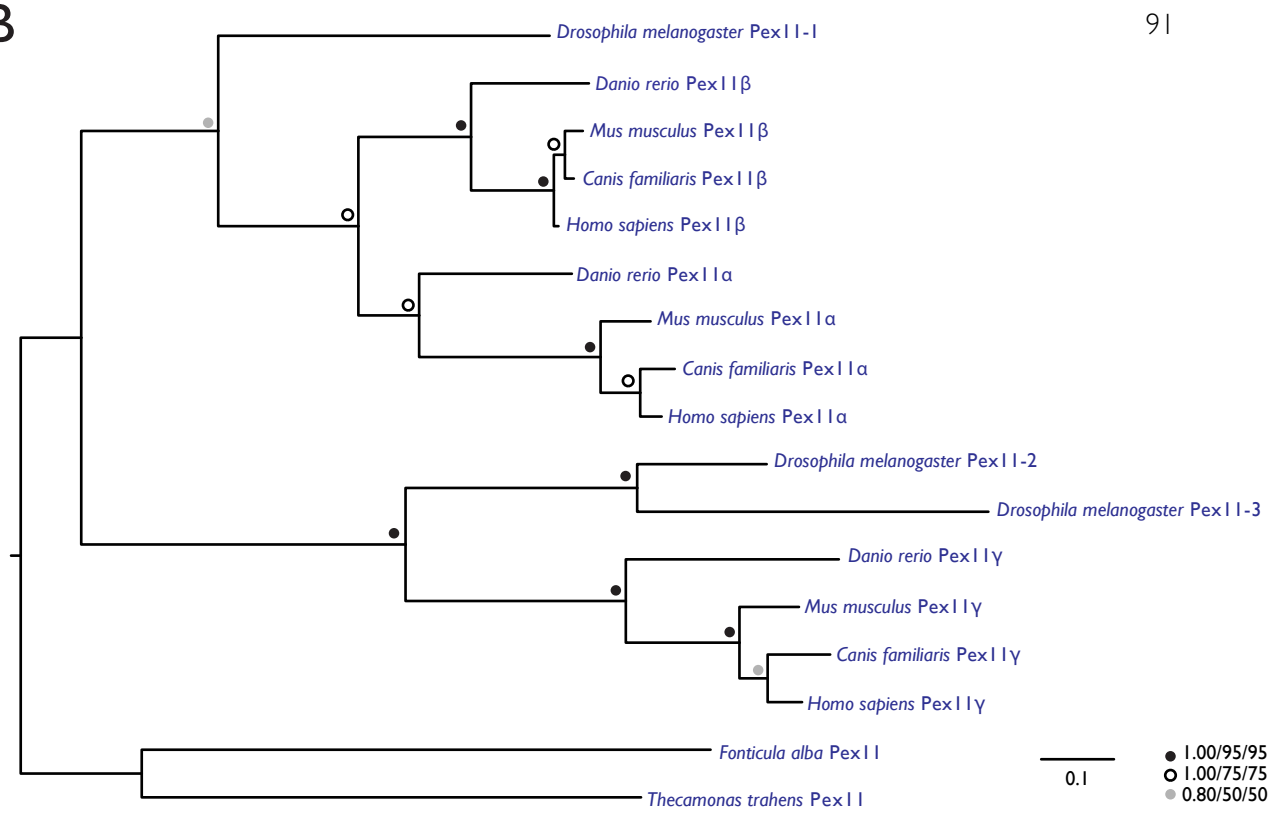
Unlike most other peroxins, the Pex11 protein family has undergone multiple expansions in diverse eukaryotic lineages (reviewed in Smith and Aitchison, 2013). Moderate sequence similarity between

Figure 3-3. Example of the application of ZORRO to masking multiple sequence alignments and phylogenetic analyses. (A) Exemplary graph of ZORRO output for an alignment of Pex11 proteins. Relative frequency of HMM scores (%) is expressed as a function of raw HMM scores. (B-K) Exemplary phylogenetic analyses using alignments from (A), following masking and trimming alignments using ZORRO thresholds from 0 to 90%, with increasing increments of 10%. Node values indicate statistical support by MrBayes/PhyML/RAxML (posterior probability/bootstrap value/bootstrap value), with statistical values for highly supported nodes replaced by symbols as indicated. Best Bayesian topology is shown rooted on the *F. alba* and *T. trahens* Pex11p sequences as outgroups. Species names are color-coded according to Figure 3-1.

A

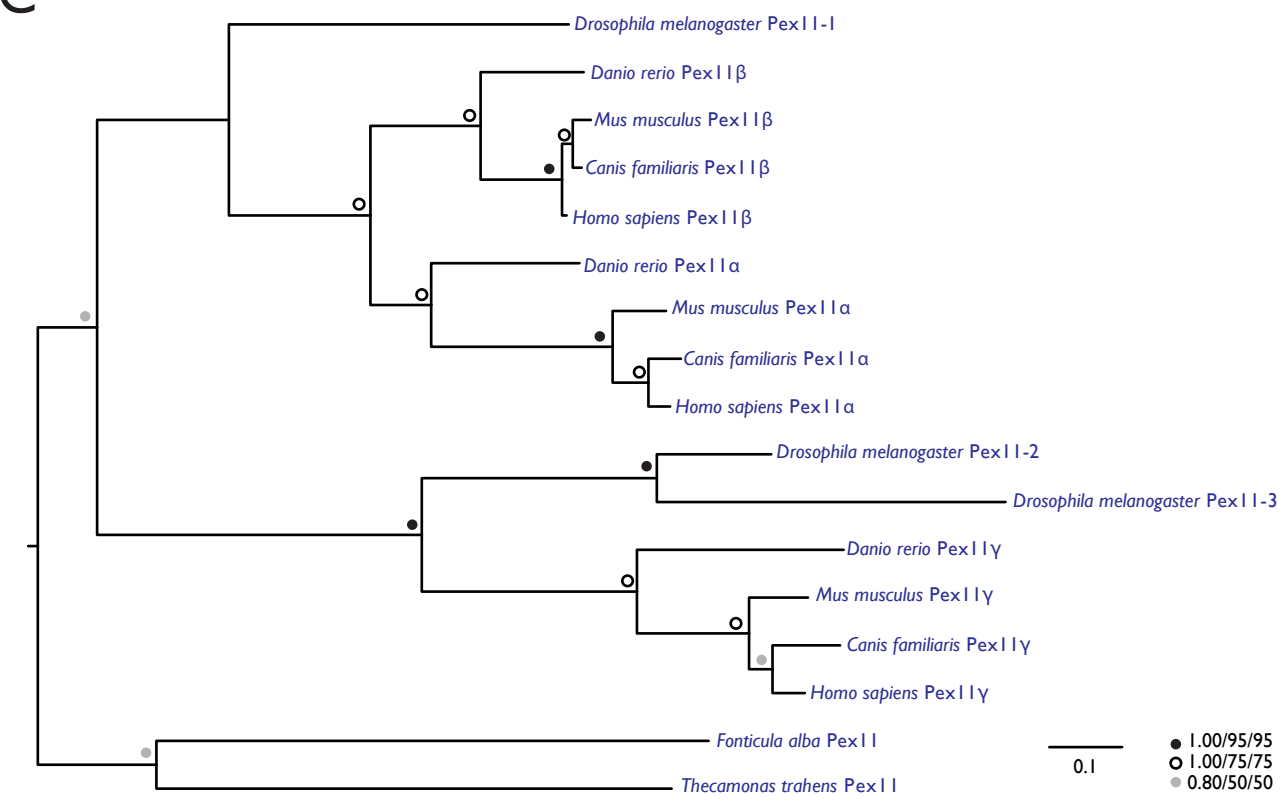

100
90
80
70
60
50
40
30
20
10
0
Relative frequency of HMM scores (%)
0
1
2
3
4
5
6
7
8
9
10
HMM scores

B



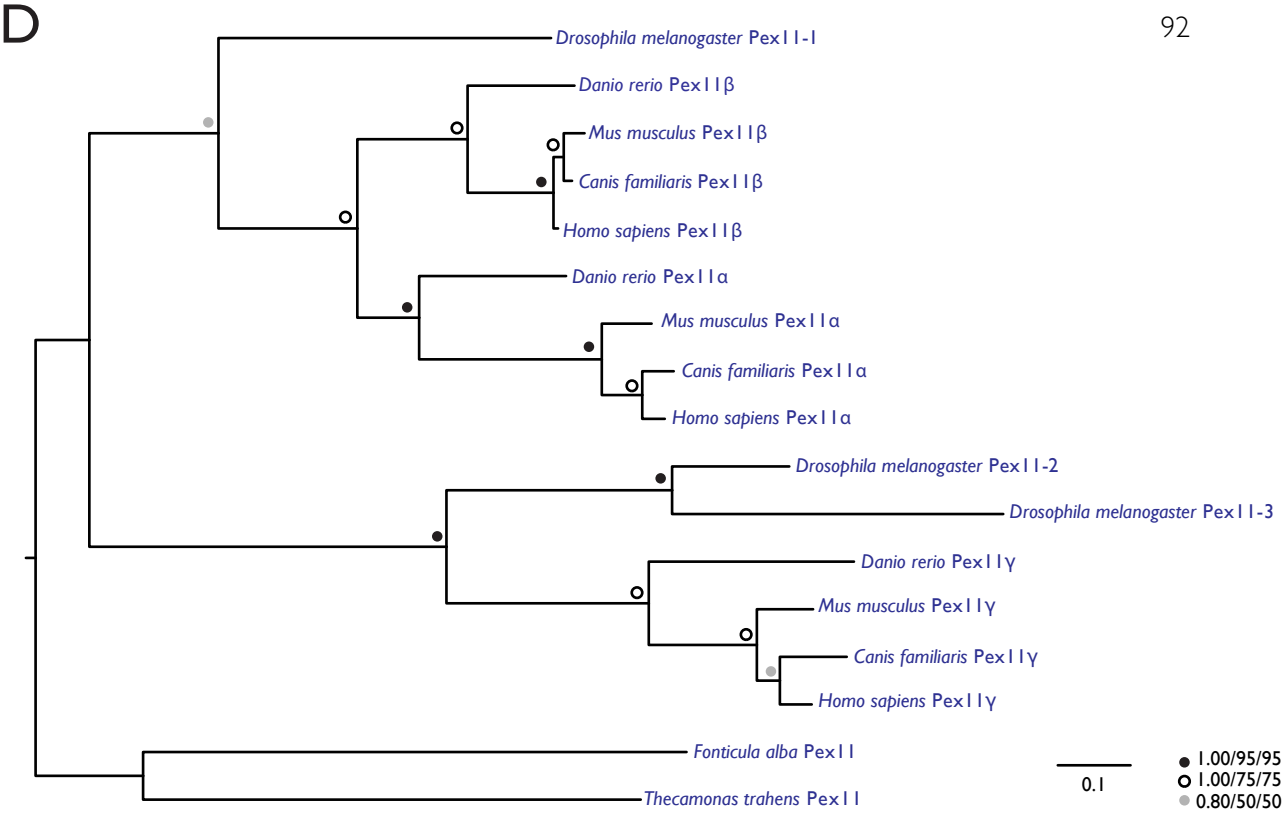
ZORRO - 0% Mask

C



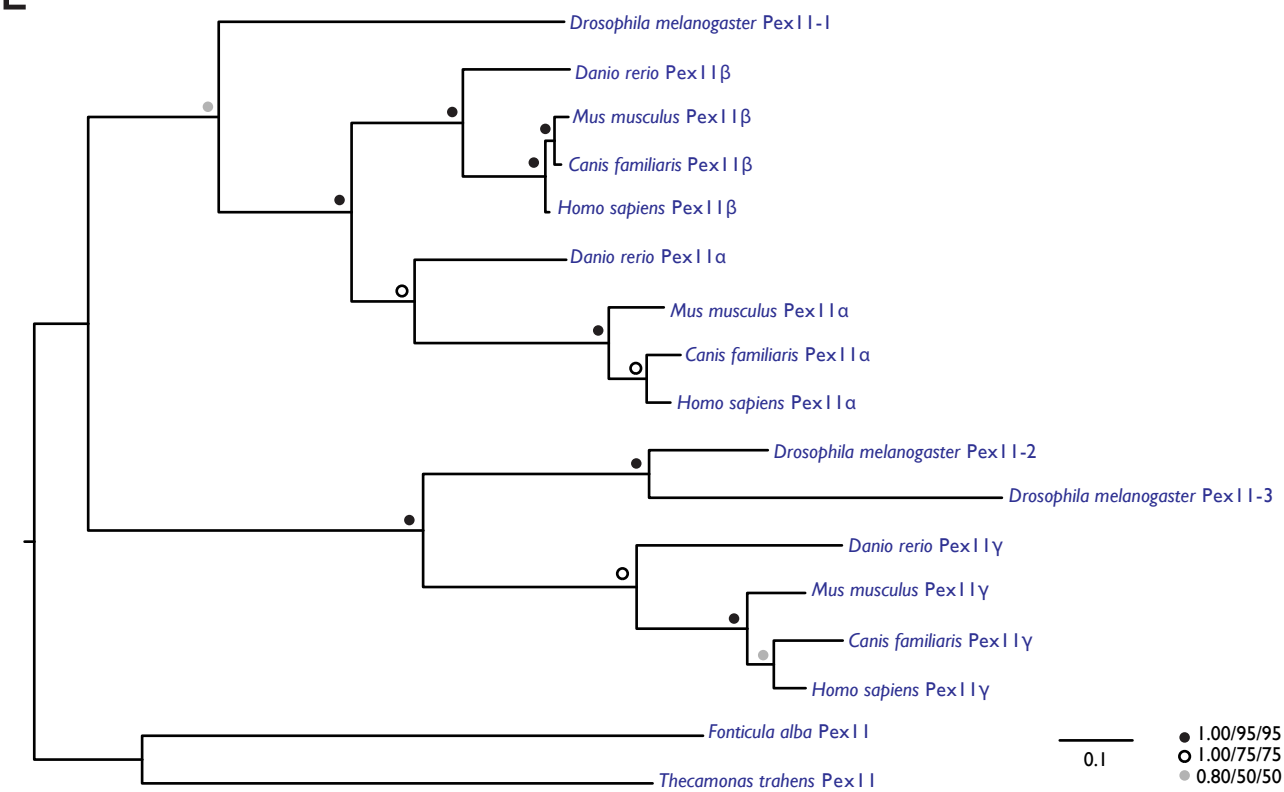
ZORRO - 10% Mask

D



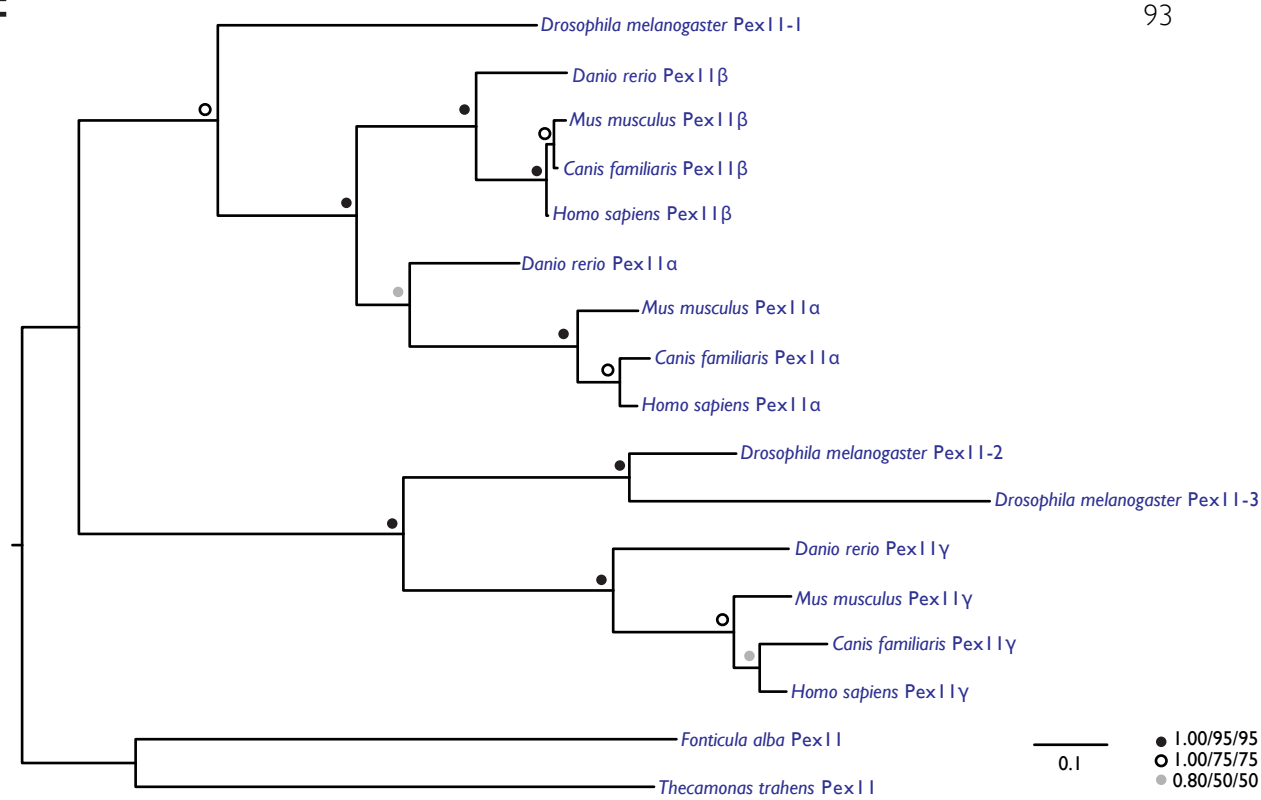
ZORRO - 20% Mask

E



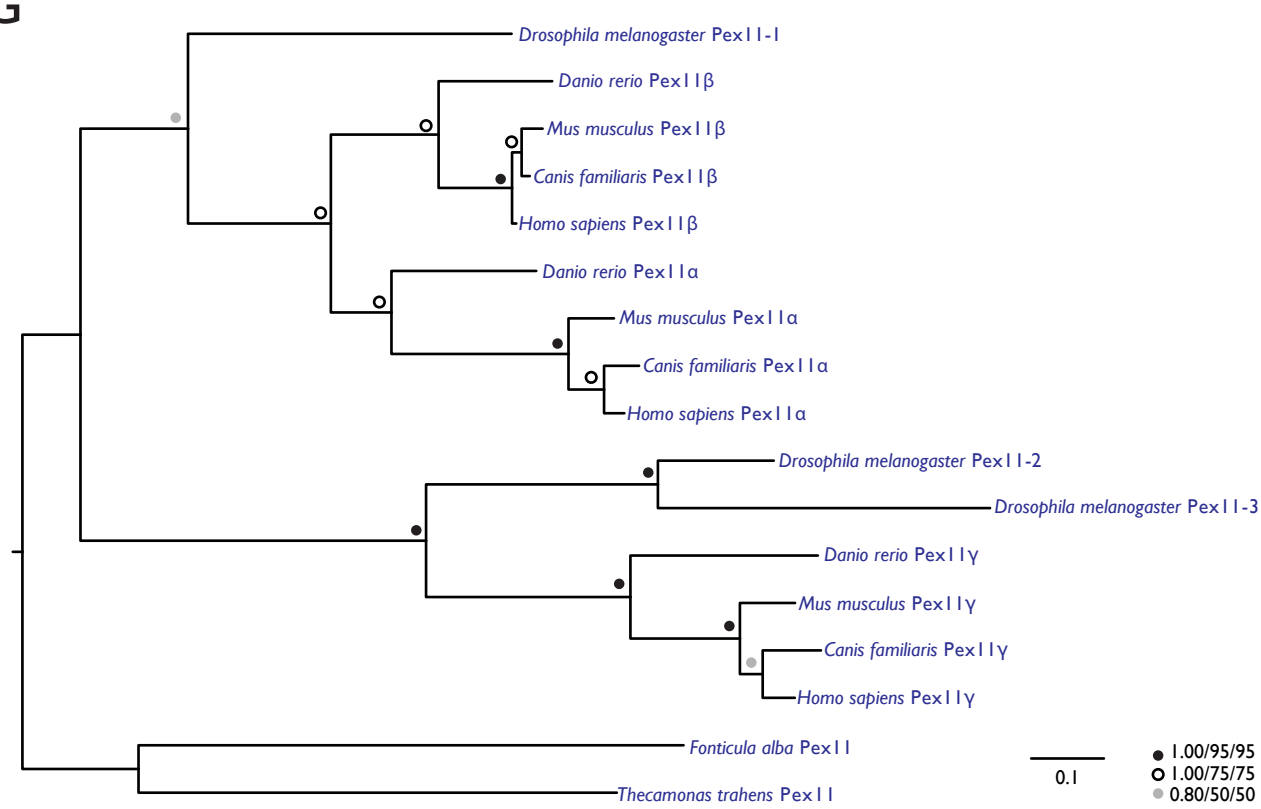
ZORRO - 30% Mask

F



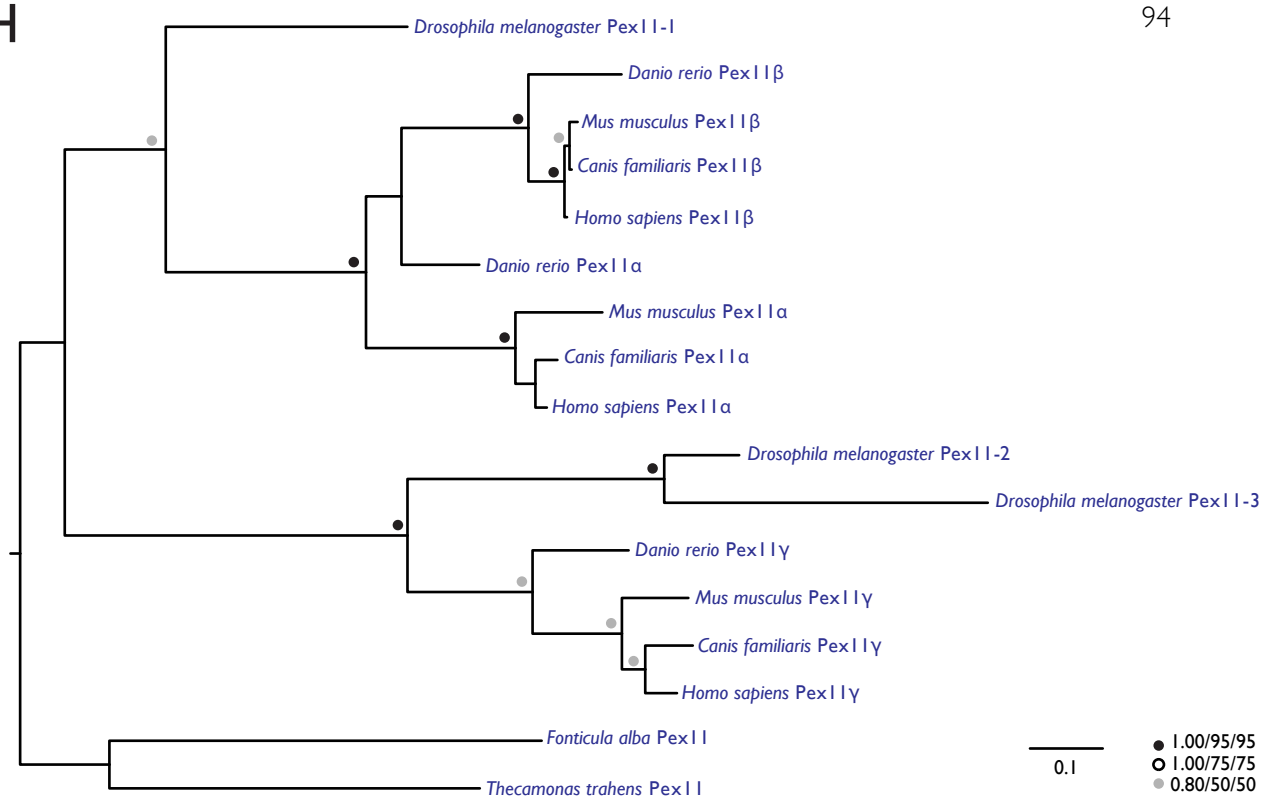
ZORRO - 40% Mask

G



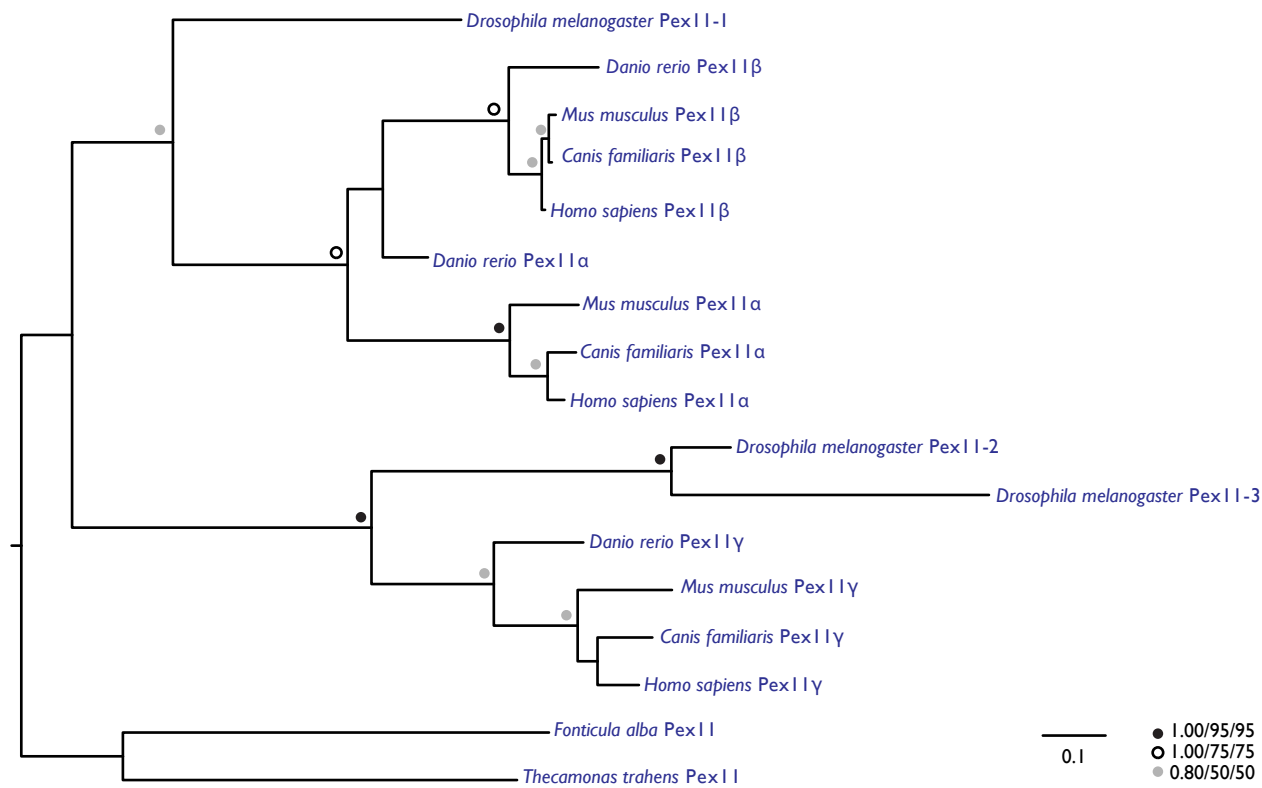
ZORRO - 50% Mask

H



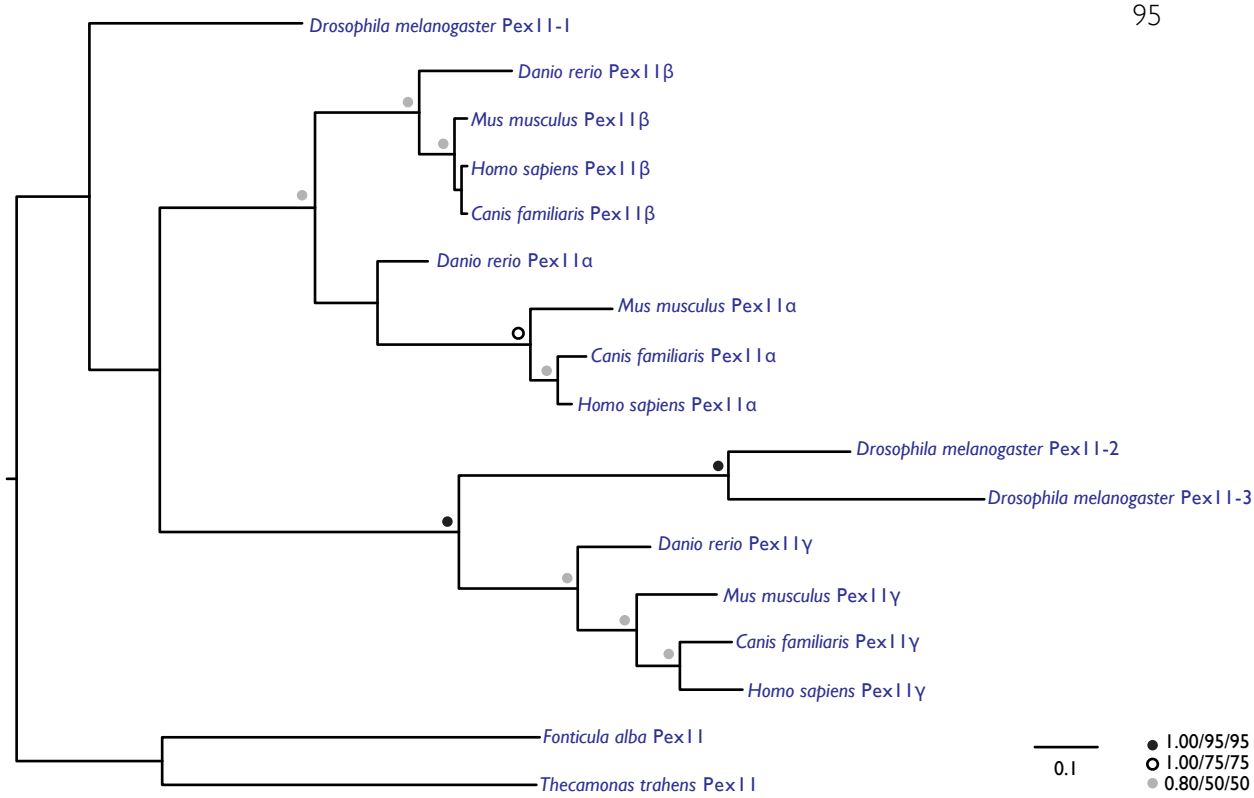
ZORRO - 60% Mask

I



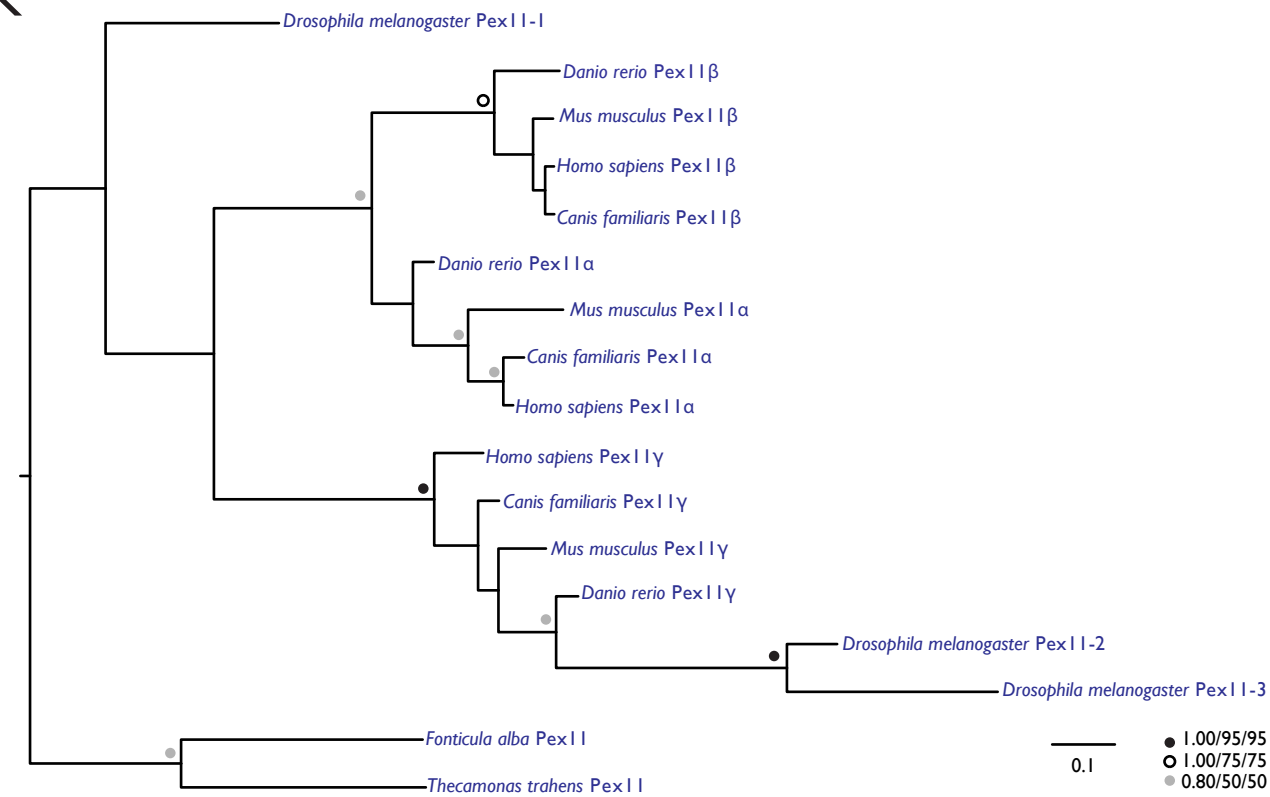
ZORRO - 70% Mask

J



ZORRO - 80% Mask

K



ZORRO - 90% Mask

Pex11p, Pex25p and Pex27p has been previously noted (Tam et al., 2003), suggestive of a potential evolutionary relationship. Our comparative genomics survey confirmed that *S. cerevisiae* Pex25p and Pex27p are homologues, as a pHMMer search into the *S. cerevisiae* genome with Pex25p as query retrieved Pex25p but also Pex27p as the next-best hit (E-value = 5.3×10^{-10}), while a pHMMer search with Pex27p as query retrieved Pex27p but also Pex25p as the next-best hit (E-value = 5.6×10^{-12}). We also investigated the specific evolutionary relationship between Pex25p and Pex27p, and Pex11p. A Pex11p hidden Markov model (HMM) retrieved *S. cerevisiae* Pex25p (E-value = 3.8×10^{-16}) and Pex27p (E-value = 0.0013), in addition to Pex11p (E-value = 2.9×10^{-63}). Similarly, a Pex25p HMM retrieved Pex11p in some Fungi. For example, the *Cryptococcus neoformans* and *Penicillium chrysogenum* genomes do not encode Pex25p homologues. However, a Pex25p HMM, retrieved *C. neoformans* Pex11p (CNAG_07341, E-value = 0.023) and *P. chrysogenum* Pex11p (XP_002557767.1, E-value = 0.00062). A Pex27p HMM did not retrieve Pex11p, probably because Pex27p is restricted to the genus *Saccharomyces*. A homologous relationship was not detected between any of the *S. cerevisiae* Pex11 family proteins and Pex34p, another protein that functions with the Pex11p family to regulate peroxisome size and number in *S. cerevisiae* (Tower et al., 2011).

Are the proposed Pex11 family proteins Pex11Bp, Pex11Cp and Pex11/25p genuine Pex11p homologues? HMMer searches with Pex11Bp or Pex11Cp as queries retrieved Pex11p homologues in yeast, in addition to Pex11Bp or Pex11Cp homologues. Likewise, HMMer searches with Pex11p queries retrieved some yeast Pex11Bp and Pex11Cp homologues, as well as Pex11p homologues. For example, Pex11Bp and Pex11Cp HMMs retrieved *Penicillium chrysogenum* Pex11p as the next best hit (XP_002557767.1, E-values = 7.9×10^{-21} and 7.2×10^{-21}, respectively). A Pex11p HMM retrieved *P. chrysogenum* Pex11Bp (XP_002564016.1, E-value = 2.8×10^{-29}) and Pex11Cp (XP_002568919.1, E-value = 8.7×10^{-17}) as next best hits. The identities of some proteins identified in our comparative genomics survey were sometimes ambiguous. For example, *Chlamydomonas reinhardtii*, *Micromonas* sp., *Volvox carteri*, *Bigelowiella natans*, *Tetrahymena thermophila* and *Naegleria gruberi* proteins were retrieved by both Pex11p and Pex11Bp HMMs

(Appendix 2). A Pex11Cp HMM retrieved proteins not only in *Nematostella vectensis*, *Bigelowiella natans*, *Ectocarpus siliculosus*, *Phytophthora ramorum* and *Guillardia theta*, but also retrieved PEX11γ in *Homo sapiens*, *Mus musculus*, *Canis familiaris* and *Danio rerio* (Appendix 3). Pex11/25p was originally identified in *Yarrowia lipolytica* as a protein with weak similarity to both Pex11p and Pex25p (Kiel et al., 2006); however, our alignments do not show evidence of a clear fusion protein (data not shown). Searches with a Pex11p HMM in the *Y. lipolytica* genome retrieved the original *Y. lipolytica* Pex11/25p query (XP_503276.1, E-value = 1.0×10^{-24}) in addition to *Y. lipolytica* Pex11p (XP_501425.1, E-value = 7.1×10^{-66}). HMMer searches with Pex11p or Pex25p HMMs occasionally retrieved other Pex11/25p homologues. For example, Pex11p and Pex25p HMMs retrieved *Batrachochytrium dendrobatidis* Pex11/25p (BDEG_04173, E-value = 2.3×10^{-11}) and *Laccaria bicolor* Pex11/25p (XP_001876771.1, E-value = 1.1×10^{-53}), respectively.

Unlike most other peroxins, which have not undergone expanisions, the Pex11 protein family is more complex in that it is comprised of a number of homologous proteins. We therefore undertook a phylogenetic analysis to elucidate further the evolutionary histories of the Pex11 family proteins in the Opisthokonta supergroup and to clarify the classification of some homologues whose identities were ambiguous based on reciprocal pHMMer searches alone (Figure 3-4). A strongly supported clade comprised of mammalian PEX11γ and fungal Pex11Cp homologues was observed. Strikingly, this suggests that these are specific orthologues of the same gene, and the product of a gene duplication present in at least the ancestor of opisthokonts. This analysis resolved each of the fungal Pex11 family proteins into well supported clades. Pex11Bp, identified in Pezizomycotina Fungi, arose within the Pex11p clade and appears to be the result of a *PEX11* gene duplication in these fungal lineages. Pex27p arose in the Pex25p clade; *PEX25* and *PEX27* have been identified as ohnologues arising from a *Saccharomyces* whole genome duplication (Byrne and Wolfe, 2005). Our phylogenetic analysis shows that Pex11p is the ancestral member of the Pex11 protein family and that other members of this family (Pex11Cp aside) are fungal-specific innovations.

Figure 3-4. Phylogenetic analysis of the Pex11 protein family in Opisthokonta. Node values indicate statistical support by MrBayes/PhyML/RAxML (posterior probability/bootstrap value/bootstrap value), with statistical values for highly supported nodes replaced by symbols as indicated. Best Bayesian topology is shown rooted on the *O. tauri* and *C. merolae* Pex11p sequences as outgroups. Species names are color-coded according to Figure 3-1. A supported clade comprised of Holozoa PEX11γ and fungal Pex11Cp is observed. Fungal Pex11/25p/Pex25p/Pex27p, Holozoa PEX11α/PEX11β, and fungal Pex11p/Pex11Bp are also resolved into separate clades.

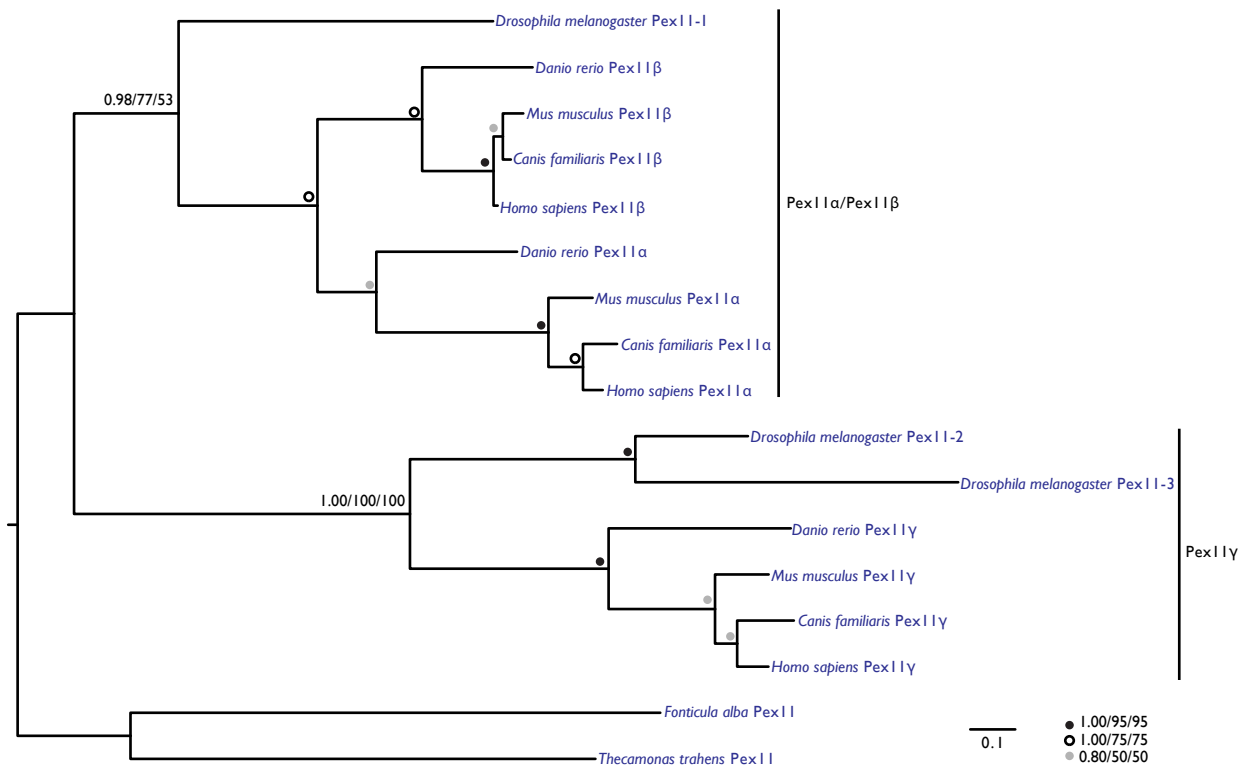

Cyanidioschyzon merolae Pex11
Ostreococcus tauri Pex11-2
Gibberella zeae Pex11C
Neurospora crassa Pex11C
Botryotinia fuckeliana Pex11C
Hysterium pulicare Pex11C
Penicillium chrysogenum Pex11C
Xanthoria parietina Pex11C
Candida albicans Pex11C
Debaryomyces hansenii Pex11C
Hansenula polymorpha Pex11C
Pichia pastoris Pex11C
Yarrowia lipolytica Pex11C
Lipomyces starkeyi Pex11C
Dacryopinax sp. Pex11C
Cryptococcus neoformans Pex11C
Sphaerobolus stellatus Pex11C
Agaricus bisporus Pex11C
Batrachochytrium dendrobatidis Pex11C
Rhizopus oryzae Pex11C
Allomyces macrogynus Pex11C
Rozella allomycis Pex11C
Pex11C, Pex11γ
1.00/100/100
Nematostella vectensis Pex11-2
Mus musculus Pex11γ
Homo sapiens Pex11γ
Canis familiaris Pex11γ
Danio rerio Pex11γ
Drosophila melanogaster Pex11-3
Drosophila melanogaster Pex11-2
Saccharomyces cerevisiae Pex27
Saccharomyces bayanus Pex27
Zygosaccharomyces rouxii Pex25
Saccharomyces cerevisiae Pex25
Saccharomyces bayanus Pex25
Kluyveromyces lactis Pex25
Ashbya gossypii Pex25
Hansenula polymorpha Pex25
Pichia pastoris Pex25
Candida albicans Pex25
Debaryomyces hansenii Pex25
Sphaerobolus stellatus Pex11/25
Agaricus bisporus Pex11/25
Yarrowia lipolytica Pex11/25
Rhizopus oryzae Pex11/25
Rozella allomycis Pex11/25
Batrachochytrium dendrobatidis Pex11/25
Allomyces macrogynus Pex11/25
Pex25, Pex27, Pex11/25
1.00/100/99
1.00/92/88
Drosophila melanogaster Pex11α
Nematostella vectensis Pex11-1
Danio rerio Pex11β
Mus musculus Pex11β
Canis familiaris Pex11β
Homo sapiens Pex11β
Danio rerio Pex11α
Mus musculus Pex11α
Homo sapiens Pex11α
Canis familiaris Pex11α
Pex11α, Pex11β
Botryotinia fuckeliana Pex11B
Gibberella zeae Pex11B
Neurospora crassa Pex11B
Penicillium chrysogenum Pex11B
Allomyces macrogynus Pex11
Batrachochytrium dendrobatidis Pex11
Rhizopus oryzae Pex11
0.95/73/80
Cryptococcus neoformans Pex11
Dacryopinax sp. Pex11
Sphaerobolus stellatus Pex11
Agaricus bisporus Pex11
Gibberella zeae Pex11
Hysterium pulicare Pex11
Penicillium chrysogenum Pex11
Xanthoria parietina Pex11
Botryotinia fuckeliana Pex11
Neurospora crassa Pex11
Hansenula polymorpha Pex11
Pichia pastoris Pex11
Zygosaccharomyces rouxii Pex11
Saccharomyces cerevisiae Pex11
Saccharomyces bayanus Pex11
Kluyveromyces lactis Pex11
Ashbya gossypii Pex11
Candida albicans Pex11
Debaryomyces hansenii Pex11
Yarrowia lipolytica Pex11
Pex11, Pex11B
0.1
1.00/95/95
1.00/75/75
0.80/50/50

It was intriguing to predict three Pex11p homologues in *D. melanogaster*, as previous work only identified one or two homologues (Mast et al., 2011; Faust et al., 2012). Reciprocal pHMMer searches with these proteins into the human genome revealed that one homologue (CG8315) most closely resembles PEX11β, while the remaining two homologues (CG13827 and CG33474) retrieve PEX11γ as the top hit. Therefore, these results do not support these proteins being equivalent to PEX11α, PEX11β and PEX11γ as in mammalian genomes. A phylogenetic analysis of mammalian and *D. melanogaster* Pex11 proteins revealed supported PEX11α/PEX11β and PEX11γ clades (Figure 3-5). As suggested by reciprocal pHMMer results, one homologue is more closely related to PEX11α/PEX11β and the remaining two homologues are more closely related to PEX11γ, and likely being a *D. melanogaster*-specific gene duplication.

3.4 Conserved features of the *Y. lipolytica* Pex11 protein family

Our current and early (Kiel et al., 2006; Schlüter et al., 2006; Gabaldón et al., 2006) comparative genomics surveys and phylogenetic analyses have elucidated the evolution of the Pex11 protein family in Fungi and other eukaryotes. We chose to focus on the *Y. lipolytica* Pex11 protein family, *i.e.* Pex11p, Pex11Cp and Pex11/25p, for additional analyses. *Y. lipolytica* shares characteristics with both the pezizomycete and the saccharomycete yeasts, and its genome has been shown previously to be taxonomically informative for understanding the evolution of Rab-GTPases in yeasts (Pereira-Leal, 2008). *Y. lipolytica* is also a very effective model for peroxisome biogenesis because it efficiently utilizes hydrophobic substrates, accompanied by extensive peroxisome proliferation (Nicaud, 2012). Curiously, although all three *Y. lipolytica* Pex11 family proteins are retrieved by a Pex11p HMM, only Pex11p, and not Pex11Cp or Pex11/25p, was positively verified with a pHMMer search against the *S. cerevisiae* genome (Appendix 1). *Y. lipolytica* Pex11p and Pex11Cp were positively verified with pHMMer searches against the *H. sapiens* genome (E-values = 2.7×10^{-13} and 0.04, respectively). However, I have demonstrated that there is evidence for Pex11Bp, Pex11Cp and Pex11/25p being genuine Pex11p homologues (Section 3.3).

Figure 3-5. Phylogenetic analysis of the Pex11 protein family in *Drosophila melanogaster* and other Opisthokonta. Node values indicate statistical support by MrBayes/PhyML/RAxML (posterior probability/bootstrap value/bootstrap value), with statistical values for highly supported nodes replaced by symbols as indicated. Best Bayesian topology is shown rooted on the *F. alba* and *T. trahens* Pex11p sequences as outgroups. Species names are color-coded according to Figure 3-1. PEX11α/PEX11β and PEX11γ clades are supported, with *D. melanogaster* Pex11p homologues grouping as shown.


Drosophila melanogaster Pex11-1
Danio rerio Pex11β
Mus musculus Pex11β
Canis familiaris Pex11β
Homo sapiens Pex11β
Danio rerio Pex11α
Mus musculus Pex11α
Canis familiaris Pex11α
Homo sapiens Pex11α
Pex11α/Pex11β
Drosophila melanogaster Pex11-2
Drosophila melanogaster Pex11-3
Danio rerio Pex11γ
Mus musculus Pex11γ
Canis familiaris Pex11γ
Homo sapiens Pex11γ
Pex11γ
Fonticula alba Pex11
Thecamonas trahens Pex11
0.98/77/53
1.00/100/100
0.1
1.00/95/95
1.00/75/75
0.80/50/50

3.4.1 The *Y. lipolytica PEX11* family contains putative Far binding sites

The expression search program dreg (emboss.bioinformatics.nl/cgi-bin/emboss/dreg) was used to search the promoter region of each of the *PEX11* family genes for putative oleate response elements (OREs). A region of 500 nucleotides upstream of the start codon was queried for the ORE consensus sequence C-(C/A/G)-G-N_{14-18}-C-(C/G)-G. This search allowed for deviations from the canonical ORE sequence previously published for *PEX* genes and which had been deemed not critical for its function (Rottensteiner et al., 2002; 2003; Vizeacoumar et al., 2004).

Two OREs were predicted for the *PEX11/25* gene. Both predicted OREs contain conserved TNA triplets, but neither is palindromic, which does not necessarily preclude them from functioning as true OREs (Table 3-1). Several other putative OREs identified in these searches have 14 intervening nucleotides, which has been previously published as an exception to the known consensus sequence (Rottensteiner et al., 2002).

Table 3-1. Putative oleate reponse elements (OREs) and UAS1 sites in the promoters of *Y. lipolytica PEX11* family genes

Gene	ORE (position)	UAS1 (position)
PEX11 (XP_501425.1)	CCGN$_{10}$TTGNCCG (-317)	CCCCAGN$_{15}$GTGGTG (-87)
		CTCCGTN$_{16}$ATGGTG (-168)
PEX11C (XP_501447.1)	CGGN$_{11}$TTGN$_4$CGG (-103)	-
PEX11/25 (XP_503276.1)	CAGN$_2$TCAN$_{13}$CGG (-176)	TCCCGTN$_{29}$TTGGAG (-239)
	CAGN$_8$TTAN$_3$CCG (-254)	CCCCTTN$_{32}$TTGGTG (-345)

The expression search program dreg was also used to search the promoter regions of the *PEX11* family genes for putative UAS1 binding sites. Again, a region of 500 nucleotides upstream of the start codon was queried for the UAS1 consensus sequence C-(T/C)-C-C-(A/G/T)-(A/T/G)-N_{4-36}-(A/T/C/G)-(T/C)-G-G-(A/G)-G, which contains deviations from the canonical UAS1 binding site sequence that are not critical for function (Rottensteiner et al., 2003). These searches were also repeated with the consensus sequence (C/T)-(T/C/G)-C-C-(A/G/T)-(A/T/G)-N_{4-36}-(A/T/C/G)-(T/C)-G-G-(A/G/T)-G, which allows for deviations from the canonical UAS1 binding site sequence

that would result in a non-functional element, *i.e.* allowing for a T at position one, G at position two, and T in the second to last position (Rottensteiner et al., 2003). No canonical UAS1 sites were found in the promoters of the *Y. lipolytica PEX11* family genes. Only when deviations from the consensus sequence that would exclude these motifs as functional UAS1 sites were allowed did possibilities arise, two for each in *PEX11* and *PEX11/25* (Table 3-1). Interestingly, the predicted UAS1 site for *PEX11/25* at -239 would overlap with the ORE at position at -254, which is typical of both of these elements. In summary, these motifs were not convincing OREs or UAS1 sites, and it is not known if they are functional.

The lack of canonical OREs and UAS1 sites in the genes of the *Y. lipolytica* Pex11 protein family may be due to the fact that an entirely different system is responsible for the mechanism of transcriptional regulation of peroxisome proliferation in this yeast. Previous bioinformatic (Hynes et al., 2006) and functional (Poopanitpan et al., 2010) work on this process in *Y. lipolytica* suggests that this is in fact the case, and that *Y. lipolytica* uses the FarA/B system used by *Aspergillus nidulans*.

The dreg program was used to search 1 kbp upstream of the start codon for each of the *Y. lipolytica PEX11* family genes for a conserved CCTCGG and its complement CGGAGG, the binding site for FarA/B in *A. nidulans* (Hynes et al., 2006). As previously reported (Hynes et al., 2006), CCTCGG sites were present in *PEX11* (Table 3-2). and *PEX11C*, but not in PEX11/25. Importantly, as previously noted (Poopanitpan et al., 2010), it is not known at present whether mutation of this site impairs gene function in *Y. lipolytica*. However, these results do pose the possibility that some members of the *PEX11* family in *Y. lipolytica* may be transcriptionally regulated in response to the presence of fatty acid.

Table 3-2. Putative CCTCGG sites in the *Y. lipolytica PEX11* gene family promoters

Gene	CCTCGG site(s)
PEX11	-115
PEX11C	-106
	-353
	-415
PEX11/25	-

3.4.2 Putative phosphorylation sites in the *Y. lipolytica* Pex11 protein family

Phosphorylation of *S. cerevisiae* and *P. pastoris* Pex11 proteins has been shown to be necessary for their role in peroxisome dynamics (Knoblach and Rachubinski, 2010; Joshi et al., 2012). These observations raised the question of whether or not similar regulation may be possible in *Y. lipolytica*, and if other members of the Pex11p family could have similar phosphorylation at certain residues.

Alignments of Pex11 protein sequences, using either sequences from Fungi exclusively or from other eukaryotes in addition to Fungi, revealed that the *S. cerevisiae* Pex11p sites Ser^{165} and Ser ^{167} and the *P. pastoris* Pex11p site Ser^{167} were not well conserved (data not shown). Next, NetPhos2.0 (cbs.dtu.dk/services/NetPhos-2.0) (Blom et al., 1999) was used to predict phosphorylation sites at serine, threonine and tyrosine residues in Pex11 family proteins. Phosphorylation sites at all three residues were common in the sequences queried (Table 3-3). The majority of these sites are probably not functional: residues above a specific score threshold are predicted as phosphorylation sites, but residues with higher scores represent higher confidence predictions (Blom et al., 1999). For instance, for all of the predicted phosphorylation sites in the *S. cerevisiae* and *P. pastoris* Pex11 proteins, only two sites in *S. cerevisiae* and one site in *P. pastoris* were shown to be actually functional (Knoblach and Rachubinski, 2010; Joshi et al., 2012). *Y. lipolytica* Pex11p is has a putative phosphorylation site at Ser^{174}, but pairwise alignments with both *S. cerevisiae* and *P. pastoris* indicate that this is not a conserved position (data not shown). Pairwise alignment of *S. cerevisiae* and *P. pastoris* demonstrated that Ser^{165} and Ser^{173} are conserved (data not shown). However, it is not known if any of the residues in *Y. lipolytica* are actually phosphorylated and if phosphorylation has any effect on peroxisome dynamics.

Table 3-3. Putative phosphorylation sites for Pex11 family proteins in selected yeasts

Protein	Serine sites	Threonine sites	Tyrosine sites
S. cerevisiae Pex11p	Ser^{22}, Ser^{153}, Ser^{165}, Ser^{167}, Ser^{220}	Thr^{57}, Thr^{123}, Thr^{152}	Tyr^{187}, Tyr^{209}, Tyr^{216}
S. cerevisiae Pex25p	Ser^{11}, Ser^{22}, Ser^{35}, Ser^{36}, Ser^{44}, Ser^{47}, Ser^{55}, Ser^{56}, Ser^{58}, Ser^{63}, Ser^{82}, Ser^{84},	Thr^{176}, Thr^{361}, Thr^{385}	Tyr^{18}, Tyr^{23}, Tyr^{239}, Tyr^{339}, Tyr^{366}

	Ser^{165}, Ser^{243} Ser^{247}, Ser^{257}, Ser^{302}, Ser^{304}, Ser^{346}, Ser^{370}, Ser^{388}, Ser^{391}		
S. cerevisiae Pex27p	Ser^{23}, Ser^{42}, Ser^{82}, Ser^{85}, Ser^{88}, Ser^{95}, Ser^{236}, Ser^{281}, Ser^{341}, Ser^{368}	Thr^{2}, Thr^{16}, Thr^{84}, Thr^{245}	Tyr^{102}, Tyr^{192}, Tyr^{252}, Tyr^{279}, Tyr^{297}, Tyr^{339}, Tyr^{345}
P. pastoris Pex11p	Ser^{22}, Ser^{81}, Ser^{98}, Ser^{125}, Ser^{158}, Ser^{164}, Ser^{173}, Ser^{191}	Thr^{11}, Thr^{55}, Thr^{91}, Thr^{180}	Tyr^{106}, Tyr^{246}
P. pastoris Pex11Cp	Ser^{12}, Ser^{29}, Ser^{55}, Ser^{59}, Ser^{62}, Ser^{68}, Ser^{83}, Ser^{106}, Ser^{160}, Ser^{184}, Ser^{190}, Ser^{191}	Thr^{187}	Tyr^{17}
P. pastoris Pex25p	Ser^{46}, Ser^{50}, Ser^{53}, Ser^{55}, Ser^{58}, Ser^{68}, Ser^{88}, Ser^{129}, Ser^{164}, Ser^{168}, Ser^{204}, Ser^{205}, Ser^{222}, Ser^{259}, Ser^{319}, Ser^{402}, Ser^{405}, Ser^{407}, Ser^{410}, Ser^{418}, Ser^{422}, Ser^{424}, Ser^{426}, Ser^{436}, Ser^{500}	Thr^{41}, Thr^{67}, Thr^{119}, Thr^{162}, Thr^{163}, Thr^{210}, Thr^{270}, Thr^{309}, Thr^{343}	Tyr^{279}, Tyr^{338}, Tyr^{453}, Tyr^{454}
H. sapiens Pex11α	Ser^{13}, Ser^{43}, Ser^{57}, Ser^{157}, Ser^{225}	Thr^{127}, Thr^{240}	-
H. sapiens Pex11β	Ser^{51}, Ser^{112}, Ser^{159}, Ser^{161}	Thr^{54}, Thr^{150}	Tyr^{23}, Tyr^{140}, Tyr^{215}
H. sapiens Pex11γ	Ser^{8}, Ser^{11}, Ser^{38}, Ser^{70}, Ser^{160}, Ser^{168}	Thr^{178}	Tyr^{140}
Y. lipolytica Pex11p	Ser^{63}, Ser^{84}, Ser^{136}, Ser^{174}	Thr^{39}, Thr^{51}, Thr^{98}, Thr^{130}	Tyr^{96}, Tyr^{105}, Tyr^{116}
Y. lipolytica Pex11Cp	Ser^{13}, Ser^{90}, Ser^{119}, Ser^{211}	Thr^{28}, Thr^{221}, Thr^{222}, Thr^{295}	Tyr^{134}, Tyr^{163}
Y. lipolytica Pex11/25p	Ser^{15}, Ser^{51}, Ser^{56}, Ser^{133}, Ser^{225}, Ser^{283}	Thr^{18}, Thr^{154}, Thr^{173}, Thr^{237}	Tyr^{179}

3.5 Deletion of the *Y. lipolytica PEX11* family genes

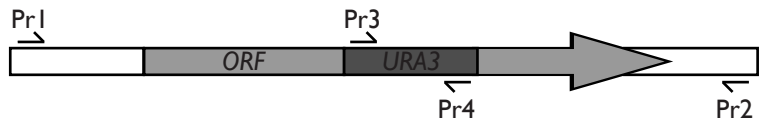
In *Y. lipolytica*, all members of the Pex11 protein family have only been predicted *in silico*, and nothing is known regarding their localization in the cell or if they have a role in peroxisome biogenesis. We therefore chose to extend our evolutionary study by functionally characterizing this Pex11 protein family. Pex11p and Pex11Cp share 9.9% identity and 25.2% similarity. Pex11p and Pex11/25p share 12.1% identity and 25.5% similarity. Pex11Cp and Pex11/25p share 9.9% identity and 28.7% identity (Figure 3-6A). The *Y. lipolytica PEX11*, *PEX11C* and *PEX11/25* genes were

Figure 3-6. PCR confirmation of *pex11*Δ, *pex11C*Δ and *pex11/25*Δ deletion strains. (A) Sequence alignment of *Y. lipolytica* Pex11p, Pex11Cp and Pex11/25p. Amino acid sequences were aligned with MUSCLE 3.6. Identical residues (black) and similar residues (gray) are shaded, and dashes represent gaps. Similarity rules: G = A = S; A = V; V = I = L = M; I = L = M = F = Y = W; K = R = H; D = E = Q = N; and S = T = Q = N. (B) PCR strategy used to confirm deletion strains. The ORF of the gene of interest with genomically integrated *URA3* is shown, together with the layout of four PCR primers used to confirm gene disruption. (C) Results of Checking PCR using genomic DNA of the deletion strains and the corresponding wild-type strain, and the primer pair indicated in parentheses. Lanes 1 and 2, PCR of deletion strain genomic DNA (~1.5 kbp product). Lanes 3 and 4, negative control with corresponding wild-type genomic DNA (no product). Lane 5, positive control with corresponding wild-type genomic DNA (~2 kbp product).

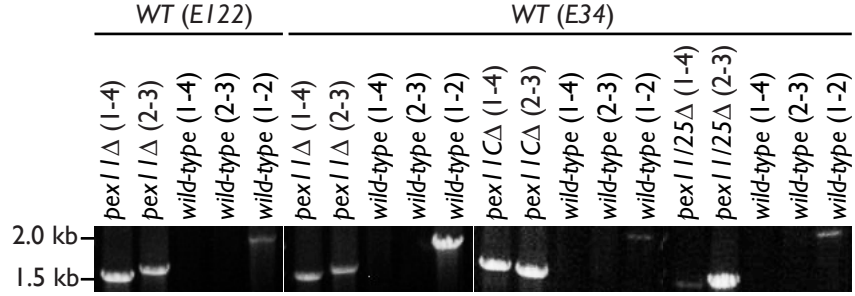
A

```
Pex11p    ------------------------------------------------------------MSVCLA  6
Pex11Cp   ------------------------------------------------------------MSLAYT  6
Pex11/25p MVSKLVGRTKPAGDSQKTVRIDEKVEIKEFQAHKPEIIDNTTEIPAPVPMSIAERSIADA 60

Pex11p    QNPTVT-----RVVKLLETHVGRDKILRSIQYFSRFLTYYLF------------------ 43
Pex11Cp   VNLLRASGRLNKVLALLSNAGTLDPTLAFVGYGSLFVCELLRLESVQMGVVVAKVLLGQV 66
Pex11/25p ARPPLHASPWRVFLTLFAEKGGLDKTIKLIQYTGRLLLWAAK------------------ 102

Pex11p    -----RKGYTK-----------DTTDIFRKIQNQFSMARKLPRVGKPIGHLKTAAVSFEN 87
Pex11Cp   TGLDIIESFEKLGLASDLTGVADSLGVLSGLISDIRIFNRLWGLVGLISWGISEYDAQQK 126
Pex11/25p -----QGWFTRHKQMMLLWKLAEMESRLNGMVSNFSQFRKIIKLGEWLG-PVEDLVTTKN 156

Pex11p    --KTLDPCLRYTTIGRNL------GYAIYLVFDSIIYINGSGIKKIDNIKTIKKVGSYFW 139
Pex11Cp   PLDVNDPYIKYDKAVRVITALQVWSNLFYQPLENVAYLGMHKLLPVSDA-----AQNSLW 181
Pex11/25p PLTSLAFQSELMEVINTI------GDDIY-CLSKIGVVKGKRLGRNGEL-----MANWGW 204

Pex11p    AFG-IFCNILNSIHKINICKKKRAAIAAEKEKDTTSAKKNDKDAAAAQKQLV-------- 190
Pex11Cp   LHSSKLWALHVVLELIRLVVEIRRNVKRQSIEKKAGKTTTTTDTKACNCTKLLPNFLKGL 241
Pex11/25p Y-GAIFTNIKVGIETYQL----------------AAKSGDEAA-------I-------- 231

Pex11p    ----WDLLDFSIPLTSL-GYIHLD----DGLV-----GLAGFATGIVGVQKAWAATA--- 233
Pex11Cp   GDQWWRDLVINLAYLPLTAHWSID----GGIVDNLIVGFLGASAAAANVGWKWKAL-TP 296
Pex11/25p ----WDAKLTLFKLANDFIGTIDCLEPEGLS-NIYQTVTGLASGSVGFYKLWRKISKKL 286

Pex11p    ------------- 233
Pex11Cp   DVPVEKKMQ---- 305
Pex11/25p DKELEEKERCKTC 299
```

B



C



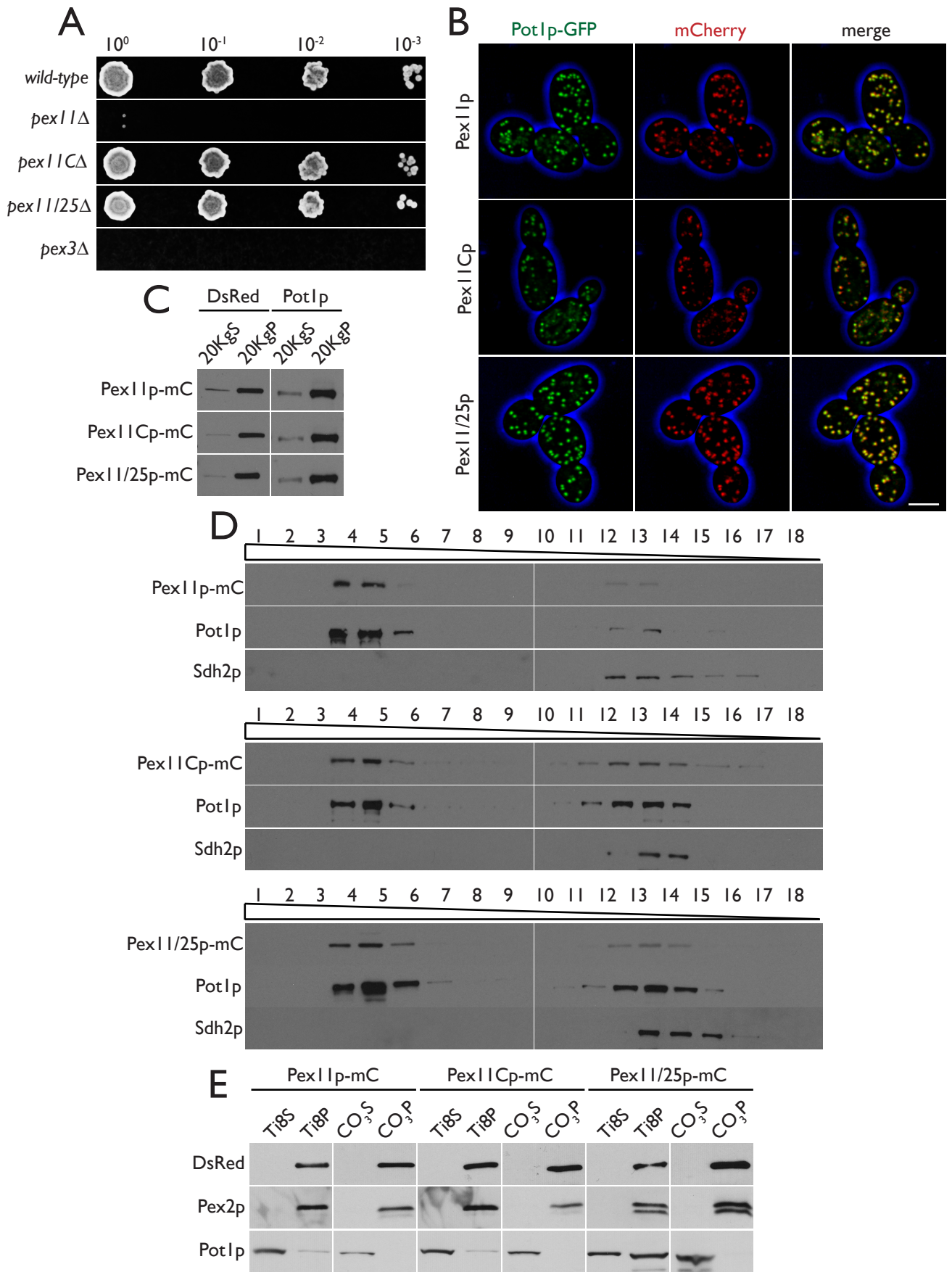
disrupted through homologous transformation of yeast using a fusion PCR-based integrative procedure (Davidson et al., 2002). To confirm correct deletion of genes, chromosomal DNA was isolated from the deletion strains, and PCR reactions were carried out using primer pairs directly either upstream or downstream of the gene ORF and within the *URA3* selectable marker (Figure 3-6B). The correctly sized PCR product (~1.5 kbp) was observed for all deletion strains (Figure 3-6C). No PCR products were observed for the negative controls with the corresponding wild-type genomic DNA, and the correctly size PCR product (~2 kbp) was observed for all positive control reactions.

3.6 *Y. lipolytica* Pex11 protein family members are peroxisomal integral membrane proteins

Deletion of genes like *PEX3* encoding a peroxin essential for formation of the peroxisomal membrane leads to abnormal peroxisome assembly and failure of yeast strains to grow on medium containing oleic acid, a carbon source whose metabolism requires functional peroxisomes (Erdmann et al., 1989). Growth of *pex11C*Δ and *pex11/25*Δ strains on oleic acid-containing medium was similar to that of the wild-type *E34* strain (Figure 3-7A). In contrast, the *pex11*Δ strain like the *pex3*Δ strain displayed no growth on oleic acid-containing medium, and thus Pex11p does not share redundant function with Pex11Cp or Pex11/25p. Our findings also suggested that *Y. lipolytica pex11*Δ cells lack functional peroxisomes, in contrast to what has been observed to date for other eukaryotic cells lacking functional Pex11p which contain fewer and enlarged peroxisomes, prompting further investigation into a possible role for Pex11p in peroxisome assembly in *Y. lipolytica*.

We next determined whether the *Y. lipolytica* Pex11 protein family members are peroxisomal proteins (Figure 3-7B). We tagged Pex11p, Pex11Cp and Pex11/25p at their C-termini with mCherry (mC) and visualized their localization by confocal microscopy. Pex11p-mC, Pex11Cp-mC and Pex11/25p-mC colocalized with the GFP-tagged peroxisomal matrix protein thiolase (Pot1p-GFP) to punctate structures characteristic of peroxisomes. Subcellular fractionation showed

Figure 3-7. *Y. lipolytica* Pex11p, Pex11Cp and Pex11/25p are peroxisomal integral membrane proteins. (A) *pex11*Δ cells, but not *pex11C*Δ or *pex11/25*Δ cells, display a growth defect on oleic acid-containing medium. Cells of the wild-type strain *E34* and the designated gene deletion strains were grown in YEPD medium, spotted onto oleic acid-containing YPBO agar at 10-fold serial dilutions from 10^0 to 10^{-3}, and grown for 2 d at 28°C. (B) Pex11p-mCherry (mC), Pex11Cp-mC and Pex11/25p-mC colocalize with the GFP-tagged peroxisomal matrix enzyme thiolase (Pot1p-GFP) by confocal microscopy. Bar, 5 μm. (C) Pex11p-mC, Pex11Cp-mC and Pex11/25p-mC localize preferentially to the 20KgP subcellular fraction enriched for peroxisomes. Immunoblotting was performed using antibodies to Pot1p and to DsRed to detect mCherry-tagged proteins. (D) Pex11p-mC, Pex11Cp-mC and Pex11/25p-mC cofractionate with the peroxisomal marker Pot1p. Organelles in the 20KgP fraction were subjected to density gradient centrifugation on a discontinuous Nycodenz gradient. Fractions were collected from the bottom of the gradient, and immunoblotting was performed using antibodies to DsRed to detect mCherry-tagged proteins. Fractions enriched for peroxisomes and mitochondria were identified using antibodies to Pot1p and Sdh2p, respectively. (E) Pex11p, Pex11Cp and Pex11/25p are peroxisomal integral membrane proteins. The 20KgP fraction from cells expressing Pex11p-mC, Pex11Cp-mC or Pex11/25p-mC was treated dilute alkali Tris buffer to rupture peroxisomes and then subjected to ultracentrifugation to yield a supernatant (Ti8S) fraction enriched for matrix proteins and a pellet (Ti8P) fraction enriched for membrane proteins. The Ti8P fractions were further treated with alkali Na_2CO_3 and separated by ultracentrifugation into a supernatant (CO_3S) fraction enriched for peripherally associated membrane proteins and a pellet (CO_3P) fraction enriched for integral membrane proteins. Pex11p-mC, Pex11Cp-mC and Pex11/25p-mC all enriched in the Ti8P and CO_3P fractions enriched for membrane proteins and integral membrane proteins, respectively, as did the known peroxisomal integral membrane protein Pex2p. Pot1p marked the fractionation profile of a peroxisomal matrix protein. Immunoblotting was performed with antibodies to DsRed, Pex2p and Pot1p.


A
10^0
10^-1
10^-2
10^-3
wild-type
pex11Δ
pex11CΔ
pex11/25Δ
pex3Δ
B
Pot1p-GFP
mCherry
merge
Pex11p
Pex11Cp
Pex11/25p
C
DsRed
Pot1p
20KgS
20KgP
20KgS
20KgP
Pex11p-mC
Pex11Cp-mC
Pex11/25p-mC
D
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Pex11p-mC
Pot1p
Sdh2p
Pex11Cp-mC
Pot1p
Sdh2p
Pex11/25p-mC
Pot1p
Sdh2p
E
Pex11p-mC
Pex11Cp-mC
Pex11/25p-mC
Ti8S
Ti8P
CO3S
CO3P
DsRed
Pex2p
Pot1p

that Pex11p-mC, Pex11Cp-mC and Pex11/25p-mC localized preferentially to a 20,000 × *g* pellet fraction (20KgP) enriched for peroxisomes and not to a supernatant fraction (20KgS) enriched for cytosol (Figure 3-7C). Density gradient centrifugation of the 20KgP fraction showed that these three proteins cofractionate with the peroxisomal marker Pot1p but not with the mitochondrial marker Sdh2p (Figure 3-7D). Therefore, these data show that the bioinformatically predicted proteins Pex11p, Pex11Cp and Pex11/25p are peroxisomal proteins in *Y. lipolytica*.

We next determined whether Pex11p, Pex11Cp and Pex11/25p associate with the peroxisomal membrane. The 20KgP fraction was subjected to hypotonic lysis in dilute alkali Tris buffer followed by centrifugation (Figure 3-7E). Pex11p-mC, Pex11Cp-mC and Pex11/25p-mC localized preferentially to the pellet fraction (Ti8P) enriched for membrane proteins like the peroxisomal integral membrane protein Pex2p (Eitzen et al. 1996) and not to the supernatant fraction (Ti8S) enriched for matrix proteins like the peroxisomal matrix protein Pot1p. Ti8P fractions were next subjected to extraction with alkali sodium carbonate followed by centrifugation (Fujiki et al., 1982). Pex11p-mC, Pex11Cp-mC and Pex11/25p-mC localized preferentially to the pellet fraction (CO_3P) enriched for integral membrane proteins, as did Pex2p and in contrast to Pot1p (Figure 3-7E). Pex11p is predicted to have two transmembrane helices (between amino acids 103-122 and 137-156) (cbs.dtu.dk/services/TMHMM-2.0). Although no transmembrane domain was predicted for Pex11Cp or Pex11/25p, these two proteins contain regions of hydrophobicity, suggesting possible integral association with the peroxisomal membrane without the presence of a defined transmembrane region (data not shown). Taken together, these observations indicate that Pex11p, Pex11Cp and Pex11/25p are peroxisomal integral membrane proteins.

3.7 Peroxisomes are absent in *pex11*Δ cells, and are larger and fewer in number in *pex11C*Δ and *pex11/25*Δ cells

As the inability to grow on medium containing oleic acid is often indicative of defective peroxisome assembly (Erdmann et al., 1989), we examined peroxisome morphology in the *pex11*Δ strain. We

also examined whether deletion of *PEX11C* or *PEX11/25* affected peroxisome dynamics, particularly peroxisome division, which is well documented for deletions of genes encoding other Pex11 protein family members. We used confocal microscopy to observe wild-type, *pex11*Δ, *pex11C*Δ and *pex11/25*Δ cells expressing genomically integrated Pot1p-GFP following a shift from growth in glucose-containing medium to oleic acid-containing medium (Figure 3-8A). Upon transfer to oleic acid-containing medium, wild-type cells, as well as *pex11C*Δ and *pex11/25*Δ cells, displayed characteristic punctate peroxisomes that increased in number with prolonged incubation in oleic acid-containing medium. In contrast, at the time of transfer to oleic acid-containing medium, the Pot1p-GFP-containing structures in *pex11*Δ cells were much more heterogeneous in appearance and fewer in number as compared to the structures observed in wild-type cells. With prolonged incubation in oleic acid-containing medium, the labeled structures in *pex11*Δ cells became highly elongated and tendrillar. Confocal microscopy of wild-type and *pex11*Δ cells expressing Pot1p-mRFP showed that Pot1p-mRFP is cytosolic in *pex11*Δ cells and lack completely the readily identifiable punctate peroxisomes still observed in wild-type cells (Figure 3-8B). We speculate that the elongated structures observed in *pex11*Δ cells expressing Pot1p-GFP may be an artifact of oligomerization of Pot1p-GFP that is mislocalized to the cytosol, a phenomenon that we have previously observed (Chang et al., 2013).

Cells of the *pex11C*Δ and *pex11/25*Δ strains expressing Pot1p-GFP and grown in glucose-containing medium showed no significant difference in peroxisome number compared to wild-type cells by confocal micrsocopy (Figure 3-8C). In contrast, *pex11C*Δ and *pex11/25*Δ cells grown in oleic acid-containing medium contained fewer peroxisomes compared to wild-type cells (Figure 3-8C).

Thin section transmission electron microscopy (EM) showed canonical circular peroxisomes in wild-type cells (Figure 3-9B) but no evidence of wild-type peroxisomes in *pex11*Δ cells (Figure 3-9A). However, clusters of small vesicles were observed in some *pex11*Δ cells (Figure 3-8A, boxed regions and insets). These findings, together with evidence from confocal microscopy (Figure 3-8),

Figure 3-8. Peroxisomes are absent in *pex11*Δ cells and fewer in number in *pex11C*Δ and *pex11/25*Δ cells. (A) Wild-type *E34*, *pex11*Δ, *pex11C*Δ and *pex11/25*Δ cells expressing genomically integrated *POT1-GFP* were grown in glucose-containing YEPD medium, transferred to oleic acid-containing YPBO medium and incubated in YPBO medium for the times indicated. Confocal fluorescence microscopy images were acquired every 2 h. Bar, 5 μm. (B) Wild-type *E122* and *pex11*Δ cells containing plasmid expressing Pot1p-mRFP were grown in YEPD medium supplemented with hygromycin B (YEPD), or grown in YEPD medium supplemented with hygromycin B and then transferred to YPBO medium supplemented with hygromycin B for 8 h (YPBO). Representative confocal microscopy images are shown. Bar, 5 μm. (C) Wild-type *E34*, *pex11C*Δ and *pex11/25*Δ cells expressing genomically integrated *POT1-GFP* were grown in YEPD medium to log phase and imaged by confocal microscopy (YEPD), or grown in YEPD medium, transferred to YPBO medium, grown to log phase and imaged (YPBO). Representative images are shown. Bar, 5 μm. The number of peroxisomes (Pot1p-GFP-containing puncta) per μm^2 was determined with Imaris software. Graphs display the results of three independent experiments for each strain. Error bars give the SEM. Significance was determined by *t*-test with respect to the wild-type strain. N.S., not significant; **, $p < 0.01$.

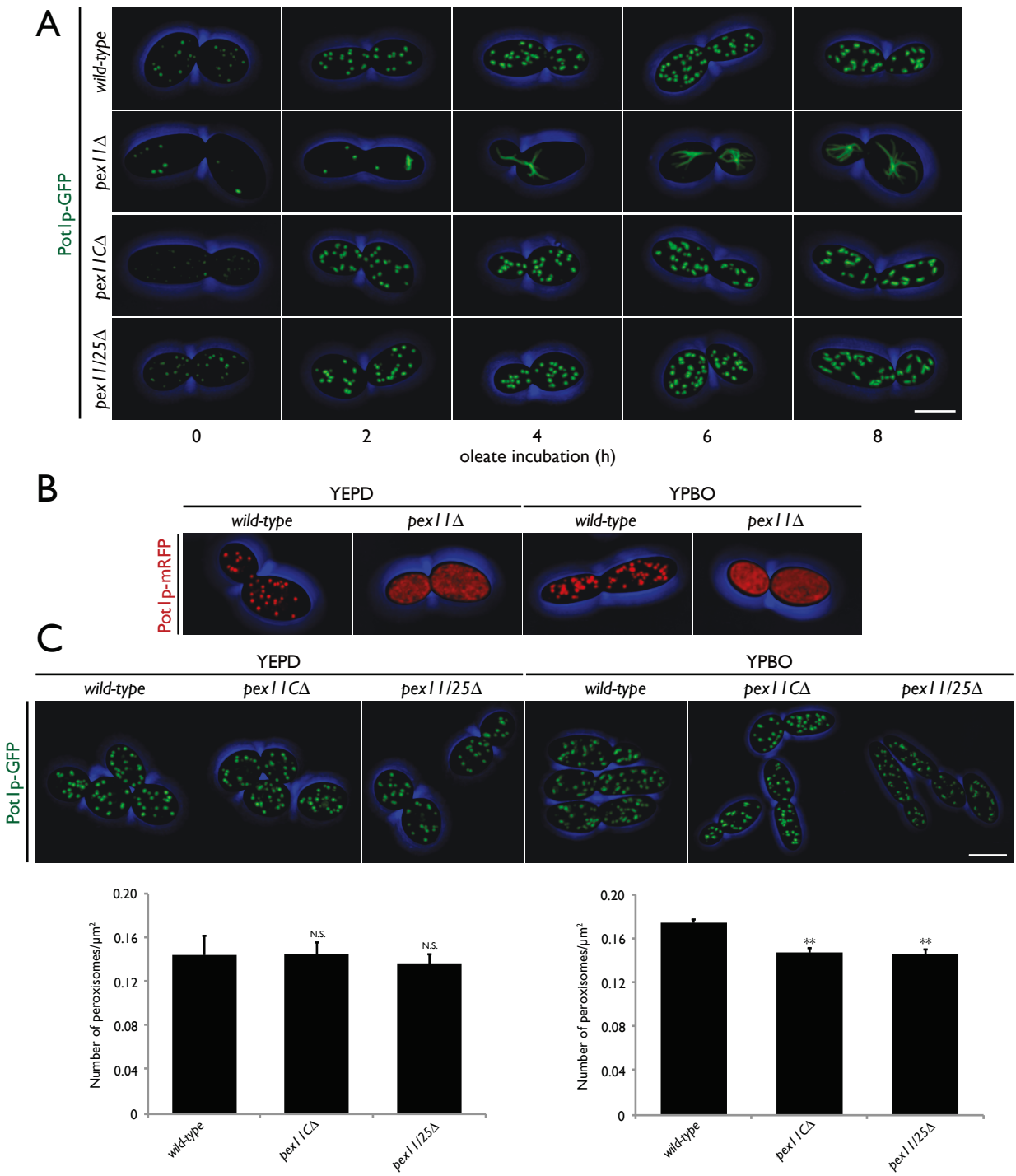
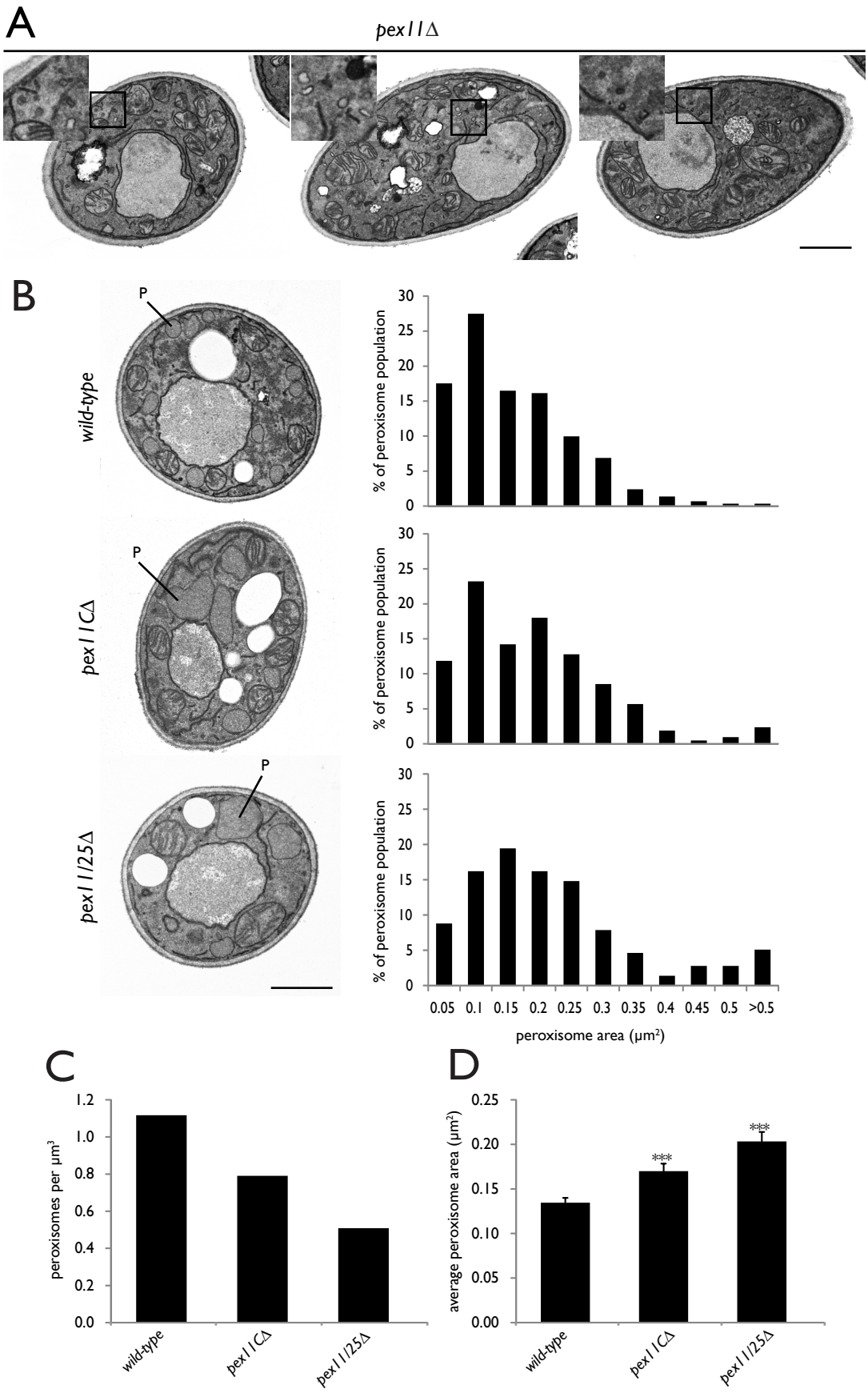

A
wild-type
pex11Δ
pex11CΔ
pex11/25Δ
Pot1p-GFP
0
2
4
6
8
oleate incubation (h)
B
YEPD
YPBO
wild-type
pex11Δ
wild-type
pex11Δ
Pot1p-mRFP
C
YEPD
YPBO
wild-type
pex11CΔ
pex11/25Δ
wild-type
pex11CΔ
pex11/25Δ
Pot1p-GFP
Number of peroxisomes/μm²
0.20
0.16
0.12
0.08
0.04
0
N.S.
N.S.
wild-type
pex11CΔ
pex11/25Δ
Number of peroxisomes/μm²
0.20
0.16
0.12
0.08
0.04
0
**
**
wild-type
pex11CΔ
pex11/25Δ

Figure 3-9. *pex11*Δ cells contain numerous small vesicles, while *pex11C*Δ and *pex11/25*Δ cells contain enlarged peroxisomes. (A) Ultrastructure of *pex11*Δ cells. Wild-type peroxisomal profiles (panel B) are absent in *pex11*Δ cells. Insets present enlarged views of boxed areas showing increased numbers of small vesicles in *pex11*Δ cells. Bar, 1 μm. (B) Ultrastructure and morphometric analysis of wild-type *E34*, *pex11C*Δ and *pex11/25*Δ cells. Cells were grown in YEPD medium, transferred to YPBO medium for 8 h and processed for EM. Representative images of each strain are shown. P, peroxisome. Bar, 1 μm. The areas of individual peroxisomes of 100 randomly selected cells were morphometrically analyzed using Olympus iTEM software. Peroxisomes were then separated into size categories. A histogram was generated for each strain depicting the percentage of total peroxisomes occupied by the peroxisomes of each category. The numbers along the *x*-axis are the maximum sizes of peroxisomes (in square micrometers) in each category, except for the last number, which represents the minimum size of peroxisomes (in square micrometers) in the last category. (C) Estimated number of peroxisomes per cubic micrometer of cells of the wild-type strain *E34* and of the mutant strains *pex11C*Δ and *pex11/25*Δ. (D) Average peroxisome size (in square micrometers) for cells of the wild-type strain *E34* and of the mutant strains *pex11C*Δ and *pex11/25*Δ. Error bars give the SEM. Significance of the comparison of the average peroxisome area of each deletion strain with that of the wild-type strain was determined by *t*-test . ***, $p < 0.001$.


A
pex11Δ
B
wild-type
pex11CΔ
pex11/25Δ
P
% of peroxisome population
peroxisome area (μm²)
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5 >0.5
C
peroxisomes per μm³
D
average peroxisome area (μm²)
***

show that *pex11*Δ cells fail to assemble morphologically identifiable, functional peroxisomes. Morphometric analysis of electron micrographs revealed that *pex11C*Δ and *pex11/25*Δ cells contain fewer and larger peroxisomes compared to wild-type cells (Figure 3-9, B-D).

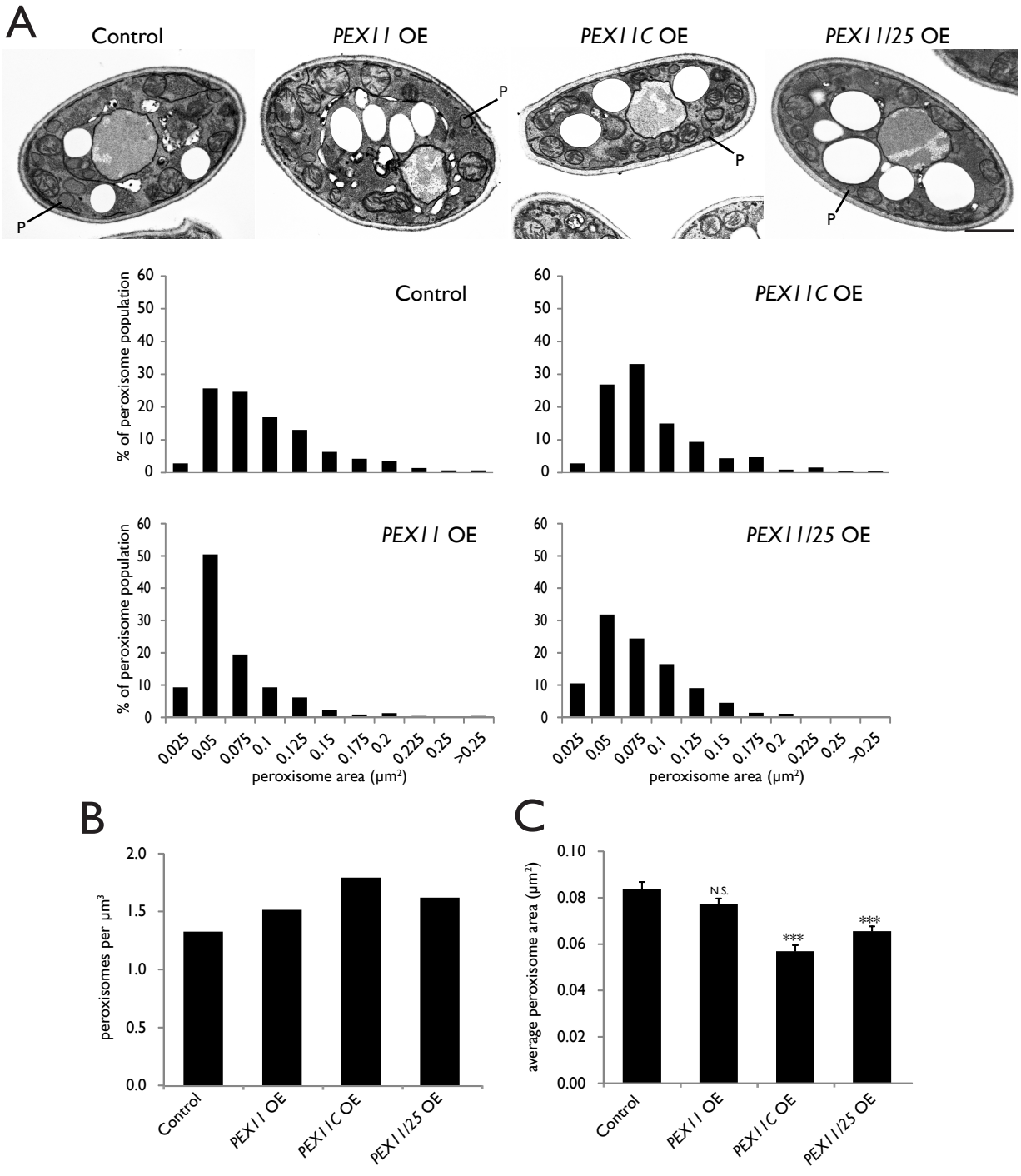
3.8 Overexpression of *PEX11C* and *PEX11/25* results in cells with increased numbers of smaller peroxisomes

We next examined the effect of overexpressing *PEX11*, *PEX11C* or *PEX11/25* on peroxisome size and number. Morphometric analysis of electron micrographs showed that cells overexpressing *PEX11C* or *PEX11/25*, but not *PEX11*, contained increased numbers of smaller peroxisomes (Figure 3-10, A-C). Taken together, the data from our gene deletion and overexpression studies demonstrated a classical Pex11-protein-family-member role in maintenance of peroxisome size and number for Pex11Cp and Pex11/25p, but also unexpectedly showed that Pex11p in *Y. lipolytica* does not function in controlling peroxisome size and number.

3.9 *pex11*Δ cells exhibit abnormal localization of peroxisomal matrix and membrane proteins

Unlike other eukaryotes investigated so far, *Y. lipolytica* cells deleted for *PEX11* do not contain morphologically identifiable peroxisomes, and cells overexpressing *PEX11* do not contain an increased number of smaller peroxisomes. These unexpected findings prompted us to examine the localization of peroxisomal matrix and membrane proteins in this strain. Cells of the wild-type strain *E34* and of the deletion strains *pex11*Δ, *pex11C*Δ and *pex11/25*Δ were subjected to differential centrifugation to yield a 20KgS fraction and a 20KgP fraction enriched for 'mature' peroxisomes (Titorenko and Rachubinski, 1998). The 20KgS fraction was subjected to ultracentrifugation at 200,000 × *g* to yield a pellet (200KgP) fraction enriched for small vesicles, including small peroxisomal vesicles (Titorenko and Rachubinski, 1998) and a supernatant (200KgS) fraction enriched for cytosol. Immunoblotting was performed with antibodies to matrix proteins with different PTSs: the PTS1-containing protein isocitrate lyase (ICL), a 62-kDa protein recognized by

Figure 3-10. Overexpression of *PEX11C* or *PEX11/25*, but not *PEX11*, results in cells with increased numbers of smaller peroxisomes. (A) Ultrastructure and morphometric analysis of cells overexpressing *PEX11*, *PEX11C* or *PEX11/25*, and a control strain containing the parental vector pTC3. Cells were grown in glucose-containing medium, transferred to oleic acid-containing medium and incubated for 20 h and processed for electron microscopy. One representative image is shown for each strain. P, peroxisome. Bar, 1 μm. Morphometric analysis was performed as described in the legend to Figure 3-9B. (B) Estimated number of peroxisomes per cubic micrometer of cells of the wild-type strain *E34* and of the mutant strains *pex11CΔ* and *pex11/25Δ*. (C) Average peroxisome size (in square micrometers) for cells of the wild-type strain *E34* and of the mutant strains *pex11CΔ* and *pex11/25Δ*. Error bars give the SEM. Significance of the comparison of the average peroxisome area of each deletion strain with that of the wild-type strain was determined by *t*-test. N.S., not significant. ***, $p < 0.001$.

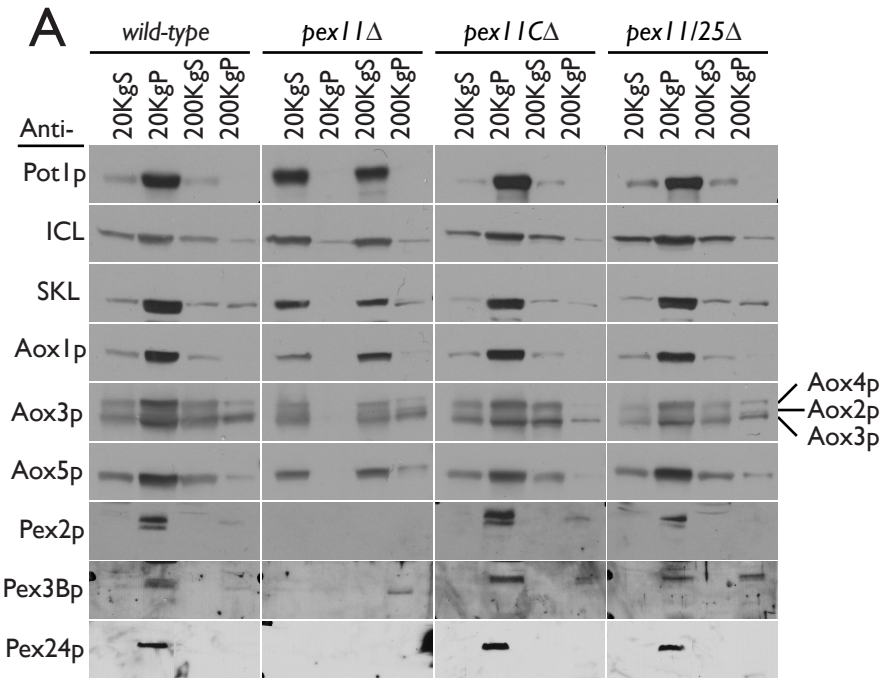

A
Control
PEX11 OE
PEX11C OE
PEX11/25 OE
P
% of peroxisome population
peroxisome area (μm²)
0.025
0.05
0.075
0.1
0.125
0.15
0.175
0.2
0.225
0.25
>0.25
B
peroxisomes per μm³
C
average peroxisome area (μm²)
N.S.
***

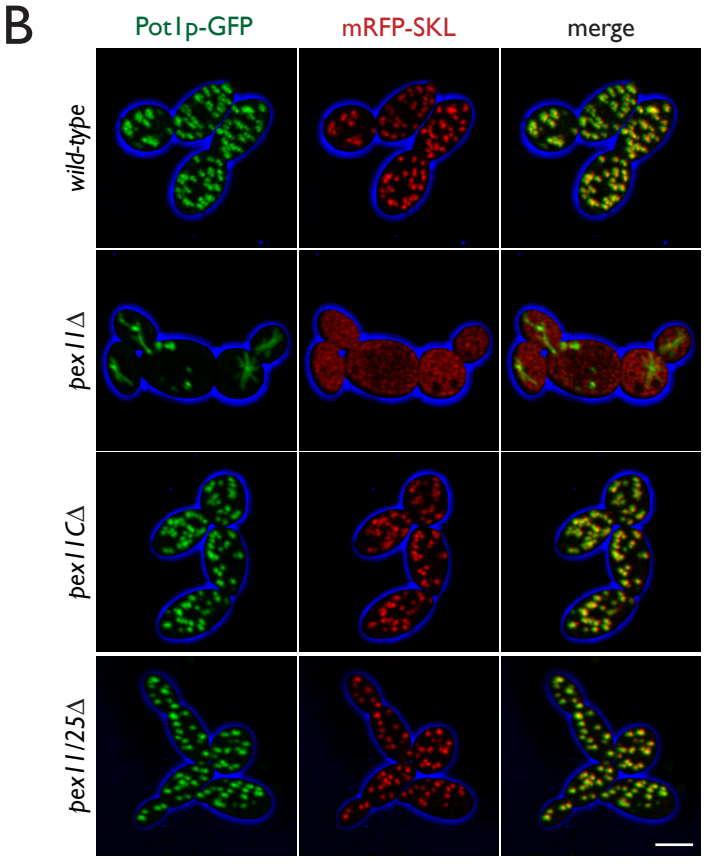
antibodies against the tripeptide Ser-Lys-Leu (SKL) PTS1 consensus sequence, the PTS2-containing protein Pot1p, and five acyl-CoA oxidase subunits (Aox1-5) that contain neither a conventional PTS1 or PTS2 sequence. These matrix proteins localized preferentially to the 20KgP fraction from wild-type, *pex11CΔ* and *pex11/25Δ* cells (Figure 3-11A). In contrast, these matrix proteins were detected predominately in the 20KgS and 200KgS fractions from *pex11Δ* cells. Confocal microscopy showed that the fluorescent PTS1-containing peroxisomal marker protein mRFP-SKL colocalized with Pot1p-GFP in punctate structures characteristic of peroxisomes in wild-type, *pex11CΔ* and *pex11/25Δ* cells (Figure 3-11B) However, mRFP-SKL in *pex11Δ* cells displayed a diffuse pattern of localization characteristic of the cytosol. These findings indicate that *pex11Δ* cells are defective in peroxisomal matrix protein import, regardless of the PTS sequence they employ.

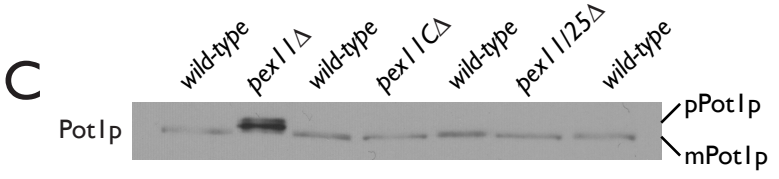
Like peroxisomal matrix proteins, the peroxisomal membrane proteins Pex2p, Pex3Bp and Pex24p enriched in the 20KgP fraction from wild-type, *pex11CΔ* and *pex11/25Δ* cells (Figure 3-11A). Suprisingly, Pex2p and Pex24p were not detected in any subcellular fraction from *pex11Δ* cells, while Pex3Bp was found to be enriched in the 200KgP fraction containing small vesicles. These data are further supportive of *pex11Δ* cells being defective in peroxisome assembly and of an absence of 'mature' peroxisomes in *pex11Δ* cells.

In wild-type cells, the enzyme Pot1p is imported into the peroxisomal matrix as a 45-kDa precursor form (pPot1p), where it undergoes proteolytic processing to its mature 43-kDa form (mPot1p) (Titorenko and Rachubinski, 1998). Only mPot1p was detected by immunoblot of cell lysates of the wild-type, *pex11Δ*, *pex11CΔ* and *pex11/25Δ* strains; however, only pPot1p was detected in the *pex11Δ* cell lysate (Figure 3-11C), further supporting that *pex11Δ* cells are defective in peroxisomal matrix protein import.

Figure 3-11. Deletion of the *PEX11* gene results in abnormal localization of peroxisomal matrix and membrane proteins. (A) Cells of wild-type strain *E34* and the deletion strains *pex11*Δ, *pex11C*Δ and *pex11/25*Δ were subjected to subcellular fractionation to yield a 20KgS fraction and a 20KgP fraction enriched for peroxisomes. The 20KgS fraction was subjected to additional ultracentrifugation at 200,000 × *g* to yield a 200KgS fraction enriched for cytosol and a 200KgP fraction enriched for small vesicles. Equivalent volumes of all fractions were analyzed by immunoblotting with antibodies to the indicated proteins. Anti-Aox3p antibodies recognize three subunits of the enzyme fatty acyl-CoA oxidase, *i.e.* Aox2p, Aox3p and Aox4p. (B) PTS1- and PTS2-containing proteins are mislocalized in *pex11*Δ cells. Cells of the wild-type strain *E34* and of the deletion strains *pex11*Δ, *pex11C*Δ and *pex11/25*Δ, expressing genomically integrated *POT1-GFP* and a mRFP-SKL-encoding plasmid, were grown in YEPD medium supplemented with hygromycin B and transferred to YPBO medium supplemented with hygromycin B for 8 h. Representative confocal microscopy images are shown. Bar, 5 μm. (C) Deletion of the *PEX11* gene prevents proteolytic processing of Pot1p to its mature form. Whole cell lysates of the wild-type strain *E34* and the deletion strains *pex11*Δ, *pex11C*Δ and *pex11/25*Δ were subjected to immunoblotting with antibodies to Pot1p. mPot1p, mature form of Pot1p; pPot1p, precursor form of Pot1p.


A
wild-type
pex11Δ
pex11CΔ
pex11/25Δ
20KgS
20KgP
200KgS
200KgP
Anti-
Pot1p
ICL
SKL
Aox1p
Aox3p
Aox5p
Pex2p
Pex3Bp
Pex24p
Aox4p
Aox2p
Aox3p


B
Pot1p-GFP
mRFP-SKL
merge
wild-type
pex11Δ
pex11CΔ
pex11/25Δ


C
wild-type
pex11Δ
wild-type
pex11CΔ
wild-type
pex11/25Δ
wild-type
Pot1p
pPot1p
mPot1p

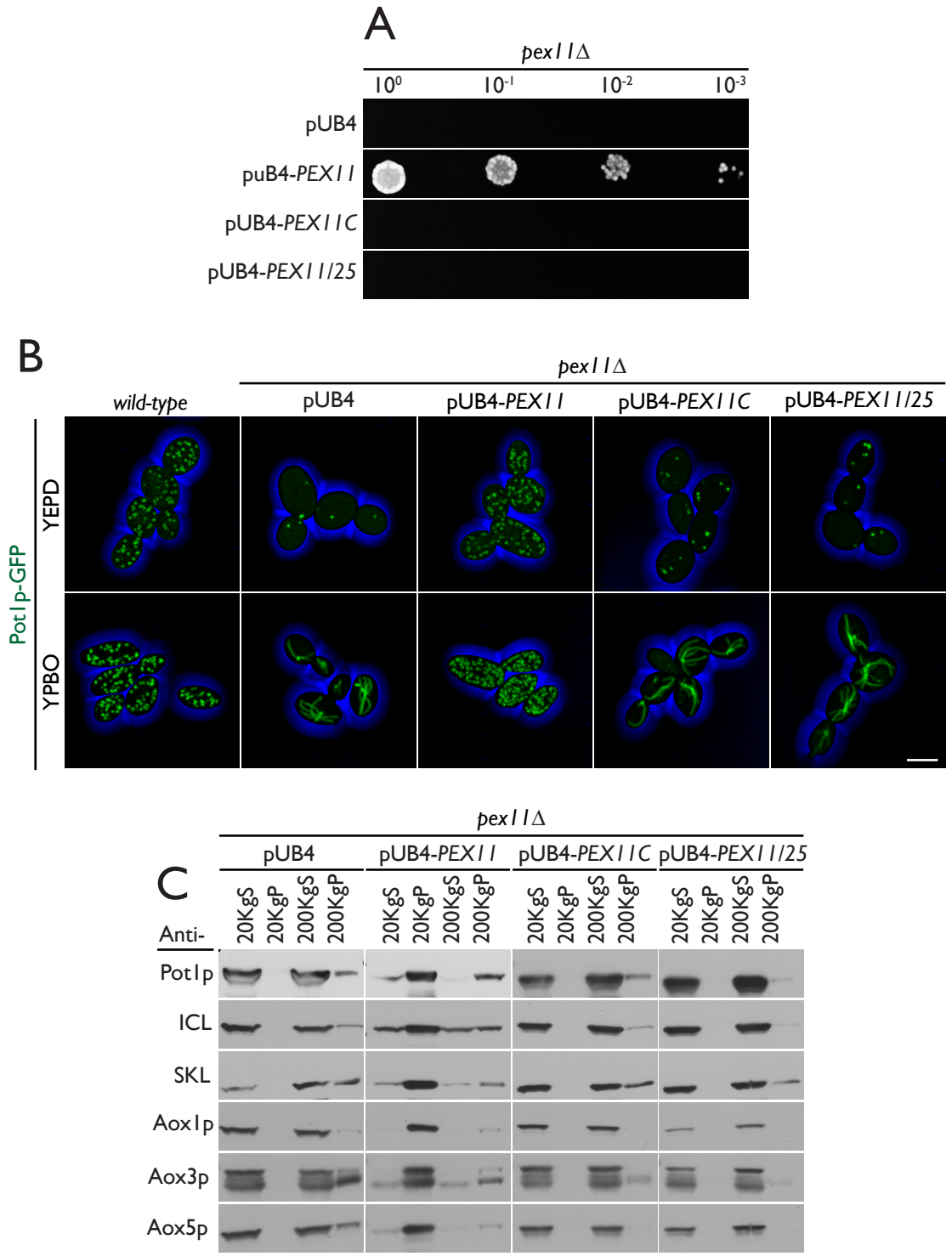
3.10 Expression of *PEX11*, but not *PEX11C* or *PEX11/25*, complements the peroxisome assembly defects of *pex11*Δ cells

We next tested whether Pex11p, Pex11Cp or Pex11/25p could complement the aberrant peroxisome assembly observed for *pex11*Δ cells, *i.e.* the inability to grow on fatty acid-containing medium (Figure 3-7A), a failure to import Pot1p-GFP into characteristic punctate peroxisomal structures (Figure 3-8A), and compromised import of peroxisomal matrix proteins and targeting of peroxisomal membrane proteins as determined by subcellular fractionation (Figure 3-11A). Plasmid expression of *PEX11*, but not of *PEX11C* or *PEX11/25*, restored growth of the *pex11*Δ strain on oleic acid-containing medium (Figure 3-12A). Likewise, only plasmid expression of *PEX11* restored the presence of characteristic wild-type fluorescent punctate peroxisomes in *pex11*Δ cells incubated in either glucose-containing YEPD medium or oleic acid-containing YPBO medium, while *pex11*Δ cells expressing plasmid-borne *PEX11C* or *PEX11/25* showed a small number of fluorescent puncta in YEPD medium and fluorescent elongated tendrillar structures in YPBO medium, like *pex11*Δ cells containing the parental plasmid (pUB4) alone (Figure 3-12B). Only plasmid-expressed *PEX11* restored normal import of peroxisomal matrix proteins as determined by subcellular fractionation (Figure 3-12C). These findings demonstrate that expression of *PEX11*, but not *PEX11C* or *PEX11/25*, complements the aberrant peroxisome assembly in *pex11*Δ cells and provide additional evidence that Pex11p has functions in peroxisome biogenesis that are different from those of both Pex11Cp and Pex11/25p.

3.11 Pex11p, Pex11Cp and Pex11/25p are mislocalized in *pex3*Δ, *pex16*Δ and *pex19*Δ cells

Our results show that *Y. lipolytica pex11*Δ cells do not assemble functional peroxisomes, do not contain morphologically identifiable peroxisomes, and mislocalize peroxisomal matrix and membrane proteins, all of which are indicative of an unprecedented role for the divisional Pex11 protein in peroxisome assembly *per se*. Peroxisomes are linked to the cellular endomembrane trafficking system via the ER. Whereas the amount *de novo* formation of peroxisomes from the ER

Figure 3-12. Expression of *PEX11*, but not *PEX11C* or *PEX11/25*, complements the peroxisome assembly defects of *pex11*Δ cells. (A) *PEX11* expression restores the ability to grow on oleic acid-containing medium to *pex11*Δ cells. *pex11*Δ cells were transformed with pUB4 or pUB4 expressing *PEX11*, *PEX11C* or *PEX11/25*, grown in YEPD medium supplemented with hygromycin B, spotted at dilutions of 10^{0} to 10^{-3} on YPBO agar plates supplemented with hygromycin B, and grown for 2 d at 28°C. (B) Normal peroxisomal morphology is restored by *PEX11* expression in *pex11*Δ cells. *pex11*Δ cells expressing genomically integrated *POT1-GFP* were transformed with pUB4 or pUB4 expressing *PEX11*, *PEX11C* or *PEX11/25*. Wild-type cells grown in YEPD medium or *pex11*Δ cells grown in YEPD medium supplemented with hygromycin B were transferred to YPBO medium (wild-type) or YPBO medium supplemented with hygromycin B (*pex11*Δ cells) for 8 h. Representative confocal microscopy images are shown. Bar, 5 μm. (C) *pex11*Δ cells containing pUB4 or pUB4 expressing *PEX11*, *PEX11C*, or *PEX11/25* were subjected to subcellular fractionation to yield 20KgS and 20KgP fractions. The 20KgS fraction was subjected to further ultracentrifugation at 200,000 × *g* to yield 200KgS and 200KgP fractions. Equivalent volumes of all fractions were analyzed by immunoblotting with antibodies to the indicated proteins.


A
pex11Δ
10^0
10^-1
10^-2
10^-3
pUB4
puB4-PEX11
pUB4-PEX11C
pUB4-PEX11/25
B
pex11Δ
wild-type
pUB4
pUB4-PEX11
pUB4-PEX11C
pUB4-PEX11/25
Pot1p-GFP
YEPD
YPBO
C
pex11Δ
pUB4
pUB4-PEX11
pUB4-PEX11C
pUB4-PEX11/25
20KgS
20KgP
200KgS
200KgP
Anti-
Pot1p
ICL
SKL
Aox1p
Aox3p
Aox5p

contributes to overall peroxisome biogenesis remains an open question, peroxisomes do rely on the ER at least to supply them with proteins and lipids to sustain their membrane components through continuous rounds of peroxisomal growth and division (reviewed in Tabak et al., 2013). Peroxisomal membrane proteins sample the ER *en route* to peroxisomes, and peroxisome growth is facilitated by proteins and lipids coming from the ER via both vesicular and non-vesicular pathways (reviewed in Smith and Aitchison, 2013). The functions of the peroxins Pex3p and Pex19p in these processes are well documented in different eukaryotes (reviewed in Smith and Aitchison, 2013; Tabak et al., 2013). In *Y. lipolytica*, *pex3*Δ cells lack any evidence of peroxisomal structures (Bascom et al., 2003). *pex16*Δ cells do not contain morphologically identifiable peroxisomes, but import a subset of peroxisomal proteins (Eitzen et al., 1997). *pex19*Δ cells, while they do contain morphologically identifiable peroxisomes, mislocalize peroxisomal matrix proteins to the cytosol (Lambkin and Rachubinski, 2001). We assessed the localization of Pex11p-mC, Pex11Cp-mC and Pex11/25p-mC in *pex3*Δ, *pex16*Δ and *pex19*Δ cells by confocal microscopy. Distinct "rosette" structures decorated by Pex11p-mC were observed in *pex3*Δ and *pex19*Δ cells grown in either YEPD medium or YPBO medium (Figure 3-13A, boxed regions and insets). Pex11-mC was often localized to a single prominent structure in *pex16*Δ cells grown in YPBO medium (Pex11-mC was not expressed in *pex16*Δ cells grown in YEPD medium) (Figure 3-13A). Pex11Cp-mC in *pex3*Δ, *pex16*Δ and *pex19*Δ cells exhibited a distinctive cortical and perinuclear localization reminiscent of the ER (Figure 3-13B) and similar in appearance to the pattern exhibited by the fluorescent, ER-localized marker protein, GFP-HDEL, in wild-type, *pex3*Δ, *pex11*Δ, *pex16*Δ and *pex19*Δ cells (Figure 3-14). In contrast, Pex11/25p-mC exhibited a generalized cortical localization in *pex3*Δ and *pex19*Δ cells (Figure 3-13C). Deletion strains expressing Pex11Cp-mC and Pex11/25p-mC did not display any fluorescent signal when grown in glucose-containing medium, and Pex11/25p-mC was not expressed in *pex16*Δ cells. Our observations suggest that Pex11p, Pex11Cp and Pex11/25p are localized to membrane structures, and in the case of Pex11Cp specifically to membranes of the

Figure 3-13. Pex11p, Pex11Cp and Pex11/25p are mislocalized in *pex3*Δ, *pex16*Δ and *pex19*Δ cells. Cells of the wild-type strain *E122* and of the deletion strains *pex3*Δ, *pex16*Δ and *pex19*Δ expressing Pex11p-mC, Pex11Cp-mC, or Pex11/25p-mC from plasmid were grown in YEPD medium and then transferred to YPBO medium for 8 h. Plasmid was maintained in cells by addition of drug (hygromycin B or nourseothricin) to media. Representative confocal microscopy images (maximum intensity projections (MIP) and middle slices) are shown for each strain. Boxes show enlarged view of rosette structures decorated by Pex11p-mC present in *pex3*Δ and *pex19*Δ cells. Intensity settings for image acquisition were optimized for each strain and for each fluorescent chimeric protein. Bar, 5 μm.

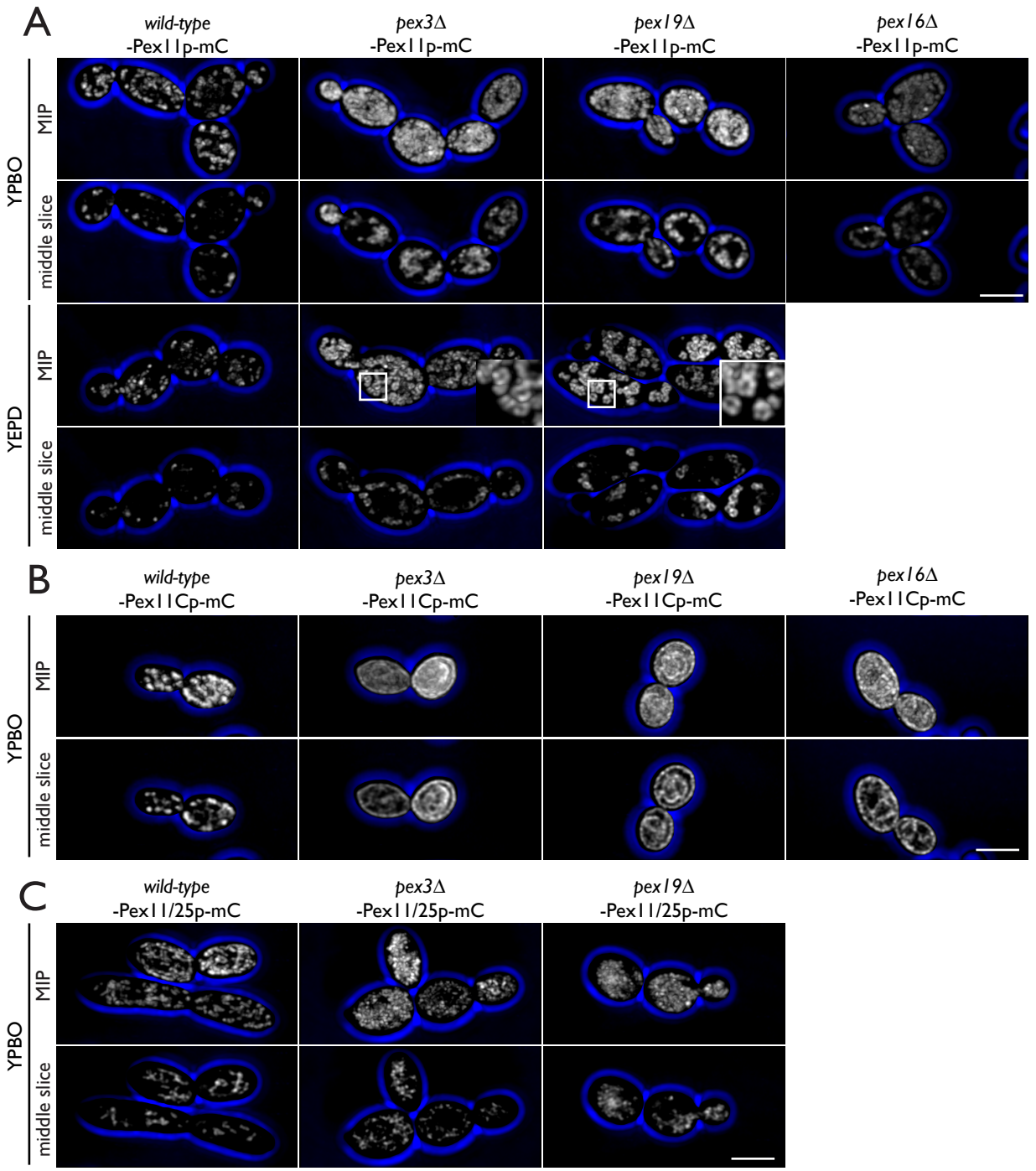
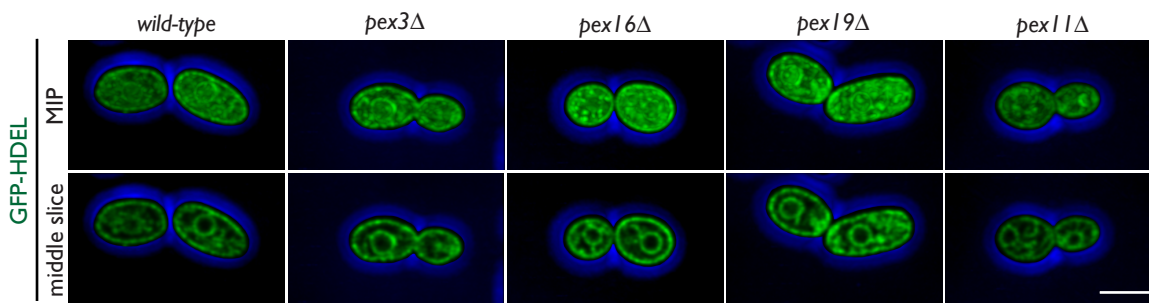

A
wild-type
-Pex11p-mC
pex3Δ
-Pex11p-mC
pex19Δ
-Pex11p-mC
pex16Δ
-Pex11p-mC
YPBO
MIP
middle slice
YEPD
MIP
middle slice
B
wild-type
-Pex11Cp-mC
pex3Δ
-Pex11Cp-mC
pex19Δ
-Pex11Cp-mC
pex16Δ
-Pex11Cp-mC
YPBO
MIP
middle slice
C
wild-type
-Pex11/25p-mC
pex3Δ
-Pex11/25p-mC
pex19Δ
-Pex11/25p-mC
YPBO
MIP
middle slice

Figure 3-14. *pex3*Δ, *pex16*Δ, *pex19*Δ and *pex11*Δ cells display normal ER morphology. Cells of the wild-type strain *E122* and of the deletion strains *pex3*Δ, *pex16*Δ, *pex19*Δ *and pex11*Δ containing plasmid expressing the fluorescent ER marker GFP-HDEL were grown in glucose-containing medium and imaged by confocal microscopy. Representative images (maximum intensity projections (MIP) and middle slices) are shown for each strain. Bar, 5 μm.


wild-type
pex3Δ
pex16Δ
pex19Δ
pex11Δ
GFP-HDEL
MIP
middle slice

secretory system, in *pex3*Δ, *pex16*Δ and *pex19*Δ cells that do not assemble functional peroxisomes.

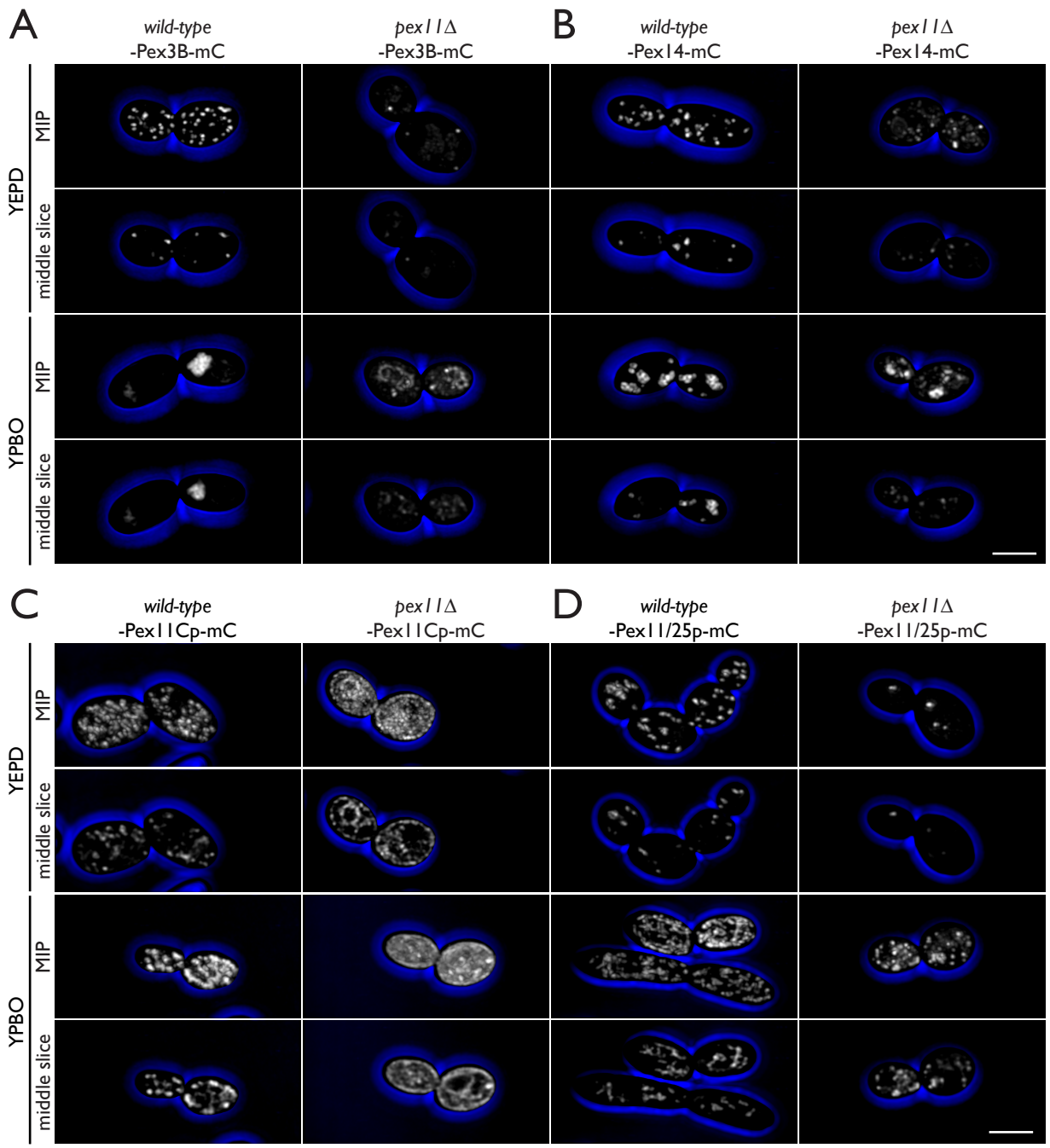
3.12 Peroxisomal membrane proteins are mislocalized in *pex11*Δ cells

We next assessed the localization of peroxisomal membrane proteins Pex3Bp-mC, Pex14p-mC, Pex11Cp-mC, and Pex11/25p-mC in *pex11*Δ cells. Although localized to punctate structures in wild-type cells grown in glucose-containing medium, Pex3Bp-mC was globular in wild-type cells grown in oleic acid-containing medium (Figure 3-15A), as previously noted (Chang et al., 2009). Pex3Bp-mC was often localized to several discrete puncta that could be cortically located in *pex11*Δ cells grown in glucose-containing medium (Figure 3-15A). Pex14p-mC was also localized to globular structures in wild-type cells grown in oleic acid-containing medium (Figure 3-15B). Pex14p-mC was found in foci, typically at the cortex, in *pex11*Δ cells grown in glucose-containing medium, and exhibited a generalized cortical localization in *pex11*Δ cells grown in oleic acid-containing medium (Figure 3-15B). Pex11Cp-mC in *pex11*Δ cells exhibited distinctive cortical and perinuclear localizations reminiscent of the ER (Figure 3-15C). Pex11/25p-mC in *pex11*Δ cells produced several prominent puncta in glucose-grown cells, and a non-specific cortical and perinuclear localization in oleic acid-grown cells (Figure 3-15D). Our observations suggest that Pex3Bp, Pex14p, Pex11Cp and Pex11/25p are localized to membrane structures, and in the case of Pex11Cp specifically to membranes of the secretory system, in *pex11*Δ cells that do not assemble functional peroxisomes.

3.13 Discussion

Limited functional work has been completed on the novel *in silico* predicted Pex11p-related proteins in Fungi (Kiel et al., 2006). The Pex11 protein family in the yeast *Penicillium chrysogenum*, consisting of Pex11p, Pex11Bp and Pex11Cp, was recently characterized (Opaliński et al., 2012). Pex11p and Pex11Cp were found to be peroxisomal, but Pex11Bp was localized to the ER. Deletion of *PEX11B* or *PEX11C* had no significant effect on peroxisome number; however, overexpression of *PEX11B* resulted in clusters of smaller peroxisomes. The authors concluded that

Figure 3-15. Membrane proteins are mislocalized in *pex11*Δ cells. Cells of the wild-type strain *E122* and of *pex11*Δ cells expressing Pex3Bp-mC, Pex14Cp-mC, Pex11Cp-mC, or Pex11/25p-mC from plasmid were grown in YEPD medium and then transferred to YPBO medium for 8 h. Plasmid was maintained in cells by addition of drug (hygromycin B or nourseothricin) to media. Representative confocal microscopy images (maximum intensity projections (MIP) and middle slices) are shown for each strain. Intensity settings for image acquisition were optimized for each strain and for each fluorescent chimeric protein. Bar, 5 μm.

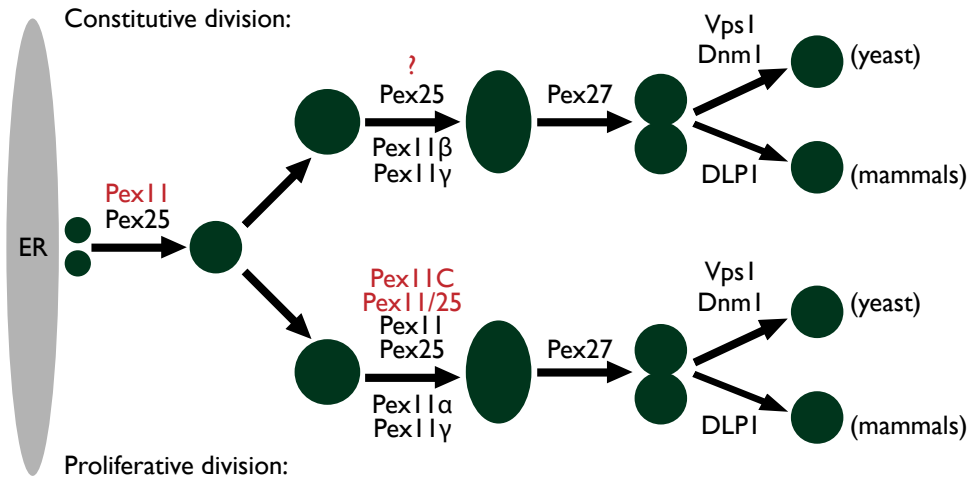

A
wild-type
-Pex3B-mC
pex11Δ
-Pex3B-mC
B
wild-type
-Pex14-mC
pex11Δ
-Pex14-mC
YEPD
YPBO
MIP
middle slice
C
wild-type
-Pex11Cp-mC
pex11Δ
-Pex11Cp-mC
D
wild-type
-Pex11/25p-mC
pex11Δ
-Pex11/25p-mC

Pex11Bp may play some role in peroxisome biogenesis, although not in *de novo* peroxisome assembly. The Pex11 protein family in the yeast *H. polymorpha* was also recently investigated, and Pex11p, Pex11Cp and Pex25p were shown to be peroxisomal (Saraya et al., 2011). Intriguingly, Pex25p in *H. polymorpha*, like Pex25p in *S. cerevisiae*, was shown to be required for *de novo* peroxisome biogenesis, as the presence of Pex25p, but not Pex11p, was needed to reintroduce peroxisomes into a peroxisome-deficient strain.

We have now demonstrated that Pex11Cp and Pex11/25p function in the regulation of peroxisome size and number. In this way, Pex11Cp and Pex11/25p function similarly to Pex11 family proteins in other eukaryotes. Quantification showed that *Y. lipolytica pex11CΔ* or *pex11/25Δ* cells grown in glucose-containing medium did not contain significantly different numbers of peroxisomes from wild-type cells. Conversely, *pex11CΔ* and *pex11/25Δ* strains grown in oleic acid-containing medium contained fewer peroxisomes than wild-type cells. These observations suggest that Pex11Cp and Pex11/25p in *Y. lipolytica* may act to regulate peroxisome size and number in a proliferating peroxisome population rather than a constitutively dividing one. Hence, we have expanded on a previous model of Pex11 family proteins in peroxisome division (Figure 3-16).

Pex11 family proteins appear to be associated with the peroxisomal membrane, but the specific membrane topology of these proteins varies with the organism (reviewed in Schrader et al., 2012). The members of the *S. cerevisiae* Pex11 protein family, consisting of Pex11p, Pex25p and Pex27p, appear to be peripherally associated with the peroxisomal membrane, although some controversy still exists with regards to the strength of the association between Pex11p and the membrane. Mammalian Pex11 proteins are integral membrane proteins with one or two transmembrane spanning helices. We show that *Y. lipolytica* Pex11p is an integral membrane protein of peroxisomes (Figure 3-7E), with two membrane-spanning domains predicted. This means that the membrane topology of *Y. lipolytica* Pex11p more closely resembles its human homologues than it does *S. cerevisiae* Pex11p. Organelle extraction also demonstrated that Pex11Cp and Pex11/25p are integral to the peroxisomal membrane (Figure 3-7E).

Figure 3-16. A revised model for the roles of Pex11 family proteins in peroxisome division. This model of peroxisome division is the same as shown in Figure 1-1, but incorporates the *Y. lipolytica* Pex11 family proteins (red). Under conditions of proliferative peroxisome division in *Y. lipolytica*, Pex11Cp and Pex11/25p act upstream of or at the stage of peroxisome elongation; the specific protein(s) needed for this process in constitutive peroxisome division is unknown. Pex11p in *Y. lipolytica* is required for *de novo* peroxisome formation from the ER. Adapted from Tower, 2012.


Constitutive division:
ER
Pex11
Pex25
?
Pex25
Pex11β
Pex11γ
Pex27
Vps1
Dnm1
(yeast)
DLP1
(mammals)
Pex11C
Pex11/25
Pex11
Pex25
Pex11α
Pex11γ
Pex27
Vps1
Dnm1
(yeast)
DLP1
(mammals)
Proliferative division:

Previous studies had concluded that Pex11p had a eukaryotic phylogenetic origin (Schlüter et al., 2006; Gabaldón et al., 2006). We have now demonstrated that Pex11p is a conserved, ancestral protein across diverse eukaryotes. One secondary structural feature of Pex11p is its N-terminal amphipathic helix, which has been proposed to induce membrane curvature like other BAR (Bin/Amphiphysin/Rvs-homology)-like domain-containing proteins (Opaliński et al., 2011). One way to induce membrane curvature is to insert an amphipathic helix into one leaflet of the lipid bilayer. This produces membrane asymmetry and bending. Membrane curvature is then stabilized by BAR domains interacting (Campelo et al., 2008). BAR domains have a core comprised of a bundle of six α-helices (Masuda and Mochizuki, 2010). N-BAR domain proteins contain a N-terminal amphipathic helix followed by a BAR domain (Cui et al., 2013). ESCRT comlexes and the retromer coat both have α-helical bundle structures resembling BAR-like domains (Field et al., 2011). The combinations of the domains that actually define membrane-deforming structures, *e.g.* a β-propeller followed by an α-solenoid to form protocoatomer, are thought to be primarily eukaryotic innovations (Field et al., 2011). Membrane-bending mechanisms are thought to have been present in at least the first eukaryotic common ancestor (FECA) (Field et al., 2011). However, the architecture of membrane-deforming complexes (β-propeller, α-solenoid and α-helix domains) has been found in prokaryotes, including β-propeller and α-solenoid proteins in the Planctomycetes-Verrucomicrobia-Chlamydiae (PVC) bacterial superphylum (Santarella-Mellwig et al., 2010). Whether there are other specific β-propeller/α-solenoid or α-helix structures in other peroxins is an open question.

*pex3*Δ, *pex16*Δ and *pex19*Δ cells of *Y. lipolytica* are defective in peroxisome biogenesis and lack wild-type peroxisomes (Eitzen et al., 1997; Lambkin and Rachubinski, 2001; Bascom et al., 2003; Chang et al., 2009). Here, we showed that Pex11p is also implicated in peroxisome biogenesis in *Y. lipolytica* and demonstrated a putative localization of Pex11 family proteins to compartments of the secretory system in *pex3*Δ, *pex16*Δ and *pex19*Δ deletion strains (Figures 3-13 and 3-14). When Pex11p-mC was expressed in *pex3*Δ and *pex19*Δ cells, distinctive fluorescent rosette structures

were observed (Figure 3-13). We suggest that Pex11p could be modifying the ER membrane in these strains; however, we did not observe evidence of modified ER in electron micrographs (Figure 3-9A) or by fluorescent microscopy of *pex11*Δ cells expressing the GFP-HDEL marker (Figure 3-14). For instance, overexpression of Pex15p in *S. cerevisiae* cells results in extensive ER proliferation that can be visualized by electron microscopy (Elgersma et al., 1997). This is thought to be the result of Pex15p accumulation in the ER, resulting in membrane proliferation (Elgersma et al., 1997). However, we did not observe this drastic ER proliferation in electron micrographs of *pex11*Δ cells. Similarly, there were no apparent prominent foci in *pex11*Δ cells expressing the GFP-HDEL marker, which may have been indicative of ER clumping. When Pex11Cp-mC was expressed in *pex16*Δ cells, a single prominent structure was often observed (Figure 3-13). This phenotype is reminiscent of the phenotype observed when the first N-terminal 46 amino acids of Pex3p are expressed in wild-type cells (Tam et al., 2005). We propose that these may be pre-peroxisomal structures. The specific nature of these structures awaits further investigation.

We have also demonstrated a putative localization of peroxisomal membrane proteins to compartments of the secretory system in *pex11*Δ cells that do not assemble functional peroxisomes (Figure 3-15). Again, the discrete cortical foci observed in *pex11*Δ cells expressing Pex3Bp, Pex14p or Pex11/25p could possibly be pre-peroxisomal structures, or membrane proteins in specific locations at the ER. Once again, the specific nature of these structures awaits further investigation.

Until recently, mutation of the *PEX11* gene had not been associated with a PBD, but the first patient with pathology resulting from a homozygous nonsense mutation in the *PEX11β* gene has been reported (Ebberink et al., 2012). Aberrant peroxisomal morphology (enlarged and elongated peroxisomes) was observed in the patient's skin fibroblasts grown at 37°C. Surprisingly, no obvious defects in the patient's peroxisomal metabolism were identified (as determined by concentrations of VLCFA, phytanic and pristanic acid, and bile acid intermediates in plasma, and plasmalogen concentration in erythrocytes). Under conditions that exacerbated this abnormal

phenotype, *e.g.* growth at 40°C, peroxisomal structures were still observed in the patient's fibroblasts, but a matrix protein was mislocalized to the cytosol. Temperature sensitivity refers to fluctuations in cell line phenotype when cultured above or below body temperature (Steinberg et al., 2006). Culturing cells at 40°C can be used to identify some peroxisomal assembly defects in cell lines displaying peroxisomal mosaicism (discussed below) (Gootjes et al., 2004). Overexpression of *PEX11β* restored matrix protein import, and overexpression of *PEX11γ* also partially complemented this deficiency. In both control and patient fibroblasts, decreased levels of *PEX11γ* mRNA were observed at 40°C, and therefore the absence of import-competent peroxisomes under this condition is probably the result of a combined *PEX11β* and *PEX11γ* deficiency. The authors proposed a model in which *PEX11β* is involved in peroxisome elongation and constriction, while *PEX11γ* plays a role in the import of matrix proteins into pre-peroxisomal vesicles. However, one caveat of this result may be that cultured skin fibroblasts may not necessary be indicative of the phenotypes observed in all tissues. The term peroxisomal mosaicism describes variation on peroxisome function within and between tissues (Steinberg et al., 2006). The first type of peroxisomal mosaicism referes a disparity between the phenotypes observed in cultured fibroblasts and the biochemical profile in a patient's bodily fluids (Pineda et al., 1999). The second type refers to variable phenotypes in cells with the same genotype (for examples, see Mandel et al., 1994; Pineda et al., 1999; Shimozawa et al., 1999).

Although PBDs are caused by mutations in *PEX* genes, a single patient with a mutation in the *DLP1* gene has also been reported (Waterham et al., 2007). This patient displayed clinical symptoms characteristic of PBDs and died after approximately one month (Waterham et al., 2007). The patient's skin fibroblasts contained fewer peroxisomes that varied in size as well as aberrant mitochondrial morphology (Waterham et al., 2007). Thus, this is another example of a mutation in peroxisomal divisional machinery resulting in human pathology.

Pex11p is also associated with pathology in other models. *PEX11β*-/- knockout mice are neonatally lethal and display severe PBD symptoms such as intrauterine growth retardation,

hypotonia, neuronal migrations defects, and developmental delays in liver and kidney tissues (Li et al., 2002). However, peroxisome function is normal and no defect in matrix protein import is observed (Li et al., 2002). Similiarly, *PEX11β*+/- mice display increased neuronal apotosis, delayed neuronal development and network formation and increased oxidative stress, with all of these phenotypes exacerbated in *PEX11β*-/- mice (Ahlemeyer et al., 2012). Deletion of *PEX11* in *S. cerevisiae* results in a defect in the β-oxidation of medium-chain fatty acids (van Roermund et al., 2000). Taken together, these findings suggest that Pex11p may play a role in a poorly understood metabolic function of peroxisomes in various organisms.

Why does the deletion of *Y. lipolytica PEX11* alone result in a defect in peroxisome assembly such that no peroxisomes are observed? One possible explanation could be the nature of peroxisome dynamics in *Y. lipolytica*, particularly in comparison to those in *S. cerevisiae*. Yeast peroxisomes increase in size and number when grown in medium containing a non-fermentable carbon course. In fact, one of the key reasons why *Y. lipolytica* is such a tractable model to study peroxisome biogenesis is because of its ability to efficiently utilize hydrophobic substrates such as oleic acid, which is accompanied by extensive proliferation of peroxisomes (Smith and Rachubinski, 2001; Chang et al., 2009; 2013). However, growth of *S. cerevisiae* in oleic acid-containing medium has been shown to result primarily in increased peroxisomal size rather than increased peroxisomal number (Tower et al., 2011). In contrast, *Y. lipolytica* cells contain greatly enlarged and much more numerous peroxisomes when grown in oleic acid-containing medium as compared to growth in glucose-containing medium (Smith and Rachubinski, 2001; Chang et al., 2009; 2013). The increased number of peroxisomes in *Y. lipolytica* cells, coupled with this yeast's responsiveness to the presence of fatty acid with regards to its increased capacity to proliferate peroxisomes, could explain why a defect in peroxisome assembly was not heretofore observed in other yeasts deleted for the *PEX11* gene. The loss of Pex11p in *Y. lipolytica* could result in the uncoupling of the process of *de novo* peroxisome biogenesis from the process of peroxisome growth and division. This uncoupling could in turn lead to dysregulation of peroxisome biogenesis and loss of the organelle, evidence for which

is observed in our inability to detect several peroxisomal membrane proteins by immunoblotting of subcellular fractions of *pex11*Δ cells (Figure 3-11A)

Another possible reason why a role in peroxisome assembly *per se* had not been attributed to the Pex11 protein before this current investigation of *Y. lipolytica* is that peroxisome assembly, rather than peroxisome division, was an ancestral function of the Pex11 protein and this function has been retained by other members of the Pex11 protein family. For instance, while humans have three Pex11 proteins, only mutation of *PEX11*β leads to mislocalization of a matrix protein and pathology associated with the PBDs (Ebberink et al., 2012). Similarly, a role in *de novo* peroxisome biogenesis has been attributed to Pex25p in *S. cerevisiae* and *H. polymorpha* (Smith et al., 2002; Saraya et al., 2011). Our phylogenetic analysis demonstrated that Pex25p did not arise by duplication of the gene encoding Pex11p at the *Y. lipolytica* branch point, and resolved the Pex11 family of proteins in Fungi into at least three well supported clades (Pex11p/Pex11Bp, Pex11Cp, and Pex11/25p/Pex25p/Pex27p) (Figure 3-4). These data point to multiple expansions of the Pex11 family, giving rise to the various Pex11 proteins in Fungi. More specifically, since both Pex11p and Pex25p have been implicated in peroxisome biogenesis, our analysis suggests that there were multiple ancestral Pex11 sequences present in the ancestor of Fungi. One of these sequences had a role in peroxisome biogenesis, and different paralogues have retained this ancestral peroxisome biogenic function.

Pex11 family proteins can form homo-oligomeric and hetero-oligomeric protein complexes (Schrader et al., 2012) Constructive neutral evolution supports retention of gene duplication events within such complexes, and allows for increased complexity combined with a dilution of responsibility for maintaining a phenotype among multiple factors (Mast et al., 2014). Thus, while there is one set of evolutionary pressures supporting the expansion of the *PEX11* family, which is possibly independent of the function of the Pex11 family protein complex, taxon-specific functional specialization could selectively retain the separate functions of the complex within distinct paralogues or divide the task among all members in an idiosyncratic manner. This hypothesis is

supported by different lines of functional evidence in Fungi, as no single Pex11 family protein is essential for peroxisome biogenesis in *P. chrysogenum* (Opaliński et al., 2012), whereas the ancestral role of *PEX11* in *de novo* peroxisome biogenesis has been retained by *PEX11* in *Y. lipolytica* and *PEX25* in other yeasts (Smith et al., 2002; Saraya et al., 2011).

Unlike most Pex proteins, the Pex11 protein family has undergone multiple expansions. Our phylogenetic analysis resolved mammalian PEX11α/PEX11β and PEX11γ into separate clades, predating the Opisthokonta, or earlier (Figures 3-4 and 3-5). The distribution observed for the Pex11 protein family and its roles in peroxisomal dynamics are indicative of a protein family that has undergone multiple, independent gene duplications and losses repeating throughout diverse eukaryotic lineages.

In closing, we have shown a role for the classical divisional peroxin Pex11p in peroxisome assembly in the yeast *Y. lipolytica*. We have also provided insight into novel members of the Pex11 protein family in Fungi by demonstrating that Pex11Cp and Pex11/25p play a conserved role in the regulation of peroxisome size and number. By undertaking an evolutionary cell biological study on the history of a protein family and studying peroxisomes in diverse eukaryotes, we were able to uncover a new and unexpected role for Pex11p in peroxisome assembly, with implications for human health through increased understanding of a new class of PBD.

Chapter 4: Comparative genomic survey of peroxisomal peroxisomal proteins in eukaryotic genomes: *Blastocystis hominis, Bodo saltans* and *Naegleria fowleri*

4.1 PTS Finder: a program for predicting PTS-containing proteins

A comparative genomics survey asks essentially whether or not a homologue of interest is present in a genome of interest. The goal of a comparative genomics survey is to establish the presence of a gene in a genome of interest that is the equivalent of another gene already characterized in another system (Klute et al., 2011).

Creating an extensive list of queries for an organelle like the peroxisome can be extremely time-consuming and is not necessarily informative if the genome of interest is rather divergent from the organism from which the queries are taken. Another approach to this issue for organelles with defined targeting signals, *e.g.* PTSs, is to generate a list of all possible proteins that may be targeted to the organelle of interest. I applied this approach to the comparative genomic projects presented in this thesis.

There are a number of different programs already available that can predict targeting to the peroxisome. PSORT is a program for predicting protein subcellular localization based on sorting signals, but the PTS1 motif is too restrictive (Nakai and Horton, 1999). PROSITE is a database of protein domains, families and functional sites, but the consensus sequence for PTS1 motifs is too generalized (Sigrist et al., 2002). Neuberger and colleagues developed a predictor for PTS1-containing proteins (Neuberger et al., 2003b) that searches for a C-terminal tripetide fitting a consensus sequence and which also considers upstream amino acid residues, making the motif a dodecamer overall (see Section 1.4). PeroxiP is another example of a PTS1 predictor (Emanuelsson et al., 2003).

The perl program PTS Finder was written in collaboration with Emily Herman, Department of Cell Biology, University of Alberta. PTS Finder searches for three types of PTS: 1) PTS1 sequences fitting the regular expression: (SAC)-(KRH)-(LA), 2) PTS2 sequences fitting the regular expression: (RK)-(LVIQ)-XX-(LVIHQ)-(LSGAK)-X-(HQ)-(LAF), and 3) non-canonical plant-like PTS sequences fitting one of the regular expressions (SAPCFVGTLKIQ)-(KR)-(LMI) or (SA)-

(KRSNLMHGETFPQCYD)-(LMI) or (SA)-(KR)-(LMIVYF) with a leader sequence comprised of targeting-enhancing peptides (see Section 1.4). The code is included in this thesis.

There are several advantages to using PTS Finder over other PTS-predicting programs. Rather than only accepting a single sequence, PTS Finder accepts an entire file of sequences. While other PTS-predictors are limited to searching for the PTS1 motif alone, PTS Finder can search for three different types of PTS motif at once. PTS Finder prints a table with the identification numbers of the proteins containing PTS motifs, and prints and categorizes the PTS motif found for each protein. Separate files containing the entire protein sequences are also printed for PTS1, PTS2, and plant-like PTS1 sequences. However, PTS Finder will not be able to find highly divergent PTS motifs. Not all proteins predicted will actually be targeted to the peroxisome, and false positives are likely. Proteins can also be targeted to the peroxisome without a recognizable PTS (see Section 1.4).

```
#!/usr/bin/perl


use warnings;
use strict;
use Bio::SeqIO;
use Bio::Seq;

## The PTS Finder program finds canonical peroxisomal targeting signals (PTS1 and
## PTS2), and non-canonical plant peroxisome targeting signals (PTS1p) with a defined leader
## sequence, in a set of sequences.

my $usage = "
        PTS Finder
        ----------
        Finds peroxisomal targeting signals in a set of sequences
        and outputs the signal sequence and the fasta sequences of
        PTS-containing proteins.
        ----------
        Usage: PTS_Finder.pl [input]
        Output: PTS1_sequences.fa, PTS2_sequences.fa, PTS1p_sequences.fa,
PTS1p_VAR_sequences.fa, PTS_table.txt
";

if (@ARGV == 0 && -t STDIN && -t STDERR) {
   die "$usage\n";
}

open (OUTFILE, '>', 'PTS_table.txt') or die "Cannot open outfile for writing: $!\n";
my $PTS1file = 'PTS1_sequences.fa';
my $seqoutPTS1 = Bio::SeqIO->new(-file => ">$PTS1file", -'format' => 'fasta');

my $PTS2file = 'PTS2_sequences.fa';
my $seqoutPTS2 = Bio::SeqIO->new(-file => ">$PTS2file", -'format' => 'fasta');

my $PTS1pfile = 'PTS1p_sequences.fa';
my $seqoutPTS1p = Bio::SeqIO->new(-file => ">$PTS1pfile", -'format' => 'fasta');

my $PTS1pVARfile = 'PTS1p_VAR_sequences.fa';
my $seqoutPTS1p_VAR = Bio::SeqIO->new(-file => ">$PTS1pVARfile", -'format' => 'fasta');


# Load sequences from file

my $infile = $ARGV[0];

my $inseq = Bio::SeqIO->new(-file => $infile, -'format' => 'fasta');

while (my $seq = $inseq->next_seq()) {
my $len = $seq->length();
```

```
# Find PTS1 signal
## PTS1 is (S|A|C)(K|R|H)(L|A) and found at the extreme C-terminus
my $Cterm_minus2=$len-2;
my $Cterm = $seq->subseq($Cterm_minus2,$len);
        if ($Cterm=~m/[SAC][KRH][LA]/) {
        my $accession = $seq->primary_id;
        print OUTFILE "$accession\t$Cterm\t-\t-\t-\n";

        $seqoutPTS1->write_seq($seq);
                }


# Find PTS2 signal
## PTS2 is (R|K)(L|V|I|Q)XX(L|V|I|H|Q)(L|S|G|A|K)X(H|Q)(L|A|F) and is found within the first 20
## amino acids of the sequence
if ($len >= 20) {
my $Nterm = $seq->subseq(1,20);
if ($Nterm=~m/[RK][LVIQ]..[LVIHQ][LSGAK].[HQ][LAF]/) {
        my $accession= $seq->primary_id;
        print OUTFILE "$accession\t-\t$Nterm\t-\t-\n";

        $seqoutPTS2->write_seq($seq);
        }
        }


# Find PTS1p signal
## PTS1p is found at the extreme C-terminus, there are three options
## Option 1 is (S|A|P|C|F|V|G|T|L|K|I|Q)(K|R)(L|M|I)
## Option 2 is (S|A)(K|R|S|N|L|M|H|G|E|T|F|P|Q|C|Y|D)(L|M|I)
## Option 3 is (S|A)(K|R)(L|M|I|V|Y|F)
my $Cterm_2_minus2=$len-2;
my $Cterm_2 = $seq->subseq($Cterm_2_minus2,$len);
        if (($Cterm_2=~m/[SAPCFVGTLKIQ][KR][LMI]/) |
($Cterm_2=~m/[SA][KRSNLMHGETFPQCYD][LMI]/) | ($Cterm_2=~m/[SA][KR][LMIVYF]/)) {
        my $accession = $seq->primary_id;
        my $PTS1pVAR_found = 0;

        ## Find 4-5 residues within the nine residues before SKL (KRHPSTAVDE)
        my $SKL_minus9 = $len - 11;
        my $Cterm_2_VAR = $seq->subseq($SKL_minus9,($Cterm_2_minus2-1));
        if (($Cterm_2_VAR =~ m/[KRHPSTAVDE][KRHPSTAVDE][KRHPSTAVDE]
        [KRHPSTAVDE]/) | ($Cterm_2_VAR=~m/[KRHPSTAVDE][KRHPSTAVDE]
        [KRHPSTAVDE][KRHPSTAVDE][KRHPSTAVDE]/)) {
                $accession = $seq->primary_id;
                $PTS1pVAR_found = 1;
                }

        if ($PTS1pVAR_found ==1) {
                print OUTFILE "$accession\t-\t-\t$Cterm_2\t$Cterm_2_VAR\n";
                $seqoutPTS1p_VAR->write_seq($seq);
```

```
                }
        else {
                print OUTFILE "$accession\t-\t-\t$Cterm_2\t-\n";
                }
        $seqoutPTS1p->write_seq($seq);

        }
}
```


The PTS Finder code has been re-paginated to fit this thesis.

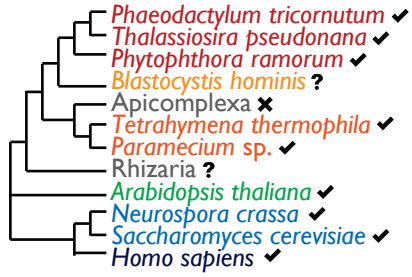
4.2 Comparative genomic survey of peroxins in *Blastocystis hominis*

In order to determine the peroxisomal protein complement of *Blastocystis*, it is critical to have an understanding of the conservation of peroxins in other genomes from the same eukaryotic supergroup and in the ancestral lineages giving rise to *Blastocystis*. Although the majority of *PEX* genes are thought to have a eukaryotic origin, previous comparative genomic studies concluded that the majority of peroxins were not present in the ancestor of eukaryotes (Gabaldón et al., 2006; Schlüter et al., 2006). However, due to the unavailability of a taxonomically broad range of eukaryotic genome sequences, taxa from the SAR supergroup were essentially not sampled. Additionally, novel peroxins whose evolutionary distribution has never been examined have been identified since this time (Managadze et al., 2010; Tower et al., 2011).

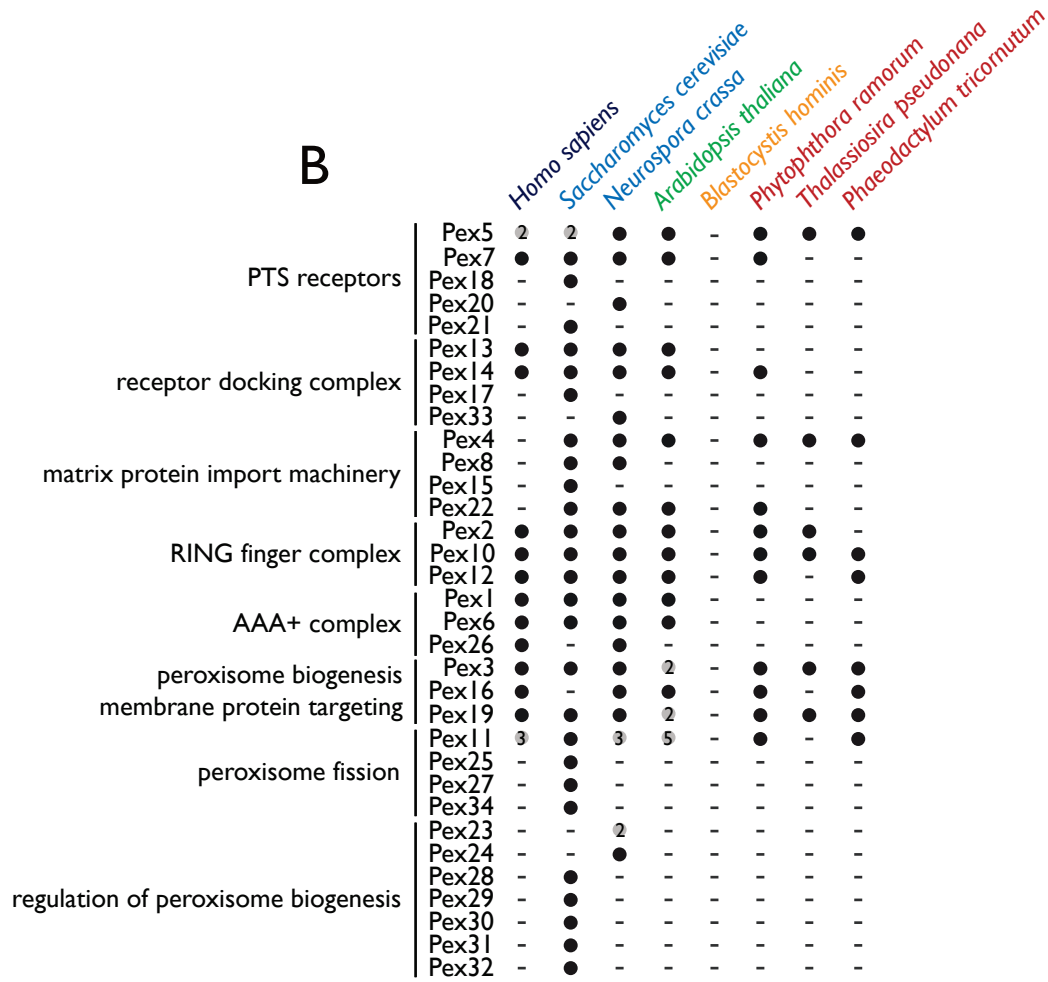
As a starting point for investigation into the peroxisomal protein complement of *Blastocystis*, we undertook a comparative genomic survey of *PEX* genes in the evolutionarily relevant stramenopile genomes *Phytophthora ramorum*, *Phaedactylum tricornutum* and *Thalassiosira pseudonana*. These taxa were key sampling points, as peroxisomes have been characterized or predicted in these organisms (Philippi et al., 1975; Armbrust et al., 2004; Gonzalez et al., 2011). As previously reported, a number of peroxins are Fungi-specific and were not identified in stramenopiles (Figure 4-1) (Appendices 9, 10, 11). These genomes did encode a number of *PEX* genes whose proteins (Pex1p, Pex2p, Pex4p, Pex5p, Pex10p, Pex14p) were previously identified as being part of the ancestral eukaryotic peroxisomal proteome (Gabaldón et al., 2006). The peroxins Pex3p, Pex19p, Pex10p and Pex12p were previously denoted as peroxisomal markers, because they were ubiquitously found in organisms with peroxisomes and absent from organisms without peroxisomes (Schlüter et al., 2006). In agreement with a previous analysis (Gonzalez et al., 2011), the receptor for PTS1-containing proteins, Pex5p, is encoded in all three genomes, but Pex7p, the recptor recognizing PTS2-containing proteins was not identified in *T. pseudonana* and *P. tricornutum*. The RING finger complex made up of Pex2p, Pex10p and Pex12p was relatively well conserved,

Figure 4-1. Comparative genomic survey of proteins encoded by *PEX* genes in *Blastocystis hominis* and selected genomes from the SAR supergroup. (A) An overview of our knowledge of peroxisome presence or absence in selected genomes from the SAR eukaryotic supergroup. Check marks indicate the presence of peroxisomes based on ultrastructural or biochemical evidence, crosses indicate that peroxisomes are absent and question marks indicate that peroxisome presence or absence has not been determined. The evolutionary relationships between the genomes are given at left, but the rooting of the tree is arbitrary. (B) Each row represents a peroxin that has been characterized in *H. sapiens*, *S. cerevisiae* and/or *N. crassa*. *PEX* genes are grouped according to general function. A black circle indicates the presence of a protein in the indicated taxon based on positive reciprocal BLAST or pHMMer searches, grey circles with numbers indicate multiple paralogues of a protein and dashes indicate that proteins were not identified in the indicated taxon.

A


Phaeodactylum tricornutum ✔
Thalassiosira pseudonana ✔
Phytophthora ramorum ✔
Blastocystis hominis ?
Apicomplexa ✖
Tetrahymena thermophila ✔
Paramecium sp. ✔
Rhizaria ?
Arabidopsis thaliana ✔
Neurospora crassa ✔
Saccharomyces cerevisiae ✔
Homo sapiens ✔

B


Homo sapiens
Saccharomyces cerevisiae
Neurospora crassa
Arabidopsis thaliana
Blastocystis hominis
Phytophthora ramorum
Thalassiosira pseudonana
Phaeodactylum tricornutum
PTS receptors
Pex5 2 2 ● ● - ● ● ●
Pex7 ● ● ● ● - ● - -
Pex18 - ● - - - - - -
Pex20 - - ● - - - - -
Pex21 - ● - - - - - -
receptor docking complex
Pex13 ● ● ● ● - - - -
Pex14 ● ● ● ● - ● - -
Pex17 - ● - - - - - -
Pex33 - - ● - - - - -
matrix protein import machinery
Pex4 - ● ● ● - ● ● ●
Pex8 - ● ● - - - - -
Pex15 - ● - - - - - -
Pex22 - ● ● ● - ● - -
RING finger complex
Pex2 ● ● ● ● - ● ● -
Pex10 ● ● ● ● - ● ● ●
Pex12 ● ● ● ● - ● - ●
AAA+ complex
Pex1 ● ● ● ● - - - -
Pex6 ● ● ● ● - - - -
Pex26 ● - ● - - - - -
peroxisome biogenesis
membrane protein targeting
Pex3 ● ● ● 2 - ● ● ●
Pex16 ● - ● ● - ● - ●
Pex19 ● ● ● 2 - ● ● ●
peroxisome fission
Pex11 3 ● 3 5 - ● - ●
Pex25 - ● - - - - - -
Pex27 - ● - - - - - -
Pex34 - ● - - - - - -
regulation of peroxisome biogenesis
Pex23 - - 2 - - - - -
Pex24 - - ● - - - - -
Pex28 - ● - - - - - -
Pex29 - ● - - - - - -
Pex30 - ● - - - - - -
Pex31 - ● - - - - - -
Pex32 - ● - - - - - -

with Pex10p and Pex12p encoded by all three genomes. The ubiquitin-conjugating enzyme Pex4p was also encoded by all three genomes. However, Pex13p and Pex14p, components of the peroxisomal protein docking complex, and Pex1p and Pex6p, components of the AAA+ ATPase complex, were not identified. The peroxins Pex3p, Pex16p and Pex19p implicated in *de novo* peroxisome formation from the ER were well conserved, as was Pex11p.

A previous genome project for *Blastocystis hominis* subtype 7 has been completed (Denoeud et al., 2011). The *B. hominis* subtype I, NandII strain is a proprietary dataset to which access was granted by Eleni Gentekaki (Dalhousie University). It was noted during early annotation of this new *Blastocystis* genome (which is also a different strain) that a number of components absent from the genome project done by Denoeud and colleagues were actually present. For example, a complete set of ESCRT complexes, membrane trafficking machinery involved in multivesciular body biogenesis, was identified in the new *Blastocystis* genome (Emily Herman, Department of Cell Biology, University of Alberta, unpublished data). Thus, we proceeded with comparative genomic searches for potential peroxisomal proteins encoded by the *Blastocystis* genome.

However, no *PEX* genes were identified in *Blastocystis* (Figure 4-1), with the exception of a possible gene encoding a Pex4p homologue (Appendix 12) even after using queries from the stramenopile genomes included in this study (Section 2.11.2). Pex4p is a challenging bioinformatic query because it contains a conserved ubiquitin-ligase domain. Thus, we concluded that this protein was not a convincing homologue of Pex4p, and the presence of only a single peroxin was not likely for this genome. This conclusion is in agreement with peroxisomes not being predicted ultrastructurally for most parasitic organisms (reviewed in Gabaldón, 2010) and peroxisomes not being morphologically identified in transmission electron micrographs of *Blastocystis* cells (personal communication, Eleni Gentekaki, Dalhousie University).

4.3 Comparative genomic survey of peroxins and glycosomal proteins in *Bodo saltans*

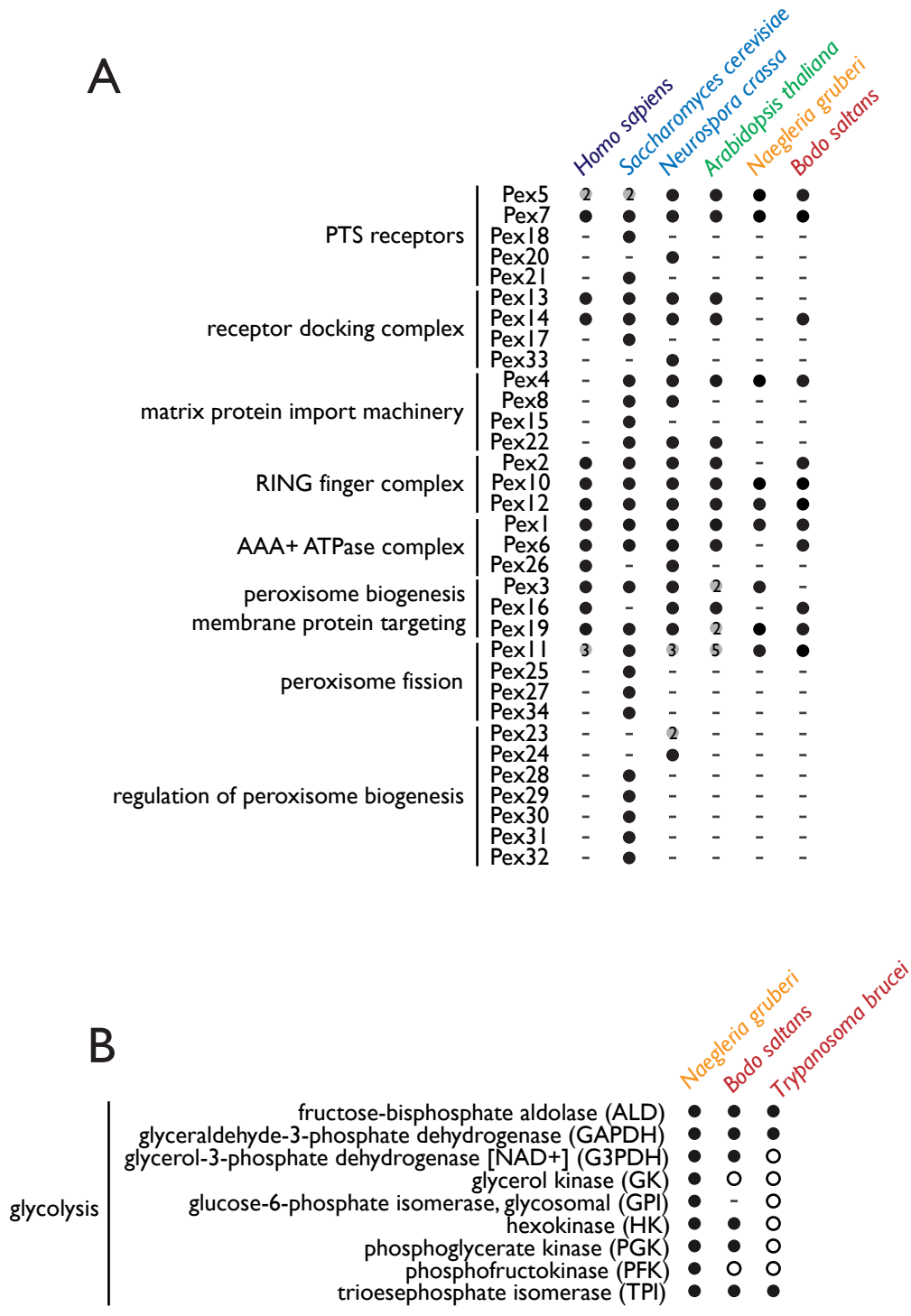
As in *Blastocystis*, when studying the glycosomal protein complement of parasitic trypanosomes it is beneficial to understand the glycosomal and/or peroxisomal complement of the ancestral taxa giving rise to this lineage. We wanted to establish an ancestral excavate peroxisomal protein complement in order to understand the complement of other genomes from this supergroup. Thus, we completed a comparative genomic survey of the peroxisomal proteome in *Naegleria gruberi*, a sister taxon to *Bodo saltans*.

N. gruberi has a relatively complete complement of *PEX* genes (Figure 4-2A) (Appendix 13A). Pex5p and Pex7p, the receptors for PTS1- and PTS2-containing proteins, respectively, were identified, as were some of the peroxins comprising the RING finger and AAA+ ATPase complexes. Pex11p was present, as were Pex3p, Pex16p and Pex19p required for *de novo* peroxisome formation. Surprisingly, Pex13p and Pex14p, components of the peroxisomal protein docking complex were not identified.

B. saltans also encodes a relatively complete complement of *PEX* genes, including some for peroxins not found in *N. gruberi* (Figure 4-2A) (Appendix 14A). *B. saltans* encodes an incomplete peroxisomal protein docking complex (Pex14p only). Like *N. gruberi*, a homologue of Pex26p, the peroxisomal receptor for the AAA+ ATPase complex, was not identified in *B. saltans*. The biogenic peroxins Pex16p and Pex19p were found in *B. saltans*, but both could be considered divergent. Pex16p did not retrieve homologues in well characterized models, and Pex19p lacked the well conserved CAAX box domain. Surprisingly, a Pex3p homologue was not found in *B. saltans*; whereas other organisms lacking Pex3p also lack peroxisomes, glycosomes are still present in *B. saltans.*

We also completed comparative genomic searches for 201 high-confidence glycosomal proteins from *T. brucei* (Guther et al., 2014) in *B. saltans* (Appendix 15A-C). In *B. saltans*, 115 named *T. brucei* proteins retrieved homolgoues in *B. saltans*. 54 unnamed, hypothetical *T. brucei* proteins retrieved homologues in *B. saltans*. 32 named or unnamed *T. brucei* queries did not have

Figure 4-2. Comparative genomic survey of proteins encoded by *PEX* genes and glycolysis enzymes in *Naegleria gruberi* and Bodo *saltans*. (A) Each row represents a peroxin that has been characterized in *H. sapiens*, *S. cerevisiae* and/or *N. crassa*. *PEX* genes are grouped according to general function. A black circle indicates the presence of a protein in the indicated taxon based on positive reciprocal pHMMer searches, grey circles with numbers indicate multiple paralogues of a protein, and dashes indicate that proteins were not identified in the indicated taxon. (B) Glycolysis enzymes in *Naegleria gruberi*, *Bodo saltans* and *T. brucei*. A black circle indicates the presence of a protein in the indicated taxon based on positive reciprocal pHMMer searches and dashes indicate that proteins were not identified in the indicated taxon. White circles indicate that the protein has a predicted PTS. Paralogue numbers are not included.


A
Homo sapiens
Saccharomyces cerevisiae
Neurospora crassa
Arabidopsis thaliana
Naegleria gruberi
Bodo saltans
PTS receptors
Pex5
Pex7
Pex18
Pex20
Pex21
receptor docking complex
Pex13
Pex14
Pex17
Pex33
matrix protein import machinery
Pex4
Pex8
Pex15
Pex22
RING finger complex
Pex2
Pex10
Pex12
AAA+ ATPase complex
Pex1
Pex6
Pex26
peroxisome biogenesis
membrane protein targeting
Pex3
Pex16
Pex19
peroxisome fission
Pex11
Pex25
Pex27
Pex34
regulation of peroxisome biogenesis
Pex23
Pex24
Pex28
Pex29
Pex30
Pex31
Pex32
B
Naegleria gruberi
Bodo saltans
Trypanosoma brucei
glycolysis
fructose-bisphosphate aldolase (ALD)
glyceraldehyde-3-phosphate dehydrogenase (GAPDH)
glycerol-3-phosphate dehydrogenase [NAD+] (G3PDH)
glycerol kinase (GK)
glucose-6-phosphate isomerase, glycosomal (GPI)
hexokinase (HK)
phosphoglycerate kinase (PGK)
phosphofructokinase (PFK)
trioesephosphate isomerase (TPI)

homolgoues in *B. saltans*. However, we cannot exclude the possibility that this result is due to sequence divergence between *B. saltans* and *T. brucei* (*i.e. B. saltans* may in fact have functional homologues of high confidence glycosomal proteins that cannot be identified using a single query for a comparative genomic search). Although glycosomes are present in *B. saltans*, these results indicate that there are some proteins specific to *T. brucei*, and possibly other parasitic trypanosomatidae. We also searched for the presence of PTSs in *B. saltans* homologues of *T. brucei* high-confidence glycosomal proteins using PTS Finder. Strikingly, only 14 of the *B. saltans* proteins contained recognizable PTSs (Appendix 15A-B). This result is suggestive of reduced targeting to the glycosome in *B. saltans*.

The compartmentalization of glycolysis is the defining feature of the glycosome. Hence, glycolysis enzymes, most of which are targeted to the glycosome, were among the high-confidence *T. brucei* glycosomal proteins. Homologues of these enzymes, with the exception of glucose-6-phosphate isomerase, were present in the *B. saltans* genome, but only two contained PTSs (Figure 4-2B) (Appendix 14B). A *B. saltans* glucose-6-phosphate isomerase homologue could not be identified using the corresponding *N. gruberi* query. This result is indicative of a possible reduction in the glycolysis pathway in *B. saltans*. Previous work showed that the *N. gruberi* genome encodes a complete complement of glycosomal proteins (Figure 4-2B) (Appendix 13B) (Fritz-Laylin et al., 2010). PTS Finder was used to confirm that none of the *N. gruberi* glycolysis enzymes contain PTSs (Appendix 13B).

PTS Finder was used to predict PTSs in all proteins in *B. saltans*, *N. gruberbi* and *N. fowleri*, as well as all finished genomes on TryTrypDB. PTS-containing proteins in *B. saltans* were classified based on the top reciprocal pHMMer hit in the *T. brucei* genome, and many of these proteins did not have homologues in *T. brucei* (Appendix 16). Although this analysis did not take into account genome size, a definitive reduction in the number of putative peroxisome-targeted proteins in *N. gruberi* as compared to the remaining excavate genomes was evident (Table 4-1). The total number of PTS-containing predicted proteins across parasite genomes and their closely related non-parasitic

relative *C. fasciculata* was similar, with PTS1 appearing to be the dominant pathway. We predict that there may be an issue with the *T. cruzi* CLBrener genome due to the low number of PTS-containing proteins identified.

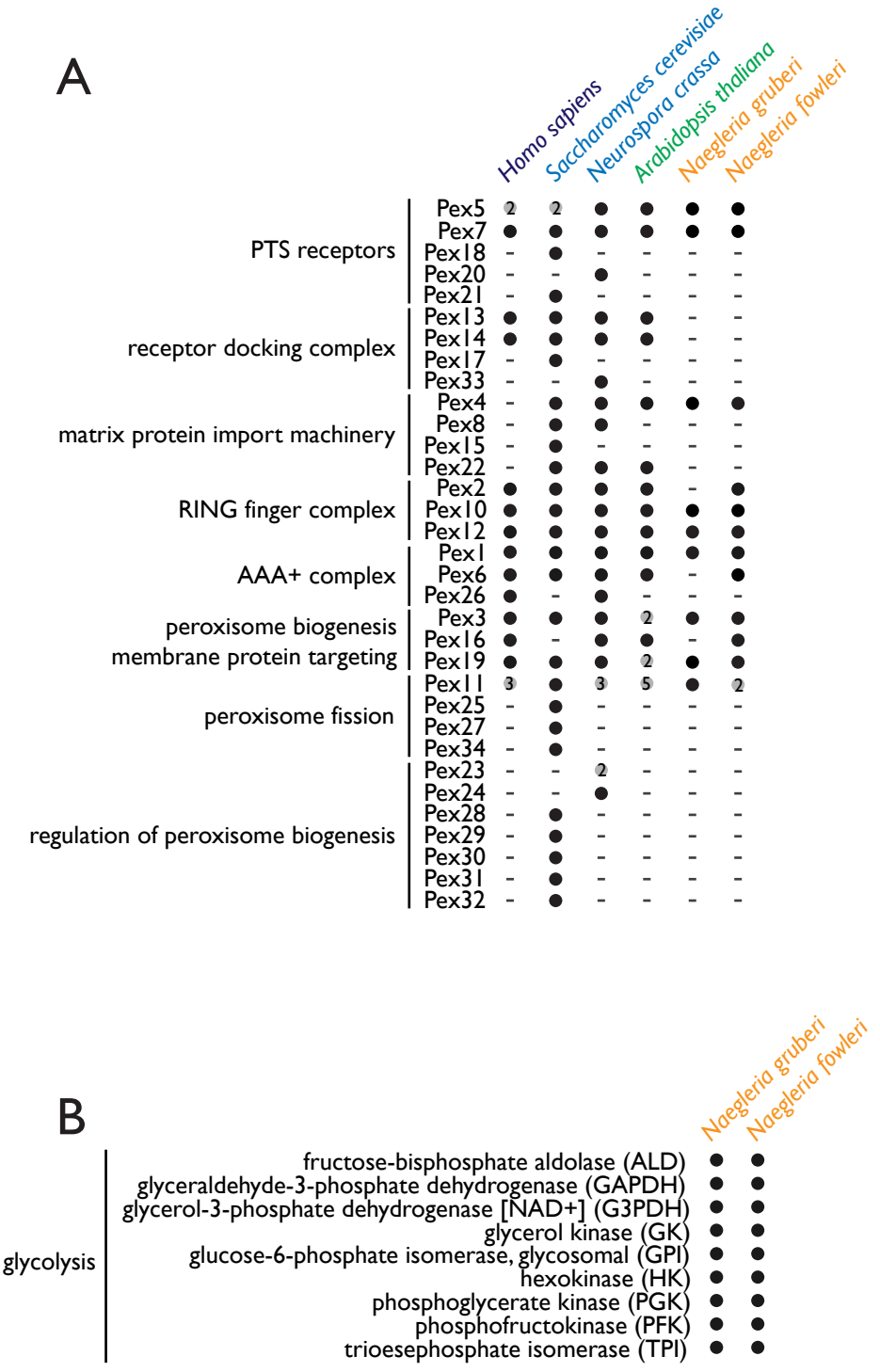
Table 4-1. Number of PTS-containing proteins in trypanosomatid genomes and outgroups

Genome (Strain)	PTS1	PTS2	PTS1p	Total PTS
Bodo saltans	113	6	292	411
Crithidia fasciculata CfCl	123	16	212	351
Leishmania braziliensis MHOMBR75M2903	73	3	141	217
Leishmania braziliensis MHOMBR75M2904	73	2	138	213
Leishmania donovani BPK282A1	61	6	147	214
Leishmania infantum JPCM5	62	6	149	217
Leishmania major Friedlin	66	7	151	224
Leishmania mexicana MHOMGT2001U1103	70	7	142	219
Leishmania tarentolae ParrotTarII	64	6	129	199
Naegleria fowleri	74	7	138	219
Naegleria gruberi	3	8	9	20
Trypanosoma brucei GambienseDAL972	54	2	147	203
Trypanosoma brucei Lister427	50	2	140	192
Trypanosoma brucei TREU927	64	7	172	243
Trypanosoma congolense IL3000	61	6	150	217
Trypanosoma cruzi CLBrener	3	0	26	29
Trypanosoma cruzi CLBrenerEsmeraldo	59	3	143	205
Trypanosoma cruzi CLBrenerNon-Esmeraldo	58	4	152	214
Trypanosoma evansi STIB05	56	2	139	197
Trypanosoma grayi ANR4	83	7	166	256
Trypanosoma vivax Y486	59	3	118	180

4.4 Comparative genomic survey of peroxins in *Naegleria fowleri*

A comparison between the peroxin complement of *N. gruberi* and *N. fowleri* is potentially informative of the unique biology of a deadly pathogen. *N. gruberi PEX* genes were identified in Section 4.3. The complement of *N. fowleri PEX* genes is similar to that of *N. gruberi* (Figure 4-3A) (Appendix 17A). The biogenic peroxins Pex3p, Pex16p and Pex19p are present, as is the divisional peroxin Pex11p. Both PTS1 and PTS2 receptors are conserved, as are components of the RING finger and AAA+ ATPase complexes. Interestingly, both *N. gruberi* and *N. fowleri* appear to be missing all of the components of the receptor docking complex. Comparative genomics searches

Figure 4-3. Comparative genomic survey of proteins encoded by *PEX* genes and glycolysis enzymes in *Naegleria fowleri*. Each row represents a peroxin that has been characterized in *H. sapiens*, *S. cerevisiae* and/or *N. crassa*. *PEX* genes are grouped according to general function. A black circle indicates the presence of a protein in the indicated taxon based on positive reciprocal BLAST or pHMMer searches, grey circles with numbers indicate multiple paralogues of a protein and dashes indicate that proteins were not identified in the indicated taxon. (B) Glycolysis enzymes in *Naegleria gruberi* and *Naegleria fowleri*. A black circle indicates the presence of a protein in the indicated taxon based on positive reciprocal pHMMer searches and dashes indicate that proteins were not identified in the indicated taxon. White circles indicate that the protein has a predicted PTS. Paralogue numbers are not included.


A
Homo sapiens
Saccharomyces cerevisiae
Neurospora crassa
Arabidopsis thaliana
Naegleria gruberi
Naegleria fowleri
PTS receptors
Pex5
Pex7
Pex18
Pex20
Pex21
receptor docking complex
Pex13
Pex14
Pex17
Pex33
matrix protein import machinery
Pex4
Pex8
Pex15
Pex22
RING finger complex
Pex2
Pex10
Pex12
AAA+ complex
Pex1
Pex6
Pex26
peroxisome biogenesis membrane protein targeting
Pex3
Pex16
Pex19
peroxisome fission
Pex11
Pex25
Pex27
Pex34
regulation of peroxisome biogenesis
Pex23
Pex24
Pex28
Pex29
Pex30
Pex31
Pex32
B
Naegleria gruberi
Naegleria fowleri
glycolysis
fructose-bisphosphate aldolase (ALD)
glyceraldehyde-3-phosphate dehydrogenase (GAPDH)
glycerol-3-phosphate dehydrogenase [NAD+] (G3PDH)
glycerol kinase (GK)
glucose-6-phosphate isomerase, glycosomal (GPI)
hexokinase (HK)
phosphoglycerate kinase (PGK)
phosphofructokinase (PFK)
trioesephosphate isomerase (TPI)

for glycolysis enzymes in *N. fowleri* revealed an identical complement to *N. gruberi*, with homologues of all glycolysis enzymes present (Figure 4-3B) (Appendix 16B). PTS Finder was used to show that like *N. gruberi*, none of the *N. fowleri* glycolysis enzymes contain PTSs (Appendix 17B).

4.5 Discussion

4.5.1 Comparative genomic survey of peroxins in Blastocystis hominis

The absence of peroxins is consistent with no observed morphologically observed peroxisomes in *Blastocystis*. However, we have added to our understanding of peroxisome evolution by completing the first comparative genomic survey of *PEX* genes in non-parasitic genomes from the SAR clade: *P. tricornutum*, *P. ramorum* and *T. pseudonana*. Taken together, the peroxisome complement of these SAR genomes is actually quite comparable to that of humans or of *A. thaliana*. It would not be unexpected to find *bona fide* peroxisomes in these organisms.

The receptor for PTS2-containing proteins, Pex7p, and its co-receptors Pex18p, Pex20p and and Pex21p, were absent from several of the SAR genomes queried. Previous work showed that the PTS2 import pathway is absent in the diatom *P. tricornutum* (Gonzalez et al., 2011). The authors also noted that another diatom species, *T. pseudonana*, also lacks Pex7p, and like *P. tricornutum*, its peroxisomal thiolase has a PTS1 motif instead of the usual PTS2 motif. The authors suggested that chromalveolate morphology, and more specifically how intracellular trafficking occurs in diatoms, could be the reason for a switch to the PTS1 system exclusively (Gonzalez et al., 2011). Plastid-targeted proteins encoded by the nuclear genome encode N-terminal bipartite targeting sequences so that these proteins can enter the four-membrane-bound chloroplast. However, some chromalveolates have a functional PTS2 pathway and complex plastids, so the presence of these leader sequences does not necessarily mean that the PTS2 motif cannot be functional. The cell wall of diatoms is also quite complex, and intracellular trafficking has been modified to deal with this differing morphology.

Diatoms and other eukaryotes from across the diversity of eukaryotes also lack the PTS2 system. PTS2-containing proteins were not identified in the red alga *Cyanidiochyzon merolae*,

although PTS2-containing proteins are successfully targeted to the peroxisome in green algae like *Chlamydomonas reinhardtii* and *Closterium ehrenbergii* (Shinozaki et al., 2009). Although the *Drosophila melanogaster* genome encodes homologues of the PTS2 receptor Pex7p and of Pex20p, the Pex7p co-receptor in some species of yeast (Mast et al., 2011), *D. melanogaster* lacks canonical PTS2-containing proteins and is unable to import chimeric fluorescent proteins carrying a human PTS2 into peroxisomes (Faust et al., 2012). Peroxisomal thiolase has a PTS1 motif in *D. melanogaster*. Previous work on the PTS2 system in *D. melanogaster* is limited by the inherent weaknesses of bioinformatic approaches and a lack of direct functional testing. *D. melanogaster* proteins may contain PTS motifs that do not fit the canonical PTS2 definition. Although Pex7p and Pex20p homologues have been indentified in *D. melanogaster*, it is possible that there are other highly divergent paralogues of these proteins. Also, sequence and/or structural conservation of *D. melanogaster* Pex7p is not by itself indicative of PTS2 pathway functionality. One approach to testing whether or not *D. melanogaster* Pex7p is a genuine PTS2 receptor would be to determine if it could rescue import of a PTS2-containing peroxisomal matrix protein in *S. cerevisiae pex7*Δ cells.

The worm *Caenorhabditis elegans* also lacks an identifiable Pex7p homologue and encodes a PTS1-containing thiolase (Motley et al., 2000). Motley and colleagues propose a possible explanation for why both *C. elegans* and *D. melanogaster* do not appear to have a functioning PTS2 system: the genome of *C. elegans* is known to be fast-evolving, so while it has lost all remnants of the PTS2 system, *i.e.* receptor, co-receptors and actual motifs, the more slowly evolving genome of *D. melanogaster* still retains the PTS2 receptor and its co-receptors, even though the pathway is non-functional (Motley et al., 2000). The PTS1 tripeptide would have been relatively easy to acquire. The switch from using the PTS2 pathway to using the PTS1 pathway must have occurred before the loss of the PTS2 pathway, because loss of the proper localization and targeting of peroxisomal enzymes would have been detrimental to the organism. Thus, a possible explanation for multiple eukaryotes lacking PTS2 targeting is multiple independent losses of the PTS2 targeting machinery throughout eukaryotic history.

4.5.2 Comparative genomic survey of peroxins and glycosomal proteins in Bodo saltans

Preliminary results from our comparative genomic survey of putative glycosomal proteins is indicative of a reduction both in terms of the glycosomal proteome and of protein targeting to the glycosome from the trypansomatids to their closest relatives. Proteomic analysis with the genomes surveyed here will aid our understanding of the transition to parasitism. *B. saltans* is an important trypanosomatid outgroup because although it has glycosomes, it is non-parasitic. Many *T. brucei* glycosomal proteins that do not have homologues in *B. saltans* are hypothetical or unnamed proteins. Looking at the localization and/or function of these proteins could be informative as to what makes the glycosome in deadly trypanosomes distinct.

It was interesting to note that *B. saltans* lacks a homologue of Pex3p, which is required for peroxisome formation in nearly every other eukaryote studied. Like *B. saltans*, *Trypanosoma* and *Leishmania* also lack identifiable Pex3p homologues (unpublished data). The mechanism of *de novo* peroxisome biogenesis in these organisms remains an open question, and it may be different from that of the human host. Still, Pex13p was shown to traffic through the ER en route to the peroxisome in trypanosomes (Guther et al. 2014), so a form of this peroxisome biogenic pathway is present in these organisms. This pathway has high potential as a possible target for therapeutic intervention because a feature of an ideal drug target is one that will interrupt the trypanosomal biogenic pathway without interfering with peroxisome formation in the human host.

It was surprising to note the nearly complete loss of PTSs in *N. gruberi*. *N. gruberi* is typically thought of as a fairly canonical excavate representative for comparative genomic analyses; however, our results showing that the peroxisome may be quite reduced in *N. gruberi* suggest that this may not be the case.

4.5.3 Comparative genomic survey of peroxins in Naegleria fowleri

Peroxisomes have not been described ultrastructurally for *N. gruberi*, although their presence was inferred based on homologues of *PEX* genes and of genes encoding peroxisomal metabolic enzymes (Fritz-Laylin et al., 2010). Although a relatively complete set of *PEX* genes was found for

both *N. gruberi* and *N. fowleri*, it was surprising to discover that their genomes do not encode members of the peroxisomal protein docking complex. Since the PTS1 receptor Pex5p and the protein docking complex constituent Pex14p form the expandable peroxisomal importomer pore (Meinecke et al., 2010), it is difficult to conceive how matrix proteins could be imported into peroxisomes, if they do indeed exist, in *N. gruberi* and *N. fowleri* in light of the absence of components of the peroxisomal protein docking complex in these organisms. However, the docking complex was also absent from *P. tricornutum*, which contains peroxisomes (Gonzalez et al., 2011), and is absent or highly reduced in *P. ramorum* and *T. pseudonana*. It is possible that the proteins of this complex in *Naegleria* sp. are too divergent to be detected by our bioinformatics searches. For example, in *A. thaliana*, the Pex26p homologue APEM9 does not retrieve the corresponding homologues from human or yeast (Goto et al. 2011). Most ancestral *PEX* genes function in matrix protein import, and the peroxisome is defined by its numerous enzymes located in its matrix. Matrix protein import is therefore very likely a highly conserved process.

Chapter 5: Perspectives

5.1 Synopsis

This thesis took two very different approaches to the topic of peroxisome evolution. First, I completed a comprehensive comparative genomic and phylogenetic study of a single peroxin family. Second, I completed a much broader survey of all peroxins in multiple extant eukaryotic genomes. I demonstrated a conserved Pex11 protein family role in the regulation of peroxisomal size and number for the novel Pex11 family proteins Pex11Cp and Pex11/25p in *Y. lipolytica*. I also showed a new role for the divisional peroxin Pex11p in peroxisome assembly in *Y. lipolytica* and proposed an ancestral role for Pex11p in *de novo* peroxisome biogenesis as well as peroxisomal division. Finally, I contributed to three international genome projects by determining the complement of peroxins and/or peroxisomal proteins in *Blastocystis hominis* (an intestinal parasite), *Bodo saltans* (a close relative of the parasitic *Trypanosoma* and *Leishmania*) and *Naegleria fowleri* (the 'brain eating amoeba').

5.2 Future directions for the study of the Pex11 protein family in *Yarrowia lipolytica*

Many points about the precise molecular mechanism of peroxisomal dynamics, including the division of peroxisomes, remain poorly understood. Bioinformatic searches of the promoter regions of the *Y. lipolytica PEX11* gene family revealed putative Far binding sites, and many *PEX* genes in *S. cerevisiae* and *Y. lipolytica* are inducible by fatty acid-containing medium. To confirm the activity of the *Y. lipolytica* FarA/B homologue Por1, Poopantipan and colleagues demonstrated decreased RNA levels for enzymes involved in β-oxidation in *por1*Δ cells compared to wild-type cells using Northern blots (Poopanitpan et al., 2010). To confirm this finding, Western blotting could be used to determine if there is a decrease in the levels of enzymes involved in β-oxidation and/or peroxins in *por1*Δ cells. A lacZ reporter construct could also be used, utilizing a promoter from a fatty acid-inducible gene and measuring beta-galactosidase activity in *por1*Δ cells. To confirm putative Far binding sites in *Y. lipolytica* genes, including the Pex11 family, an electrophoretic mobility shift assay (EMSA) could be utilized to determine if Por1 interacts with putative elements; this technique is

based on the principle that a complex will run slower on an agarose gel compared with non-interactors.

Are the genes of the *Y. lipolytica PEX11* gene family expressed, and are these genes also inducible by incubation of *Y. lipolytica* in fatty acid-containing medium? Real-time PCR could be used to determine the relative expression levels of these genes in constitutively dividing cells and in cells exposed to fatty acids. In addition, Ebberink and colleagues found that overexpression of *PEX11*γ could partially complement the *PEX11*β defect in a PBD patient (Ebberink et al., 2012). Similarly, higher levels of *PEX11*γ mRNA were found in *PEX11*β*-/-* knock out mice (Ahlemeyer et al., 2011), suggesting a possible compensatory mechanism by *PEX11*γ for the deficiency in *PEX11*β. This finding is interesting in the context of the evolutionary relationship we have shown between *PEX11*γ and Pex11Cp. Measuring the relative expression levels of each of the *Y. lipolytica PEX11* gene family members in *pex11*Δ, *pex11C*Δ and *pex11/25*Δ strains could show if a similar compensatory mechanism of one *PEX11* gene family member by another occurs in *Y. lipolytica*.

Pex11p forms homodimers and heterodimers, and other Pex11 family proteins, *e.g.* Pex25p and Pex27p, also interact with each other (Marshall et al., 1996; Tam et al., 2003). A yeast two-hybrid screen could be used to determine whether or not *Y. lipolytica* Pex11 protein family members interact. Alternatively, because yeast two-hybrid analysis does not differentiate between direct and bridged protein interactions, a pull down assay could be used to determine if *Y. lipolytica* Pex11 protein family members interact directly. This assay for direct protein binding is based on the immobilization of GST fusion proteins of interest on glutathione resin, followed by incubation with *E. coli* cell lysates containing MBP fusion proteins of interest. These types of experiments could also determine if there are interactions between Pex11p, Pex11Cp and Pex11/25p and other peroxins in *Y. lipolytica*. A recent proteomics study showed that *S. cerevisiae* Pex11p interacts with the membrane-associated peroxins making up the peroxisomal importomer complex (Oeljeklaus et al., 2012). Defining the interactions between Pex11p and other proteins, particularly those acting at the

peroxisomal membrane, would certainly help expand on what we currently know about the molecular mechanism underlying peroxisome division.

The exact nature of the distinctive Pex11p-mC-labeled rosette structures in *pex3*Δ and *pex19*Δ cells remains unknown. It is uncertain if these structures would remain intact through the fixation protocol used for electron microscopy so that their ultrastructure could be studied. Moreover, the limited number of selectable markers available for strain construction in *Y. lipolytica*, together with the limited number of antibodies available to interrogate *Y. lipolytica*, limits our ability to study the potential colocalization of the Pex11p-mC-labelled rosettes with elements of the ER markers.

Likewise, the precise localization of the *Y. lipolytica* Pex11 family proteins in *pex3*Δ, *pex16*Δ and *pex19*Δ cells, as well as the exact localization of peroxisomal membrane proteins in *pex11*Δ cells remain undetermined. We hypothesize that Pex11p, Pex11Cp and Pex11/25p are localized to the secretory system in *pex3*Δ, *pex16*Δ and *pex19*Δ cells. We also hypothesize that peroxisomal membrane proteins may be localized to the ER and/or may be found in pre-peroxisomal structures in *pex11*Δ cells. One method to test for pre-peroxisomal structures would be to express the protein needed for peroxisome assembly (*i.e.* Pex11p) with a peroxisomal marker and study the localization of peroxisomal membrane proteins using confocal microscopy. The possible localization to the secretory system could be explored using an ER shift assay. This fractionation is based on the principle that in the presence of excess cation, the ER has increased buoyant density because of the increased association of ribosomes and the Sec61 complex (via salt bridges), and ER markers are therefore found in heavier fractions (Roberg et al., 1997). The ER shift assay consists of a fractionation followed by a discontinuous gradient utilizing homogenization and gradient buffers with cation (Mg^{2+}) or no cation (EDTA) (Mast, 2013). Possible ER markers could include the *Y. lipolytica* BiP protein, which can be detected using an antibody to its *S. cerevisiae* homologue Kar2. Proteins in the same ER compartment should cofractionate regardless of the ER distribution pattern. Subcellular fractionation could be used to investigate this question, but it is highly unlikely that

fluorescently tagged Pex11 family proteins and an ER marker protein would fractionate identically, thus leaving the question unresolved.

The most intriguing result of this study is the classical *pex* phenotype, *i.e.* failure to assemble peroxisomes, observed in *Y. lipolytica* cells deleted for *PEX11*. In *H. polymorpha* cells, Pex25p was required for the re-introduction of peroxisomes when Pex3p was expressed in *pex3*Δ/*pex25*Δ cells, but not in *pex3*Δ/*pex11*Δ cells. It would be interesting to determine if Pex11p is similarily required for peroxisome re-introduction in *Y. lipolytica pex3*Δ/*pex11*Δ cells. Do Pex11Cp and Pex11/25 have unique functions in peroxisome biogenesis in *Y. lipolytca*? Is a *pex11C*Δ/*pex11/25*Δ double deletion strain affected in peroxisome size and number? Our data suggest that Pex11Cp and Pex11/25p act to maintain peroxisome proliferation in an actively proliferating peroxisome population, and leaves open the question of what maintains the peroxisome population in a constitutively dividing *Y. lipoytica* cell. In closing, given the considerable number of fungal genomes that encode these and other Pex11 family proteins, further investigation will hopefully elucidate and clarify the functions of this family of proteins in peroxisomal biogenesis in *Y. lipolytica* and other eukaryotes.

5.3 Future directions for the PTS Finder program

Although a useful bioinformatic tool, there are several ways that PTS Finder could be improved and developed further. PTS Finder does not penalize deviations from the consensus sequence, which would make it more complex than a profile search alone (Neuberger et al., 2003b). PTS Finder does not incorporate a machine learning algorithm (*e.g.* neural networks or support vector machines) such as those incorporated into the PTS1 predictor PeroxiP (Emanuelsson et al., 2003). PeroxiP also makes use of TargetP (Emanuelsson et al., 2000) and TMHMM (cbs.dtu.dk/services/TMHMM-2.0) to remove secreted proteins and transmembrane domain-containing proteins, respectively, neither of which will be targeted to the peroxisomal matrix. PTS Finder does not search for a PTS1 dodecamer, *i.e.* the amino acids upstream of the tripeptide

consensus sequence are not considered (Neuberger et al., 2003a; Brocard and Hartig, 2006). We chose to exclude these criteria from this version of the program as there was no consensus sequence given for these amino acids and they were only studied in mammalian, fungal and plant peroxisomal matrix proteins. Obviously, as the consensus PTS1 and PTS2 sequences are updated, these changes can easily be written into future versions of the program. Finally, PTS Finder should be tested for accuracy in predicting peroxisomal proteins, and possibly benchmarked against other available programs

5.4 Future directions for the study of peroxisomes in *Blastocystis hominis*

Based on the bioinformatic analysis presented in this thesis and a lack of ultrastructural evidence to date, I conclude that the peroxisome is likely absent in the parasite *Blastocystsis.* Additional comparative genomic searches for other peroxisomal proteins other than those encoded by *PEX* genes could also be done. Does *Blastocystis* encode homologues of any major families of enzymes typically localized to peroxisomes? Are PTSs predicted for any *Blastocystis* proteins? One possibility is that peroxisomal enzymes have been retargeted to another organelle such as the mitochondrion. If homologoues of peroxisomal enzymes are in fact identified in *Blastocystis*, TargetP could be used to determine if they contain cannonical mitochondrial targeting peptides (Emanuelsson et al., 2000). If there are established protocols for imaging *Blastocystis* cells, candidate proteins could be fluorescently tagged and localized using a mitochondrial marker such as MitoTracker.

Another outcome of this project was a considerable expansion of our knowledge of peroxins in the SAR clade. There are many more SAR genomes available than those covered in this thesis. Additionally, the presence/absence of peroxisomes has been completely unresolved for the closest supergroup to the SAR clade, the CCTH. Recently completed genomes from the rhizarian *Bigelowiella natans*, the haptophyte *Emiliania huxleyi*, and the cryptomonad *Guillardia theta* could be important taxonomic sampling points to address this question.

5.5 Future directions for the study of peroxisomes in *Bodo saltans*

B. saltans is an important trypanosomatid outgroup because it is non-parasitic, although it has glycosomes. A complete proteomic analysis, either bioinformatical or experimental, with the genomes surveyed here will aid our understanding of the transition to parasitism. I identified *B. saltans* proteins that are putatively targeted to the glycosome (Appendix 15). Proteins of interest could be tagged and localized. Previous phylogenetic analyses revealed that *T. brucei* glycolysis enzymes group with plant sequences (Hannaert et al., 2003). The authors hypothesize that the ancestor of euglenids and trypanosomes may have taken up an algal endosymbiont, and while *Euglena* sp. retains a chloroplast, genes encoded by the chloroplast may have been transferred to the nucleus in organisms such as *T. brucei* (Hannaert et al., 2003). Therefore, whether or not the glycolysis enzymes in *Bodo saltans*, *N. gruberi* and *N. fowleri* also group with plants in phylogenetic analyses is an interesting question. There is a massive amount of genome sequencing data available for multiple trypanosomatid genomes, which can hopefully assist in identifying factors unique to the glycosome. More specifically, one immediate area of interest would be a comparative genomic survey of peroxins and *T. brucei* high confidence glycosomal proteins in other trypanosome genomes.

5.6 Future directions for the study of peroxisomes in *Naegleria fowleri*

Naegleria gruberi is thought of as a canonical model organism for the Excavata supergroup, but our comparative genomic analysis suggests peroxisomal targeting is highly reduced in *N. gruberi*, but not in *N. fowleri*. Ultrastructural studies providing or not providing evidence of peroxisomes in these two organisms would be very informative. The *N. gruberi* and *N. fowleri* genomes appear similar in terms of peroxin complement. We did not search for homologues of high-confidence *T. brucei* glycosomal proteins (with the exception of glycolysis enzymes) in *N. gruberi* and N. *fowleri*. Are these two genomes similar in this regard as well?

5.7 Conclusion

It is intriguing to speculate on whether peroxisomes can be predicted to be present or not in an organism simply from *in silico* data alone. Many aspects of peroxisome dynamics are still poorly understood, but we do know that different peroxins are essential and dispensable for peroxisome biogenesis in diverse eukaryotes. The study of diverse eukaryotes, particularly those lacking morphological or biochemical evidence of peroxisomes, raises the question, "What is the minimum complement of proteins required for a functional peroxisome?". The answer is probably not the same for every eukaryote and for each diverse type of peroxisome.

Firstly, addressing this question would likely begin with a new comparative genomic survey, similar to previous work (Schülter et al., 2006; Gabaldón et al., 2006). Ideally, this would combine the two approaches taken in this thesis: searching the breadth of diverse eukaryotic genomes used for the comparative genomic survey of the Pex11 family proteins for all of the peroxins. Importantly, our comparative genomic survey of the Pex11 family proteins built on previous work by demonstrating that Pex11p is an ancestral protein. Gabaldón and colleagues did not consider Pex11p to be part of the ancestral eukaryotic peroxisomal proteome complement (which included Pex1p, Pex2p, Pex4p, Pex5p, Pex10p and Pex14p) (Gabaldón et al., 2006). Likewise, Schülter and colleagues identify four peroxisomal markers (Pex3p, Pex10p, Pex12p and Pex19p), but do not recognize Pex11p in this list (Schülter et al., 2006). Results from our comparative genomic survey showed that Pex11p fits the definition of a peroxisomal marker: present in all organisms with peroxisomes and absent from all parasitic genomes without peroxisomes (taking into consideration that peroxisome presence/absence has not been established for a vast number of the genomes queried in this study). This result shows that repeating previous comparative genomic surveys of peroxins can be informative. Further study of genomes from additional different and diverse eukaryotes will help to construct an updated minimum peroxisomal protein complement.

The results of such a comparative genomic survey could guide an experimental approach to addressing this question. Starting with a model system with functional peroxisomes, *PEX* genes

could be deleted sequentially to see if there are any components that are non-essential. This approach could be extended to proteins other than peroxins as well, as peroxisomes were initially defined by enzymatic content alone. Conversely, can peroxisomes be induced or formed in an organism for which wild-type cells do not contain functional peroxisomes? This is obviously a less tractable approach experimentally. Additionally, it would be interesting to undertake this experiment in a parasite without morphologically identifiable peroxisomes. Protocols for culturing many parasites are already established. However, the precise reason(s) for why this organelle is not present in the majority organisms remains undetermined.

References

Abascal, F., R. Zardoya, and D. Posada. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics*. 21:2104–2105.

Abrahamsen, M.S., T.J. Templeton, S. Enomoto, J.E. Abrahante, G. Zhu, C.A. Lancto, M. Deng, C. Liu, G. Widmer, S. Tzipori, G.A. Buck, P. Xu, A.T. Bankier, P.H. Dear, B.A. Konfortov, H.F. Spriggs, L. Iyer, V. Anantharaman, L. Aravind, and V. Kapur. 2004. Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science*. 304:441–445.

Adams, M.D., S.E. Celniker, R.A. Holt, C.A. Evans, J.D. Gocayne, P.G. Amanatides, S.E. Scherer, P.W. Li, R.A. Hoskins, R.F. Galle, R.A. George, S.E. Lewis, S. Richards, M. Ashburner, S.N. Henderson, G.G. Sutton, J.R. Wortman, M.D. Yandell, Q. Zhang, L.X. Chen, R.C. Brandon, Y.-H.C. Rogers, R.G. Blazej, M. Champe, B.D. Pfeiffer, K.H. Wan, C. Doyle, E.G. Baxter, G. Helt, C.R. Nelson, G.L.G. Miklos, J.F. Abril, A. Agbayani, H.-J. An, C. Andrews-Pfannkoch, D. Baldwin, R.M. Ballew, A. Basu, J. Baxendale, L. Bayraktaroglu, E.M. Beasley, K.Y. Beeson, P.V. Benos, B.P. Berman, D. Bhandari, S. Bolshakov, D. Borkova, M.R. Botchan, J. Bouck, P. Brokstein, P. Brottier, K.C. Burtis, D.A. Busam, H. Butler, E. Cadieu, Angela Center, I. Chandra, J.M. Cherry, S. Cawley, C. Dahlke, L.B. Davenport, P. Davies, B. de Pablos, A. Delcher, Z. Deng, A.D. Mays, I. Dew, S.M. Dietz, K. Dodson, L.E. Doup, M. Downes, S. Dugan-Rocha, B.C. Dunkov, P. Dunn, K.J. Durbin, C.C. Evangelista, C. Ferraz, S. Ferriera, W. Fleischmann, C. Fosler, A.E. Gabrielian, N.S. Garg, W.M. Gelbart, K. Glasser, A. Glodek, F. Gong, J.H. Gorrell, Z. Gu, P. Guan, M. Harris, N.L. Harris, D. Harvey, T.J. Heiman, J.R. Hernandez, J. Houck, D. Hostin, K.A. Houston, T.J. Howland, et al. 2000. The genome sequence of *Drosophila melanogaster*. *Science*. 287:2185–2195.

Adl, S.M., A.G.B. Simpson, C.E. Lane, J. Lukes, D. Bass, S.S. Bowser, M.W. Brown, F. Burki, M. Dunthorn, V. Hampl, A. Heiss, M. Hoppenrath, E. Lara, L.L. Gall, D.H. Lynn, H. Mcmanus, E.A.D. Mitchell, S. Mozley-Standridge, L. Parfrey, J. Pawlowski, S. Rueckert, L. Shadwick, C.L. Schoch, A. Smirnov, and F.W. Spiegel. 2012. The revised classification of eukaryotes. *J. Eukaryot. Microbiol.* 59:429–493.

Ahlemeyer, B., M. Gottwald, and E. Baumgart-Vogt. 2011. Deletion of a single allele of the *PEX11* gene is sufficient to cause oxidative stress, delayed differentiation and neuronal death in mouse brain. *Dis. Models Mech*. 5:125–140.

Aime, M.C., P.B. Matheny, D.A. Henk, E.M. Frieders, R.H. Nilsson, M. Piepenbring, D.J. McLaughlin, L.J. Szabo, D. Begerow, J.P. Sampaio, R. Bauer, M. Weiss, F. Oberwinkler, and D.S. Hibbett. 2006. An overview of the higher level classification of Pucciniomycotina based on combined analyses of nuclear large and small subunit rDNA sequences. Mycologia. 98:896–905.Aitchison, J.D., R.K. Szilard, W.M. Nuttley, and R.A. Rachubinski. 1992. Antibodies directed against a yeast carboxyl-terminal peroxisomal targeting signal specifically recognize peroxisomal proteins from various yeasts. *Yeast*. 8:721–734.

al-Tawil, Y.S., M.A. Gilger, G.S. Gopalakrishna, C. Langston, and K.E. Bommer. 1994. Invasive *Blastocystis hominis* infection in a child. *Arch Pediatr Adolesc Med*. 148:882-885.

Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nuc. Acids Res*. 25:3389–3402.

Andrade-Navarro, M.A., L. Sanchez-Pulido, and H.M. McBride. 2009. Mitochondrial vesicles: an ancient process providing new links to peroxisomes. *Curr. Opin. Cell Biol*. 21:560–567.

Antonenkov, V.D., and J.K. Hiltunen. 2006. Peroxisomal membrane permeability and solute transfer. *Biochim. Biophys. Acta.* 1763:1697-1706.

Archibald, J.M. 2007. Nucleomorph genomes: structure, function, origin, and evolution. *BioEssays*. 29:392–402.

Arisue, N., T. Hashimoto, H. Yoshikawa, Y. Nakamura, G. Nakamura, F. Nakamura, T.-A. Yano, and M. Hasegawa. 2002. Phylogenetic position of *Blastocystis hominis* and of stramenopiles inferred from multiple molecular sequence data. *J. Eukaryot. Microbiol.* 49:42–53.

Armbrust, E.V., J.A. Berges, C. Bowler, B.R. Green, D. Martinez, N.H. Putnam, S. Zhou, A.E. Allen, K.E. Apt, M. Bechner, M.A. Brzezinski, B.K. Chaal, A. Chiovitti, A.K. Davis, M.S. Demarest, J.C. Detter, T. Glavina, D. Goodstein, M.Z. Hadi, U. Hellsten, M. Hildebrand, B.D. Jenkins, J. Jurka, V.V. Kapitonov, N. Kroger, W.W.Y. Lau, T.W. Lane, F.W. Larimer, J.C. Lippmeier, S. Lucas, M. Medina, A. Montsant, M. Oborník, M.S. Parker, B. Palenik, G.J. Pazour, P.M. Richardson, T.A. Rynearson, M.A. Saito, D.C. Schwartz, K. Thamatrakoln, K. Valentin, A. Vardi, F.P. Wilkerson, and D.S. Rokhsar. 2004. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science*. 306:79–86.

Aslett, M., Aurora, M. Berriman, J. Brestelli, B.P. Brunk, M. Carrington, D.P. Depledge, S. Fischer, B. Gajria, X. Gao, M.J. Gardner, A. Gingle, G. Grant, O.S. Harb, M. Heiges, C. Hertz-Fowler, R. Houston, F. Innamorato, J. Iodice, J.C. Kissinger, E. Kraemer, W. Li, F.J. Logan, J.A. Miller, S. Mitra, P.J. Myler, V. Nayak, C. Pennington, I. Phan, D.F. Pinney, G. Ramasamy, M.B. Rogers, D.S. Roos, C. Ross, D. Sivam, D.F. Smith, G. Srinivasamoorthy, C.J. Stoeckert, S. Subramanian, R. Thibodeau, A. Tivey, C. Treatman, G. Velarde, and H. Wang. 2009. TriTrypDB: a functional genomic resource for the Trypanosomatidae. *Nuc. Acids Res*. 38:D457–D462.

Aurora, J. Brestelli, B.P. Brunk, J.M. Carlton, J. Dommer, S. Fischer, B. Gajria, X. Gao, A. Gingle, G. Grant, O.S. Harb, M. Heiges, F. Innamorato, J. Iodice, J.C. Kissinger, E. Kraemer, W. Li, J.A. Miller, H.G. Morrison, V. Nayak, C. Pennington, D.F. Pinney, D.S. Roos, C. Ross, C.J. Stoeckert, S. Sullivan, C. Treatman, and H. Wang. 2009. GiardiaDB and TrichDB: integrated genomic resources for the eukaryotic protist pathogens *Giardia lamblia* and *Trichomonas vaginalis*. *Nuc. Acids Res*. 37:D526–D530.

Ausubel, F.J., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl. 1989. Current Protocols in Molecular Biology. Greene Publishing Associates, New York, NY.

Bagattin, A., L. Hugendubler, and E. Mueller. 2010. Transcriptional coactivator PGC-1alpha promotes peroxisomal remodeling and biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 107:20376–20381.

Barnett, P., H.F. Tabak, and E.H. Hettema. 2000. Nuclear receptors arose from pre-existing protein modules during evolution. *Trends Biochem Sci*. 227–228.

Barth, G., and C. Gaillardin. 1996. *Yarrowia lipolytica*. *In* Non-Conventional Yeasts in Biotechnology. K. Wolf, editor. Springer. 313–388.

Baudhuin, P., H. Beaufay, and C. de Duve. 1965. Combined biochemical and morphological study of particulate fractions from rat liver. *J. Cell Biol*. 26:219–243.

Berka, R.M., I.V. Grigoriev, R. Otillar, A. Salamov, J. Grimwood, I. Reid, N. Ishmael, T. John, C. Darmond, M.-C. Moisan, B. Henrissat, P.M. Coutinho, V. Lombard, D.O. Natvig, E. Lindquist, J.

Schmutz, S. Lucas, P. Harris, J. Powlowski, A. Bellemare, D. Taylor, G. Butler, R.P. de Vries, I.E. Allijn, J. van den Brink, S. Ushinsky, R. Storms, A.J. Powell, I.T. Paulsen, L.D.H. Elbourne, S.E. Baker, J. Magnuson, S. LaBoissiere, A.J. Clutterbuck, D. Martinez, M. Wogulis, A.L. de Leon, M.W. Rey, and A. Tsang. 2011. Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. *Nature Biotechnol*. 29:922–927.

Berriman, M., E. Ghedin, C. Hertz-Fowler, G. Blandin, H. Renauld, D.C. Bartholomeu, N.J. Lennard, E. Caler, N.E. Hamlin, B. Haas, B. Ulrike, L. Hannick, M.A. Aslett, J. Shallom, L. Marcello, L. Hou, B. Wickstead, U.C.M. Alsmark, C. Arrowsmith, R.J. Atkin, A.J. Barron, F. Bringaud, K. Brooks, M. Carrington, I. Cherevach, T.-J. Chillingworth, C. Churcher, L.N. Clark, C.H. Corton, A. Cronin, R.M. Davies, J. Doggett, A. Djikeng, T. Feldblyum, M.C. Field, A. Fraser, I. Goodhead, Z. Hance, D. Harper, B.R. Harris, H. Hauser, J. Hostetler, Al Ivens, K. Jagels, D. Johnson, J. Johnson, K. Jones, A.X. Kerhornou, H. Koo, N. Larke, S. Landfear, C. Larkin, V. Leech, A. Line, A. Lord, A. MacLeod, P.J. Mooney, S. Moule, D.M.A. Martin, G.W. Morgan, K. Mungall, H. Norbertczak, D. Ormond, G. Pai, C.S. Peacock, J. Peterson, M.A. Quail, E. Rabbinowitsch, M.-A. Rajandream, C. Reitter, S.L. Salzberg, M. Sanders, S. Schobel, S. Sharp, M. Simmonds, A.J. Simpson, L. Tallon, C.M.R. Turner, A. Tait, A.R. Tivey, S. Van Aken, D. Walker, D. Wanless, S. Wang, B. White, O. White, S. Whitehead, J. Woodward, J. Wortman, M.D. Adams, T.M. Embley, K. Gull, E. Ullu, J.D. Barry, A.H. Fairlamb, F. Opperdoes, B.G. Barrell, J.E. Donelson, et al. 2005. The genome of the African trypanosome *Trypanosoma brucei*. *Science*. 309:416–422.

Blom, N., S. Gammeltoft, and S. Brunak. 1999. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol*. 294:1351–1362.

Boehm, E.W.A., G.K. Mugambi, A.N. Miller, S.M. Huhndorf, S. Marincowitz, J.W. Spatafora, and C.L. Schoch. 2010. A molecular phylogenetic reappraisal of the Hysteriaceae, Mytilinidiaceae and Gloniaceae (Pleosporomycetidae, Dothideomycetes) with keys to world species. *Stud. Mycol*. 64:49–83.

Boorom, K.F., H. Smith, L. Nimri, E. Viscogliosi, G. Spanakos, U. Parkar, L.H. Li, X.N. Zhou, U.Z. Ok, S. Leelayoova, and M.S. Jones. 2008. Oh my aching gut: irritable bowel syndrome, *Blastocystis*, and asymptomatic infection. *Parasit. Vectors*. 1:40.

Bowen, P., C.S.N. Lee, H. Zellweger, and R. Lindenberg. 1964. A familial syndrome of multiple congenital defects. *Bull. Johns Hopkins Hosp*. 114:402-414.

Brade, A.M. 1992. Peroxisome Assembly in *Yarrowia lipolytica*. McMaster University, Hamilton.

Bradford, B.U., C.B. Seed, J.A. Handler, D.T. Forman, and R.G. Thurman. 1993. Evidence that catalase is a major pathway of ethanol oxidation in vivo: dose-response studies in deer mice using methanol as a selective substrate. *Arch. Biochem. Biophys*. 303:172–176.

Braschi, E., V. Goyon, R. Zunino, A. Mohanty, L. Xu, and H.M. McBride. 2010. Vps35 mediates vesicle transport between the mitochondria and peroxisomes. *Curr. Biol*. 20:1310–1315.

Brocard, C., and A. Hartig. 2006. Peroxisome targeting signal 1: is it really a simple tripeptide? *Biochim. Biophys. Acta. Mol. Cell Res*. 1763:1565–1573.

Burnette, W.N. 1981. "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem*. 112:195–203.

Byrne, K.P., and K.H. Wolfe. 2005. The yeast gene order browser: combining curated homology and syntenic context reveals gene fatein polyploid species. *Genome Res.* 15:1456–1461.

C. elegans Sequencing Consortium. 1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science*. 282:2012–2018.

Carlton, J.M., R.P. Hirt, J.C. Silva, A.L. Delcher, M. Schatz, Q. Zhao, J.R. Wortman, S.L. Bidwell, U.C.M. Alsmark, S. Besteiro, T. Sicheritz-Ponten, C.J. Noel, J.B. Dacks, P.G. Foster, C. Simillion, Y. Van de Peer, D. Miranda-Saavedra, G.J. Barton, G.D. Westrop, S. Muller, D. Dessi, P.L. Fiori, Q. Ren, I. Paulsen, H. Zhang, F.D. Bastida-Corcuera, A. Simoes-Barbosa, M.T. Brown, R.D. Hayes, M. Mukherjee, C.Y. Okumura, R. Schneider, A.J. Smith, S. Vanacova, M. Villalvazo, B.J. Haas, M. Pertea, T.V. Feldblyum, T.R. Utterback, C.L. Shu, K. Osoegawa, P.J. de Jong, I. Hrdy, L. Horvathova, Z. Zubacova, P. Dolezal, S.B. Malik, J.M. Logsdon, K. Henze, A. Gupta, C.C. Wang, R.L. Dunne, J.A. Upcroft, P. Upcroft, O. White, S.L. Salzberg, P. Tang, C.H. Chiu, Y.S. Lee, T.M. Embley, G.H. Coombs, J.C. Mottram, J. Tachezy, C.M. Fraser-Liggett, and P.J. Johnson. 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science*. 315:207–212.

Carter, R.F. 1972. Primary amoebic meningo-encephalitis. An apprasial of present knowledge. *Trans. R. Soc. Trop. Med. Hyg.* 66:193–213.

Chang, J., F.D. Mast, A. Fagarasanu, D.A. Rachubinski, G.A. Eitzen, J.B. Dacks, and R.A. Rachubinski. 2009. Pex3 peroxisome biogenesis proteins function in peroxisome inheritance as class V myosin receptors. *J. Cell Biol.* 187:233–246.

Chang, J., R.J. Tower, D.L. Lancaster, and R.A. Rachubinski. 2013. Dynein light chain interaction with the peroxisomal import docking complex modulates peroxisome biogenesis in yeast. *J. Cell Sci.* 126:4698–4706.

Cherry, J.M., C. Adler, C. Ball, S.A. Chervitz, S.S. Dwight, E.T. Hester, Y. Jia, G. Juvik, T. Roe, M. Schroeder, S. Weng, and D. Botstein. 1997. SGD: *Saccharomyces* genome database. *Nuc. Acids Res.* 26:73–79.

Chowdhary, G., A.R. Kataya, T. Lingner, and S. Reumann. 2012. Non-canonical peroxisome targeting signals: identification of novel PTS1 tripeptides and characterization of enhancer elements by computational permutation analysis. *BMC Plant Biol.* 12:1–14.

Clarke, M., A.J. Lohan, B. Liu, I. Lagkouvardos, S. Roy, N. Zafar, C. Bertelli, C. Schilde, A. Kianianmomeni, T.R. Bürglin, C. Frech, B. Turcotte, K.O. Kopec, J.M. Synnott, C. Choo, I. Paponov, A. Finkler, C.S.H. Tan, A.P. Hutchins, T. Weinmeier, T. Rattei, J.S. Chu, G. Gimenez, M. Irimia, D.J. Rigden, D.A. Fitzpatrick, J. Lorenzo-Morales, A. Bateman, C.-H. Chiu, P. Tang, P. Hegemann, H. Fromm, D. Raoult, G. Greub, D. Miranda-Saavedra, N. Chen, P. Nash, M.L. Ginger, M. Horn, P. Schaap, L. Caler, and B.J. Loftus. 2013. Genome of *Acanthamoeba castellanii* highlights extensive lateral gene transfer and earlyevolution of tyrosine kinase signaling. *Genome Biol.* 14:R11.

Cliften, P., P. Sudarsanam, A. Desikan, L. Fulton, B. Fulton, J. Majors, R. Waterston, B.A. Cohen, and M. Johnston. 2003. Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting. *Science*. 301:71–76.

Cock, J.M., L. Sterck, P. Rouze, D. Scornet, A.E. Allen, G. Amoutzias, V. Anthouard, F. Artiguenave, J.-M. Aury, J.H. Badger, B. Beszteri, K. Billiau, E. Bonnet, J.H. Bothwell, C. Bowler, C. Boyen, C.

Brownlee, C.J. Carrano, B. Charrier, G.Y. Cho, S.M. Coelho, J. Collén, E. Corre, C. Da Silva, L. Delage, N. Delaroque, S.M. Dittami, S. Doulbeau, M. Elias, G. Farnham, C.M.M. Gachon, B. Gschloessl, S. Heesch, K. Jabbari, C. Jubin, H. Kawai, K. Kimura, B. Kloareg, F.C. Küpper, D. Lang, A. Le Bail, C. Leblanc, P. Lerouge, M. Lohr, P.J. Lopez, C. Martens, F. Maumus, G. Michel, D. Miranda-Saavedra, J. Morales, H. Moreau, T. Motomura, C. Nagasato, C.A. Napoli, D.R. Nelson, P. Nyvall-Collén, A.F. Peters, C. Pommier, P. Potin, J. Poulain, H. Quesneville, B. Read, S.A. Rensing, A. Ritter, S. Rousvoal, M. Samanta, G. Samson, D.C. Schroeder, B. Ségurens, M. Strittmatter, T. Tonon, J.W. Tregear, K. Valentin, P. von Dassow, T. Yamagishi, Y. Van de Peer, and P. Wincker. 2010. The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. *Nature*. 465:617–621.

Curtis, B.A., G. Tanifuji, F. Burki, A. Gruber, M. Irimia, S. Maruyama, M.C. Arias, S.G. Ball, G.H. Gile, Y. Hirakawa, J.F. Hopkins, A. Kuo, S.A. Rensing, J. Schmutz, A. Symeonidi, M. Elias, R.J.M. Eveleigh, E.K. Herman, M.J. Klute, T. Nakayama, M. Oborník, A. Reyes-Prieto, E.V. Armbrust, S.J. Aves, R.G. Beiko, P. Coutinho, J.B. Dacks, D.G. Durnford, N.M. Fast, B.R. Green, C.J. Grisdale, F. Hempel, B. Henrissat, M.P. Höppner, K.-I. Ishida, E. Kim, L. Kořený, P.G. Kroth, Y. Liu, S.-B. Malik, U.-G. Maier, D. McRose, T. Mock, J.A.D. Neilson, N.T. Onodera, A.M. Poole, E.J. Pritham, T.A. Richards, G. Rocap, S.W. Roy, C. Sarai, S. Schaack, S. Shirato, C.H. Slamovits, D.F. Spencer, S. Suzuki, A.Z. Worden, S. Zauner, K. Barry, C. Bell, A.K. Bharti, J.A. Crow, J. Grimwood, R. Kramer, E. Lindquist, S. Lucas, A. Salamov, G.I. McFadden, C.E. Lane, P.J. Keeling, M.W. Gray, I.V. Grigoriev, and J.M. Archibald. 2012. Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature*. 492:59–65.

Davidson, R.C., J.R. Blakenshop, P.R. Krauss, M. de Jesus Berrios, C.M. Hull, C. D'Souza, P. Wang, and J. Heitman. 2002. A PCR-based strategy to generate integrative targeting alleles with large regions of homology. *Microbiology*. 148:2607–2615.

de Duve, C. 2007. The origin of eukaryotes: a reappraisal. *Nat. Rev. Genet*. 8:395–403.

de Duve, C., and P. Baudhuin. 1966. Peroxisomes (microbodies and related particles). *Physiol. Rev*. 46:323–357.

de Duve, C., H. Beaufay, P. Jacques, Y. Rahman-Li, O.Z. Sellinger, R. Wattiaux, and S. de Connick. 1960. Intracellular localization of catalase and of some oxidases in rat liver. *Biochim. Biophys. Acta*. 40:186–187.

De Schutter, K., Y.-C. Lin, P. Tiels, A. Van Hecke, S. Glinka, J. Weber-Lehmann, P. Rouze, Y. Van de Peer, and N. Callewaert. 2009. Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat. Biotechnol*. 27:561–566.

de Souza, W., A. Lanfredi-Rangel, and L. Campanati. 2004. Contribution of microscopy to a better knowledge of the biology of *Giardia lamblia*. *Microsc. Microanal*. 10:513–527.

Delille, H.K., B. Agricola, S.C. Guimaraes, H. Borta, G.H. Luers, M. Fransen, and M. Schrader. 2010. Pex11pbeta-mediated growth and division of mammalian peroxisomes follows a maturation pathway. *J. Cell Sci*. 123:2750–2762.

Denoeud, F., M. Roussel, B. Noel, I. Wawrzyniak, C. Da Silva, M. Diogon, E. Viscogliosi, C. Brochier-Armanet, A. Couloux, J. Poulain, B. Ségurens, V. Anthouard, C. Texier, N. Blot, P. Poirier, G.C. Ng, K.S. Tan, F. Artiguenave, O. Jaillon, J.-M. Aury, F. Delbac, P. Wincker, C.P. Vivares, and H. El Alaoui. 2011. Genome sequence of the stramenopile *Blastocystis*, a human anaerobic parasite. *Genome Biol.* 12:R29.

Desai, M., and J. Hu. 2008. Light induces peroxisome proliferation in *Arabidopsis* seedlings through the photoreceptor phytochrome A, the transcription factor HY5 homolog, and the peroxisomal protein PEROXIN11b. *Plant Physiol.* 146:1117-1127.

Dibrov, E., S. Fu, and B.D. Lemire. 1998. The *Saccharomyces cerevisiae TCM62* gene encodes a chaperone necessary for the assembly of the mitochondrial succinate dehydrogenase (Complex II)*. *J. Biol. Chem.* 273:32042–32048.

Dietrich, F.S., S. Voegeli, S. Brachat, A. Lerch, K. Gates, S. Steiner, C. Mohr, R. Pohlmann, P. Luedi, S. Choi, R.A. Wing, A. Flavier, T.D. Gaffney, and P.T. Philippsen. 2004. The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. *Science*. 304-307:1–5.

Ding, M., C. Clayton, and D. Soldati. 2000. *Toxoplasma gondii* catalase: are there peroxisomes in *Toxoplasma*? *J. Cell Sci.* 113:2409–2419.

Distel, B., R. Erdmann, S.J. Gould, G. Blobel, D.I. Crane, J.M. Cregg, G. Dodt, Y. Fujiki, J.M. Goodman, W.W. Just, J.A.K.W. Kiel, W.-H. Kunau, P.B. Lazarow, G.P. Mannaerts, H.W. Moser, T. Osumi, R.A. Rachubinski, A. Roscher, S. Subramani, H.F. Tabak, T. Tsukamoto, D. Valle, I. van der Klei, P.P. van Veldhoven, and M. Veenhuis. 1996. A unified nomenclature for peroxisome biogenesis factors. *J. Cell. Biol.* 135:1–3.

Dodt, G., D. Warren, E. Becker, P. Rehling, and S.J. Gould. 2001. Domain mapping of human PEX5 reveals functional and structural similarities to *S. cerevisiae* Pex18p and Pex21p. *J. Cell Biol*. 276:41769–41781.

Dujon, B., D. Sherman, G. Fishcher, P. Durrens, S. Casaregola, I. Lafontaine, J. de Montigny, C. Marck, C. Neuveglise, E. Talla, N. Goffard, L. Frangeul, M. Aigle, V. Anthouard, A. Babour, V. Barbe, S. Barnay, S. Blanchin, J.-M. Beckerich, E. Beyne, C. Bleykasten, A. Boisrame, J. Boyer, L. Cattolico, F. Confanioleri, A. de Daruvar, L. Despons, E. Fabre, C. Fairhead, H. Ferry-Dumazet, A. Groppi, F. Hantraye, C. Hennequin, N. Jauniaux, P. Joyet, R. Kachouri, A. Kerrest, R. Koszul, M. Lemaire, I. Lesur, H. Muller, J.-M. Nicaud, M. Nikolski, S. Oztas, O. Ozier-Kalogeropoulos, S. Pellenz, S. Potier, G.-F. Richard, M.-L. Straub, A. Suleau, D. Swennen, F. Tekaia, M. Wesolowski-Louvei, E. Westhof, B. Wirth, M. Zeniou-Meyer, I. Zivanovic, M. Bolotin-Fukuhara, A. Thierry, C. Bouchier, B. Caudron, C. Scarpelli, C. Gaillardin, J. Weissenbach, P. Wincker, and J.-L. Souclet. 2004. Genome evolution in yeasts. *Nature*. 430:35–44.

Durbin, R., S.R. Eddy, A. Krogh, and G. Mitchison. 1998. Biological sequence analysis: probabilistic models of proteins and nucleic acids. Cambridge University Press, Cambridge. 371 pp.

Eaton, S., K. Bartlett, and M. Pourfarzam. 1996. Mammalian mitochondrial beta-oxidation. *Biochem. J.* 320:345–357.

Ebberink, M.S., J. Koster, G. Visser, F.V. Spronsen, I. Stolte-Dijkstra, G.P.A. Smit, J.M. Fock, S. Kemp, R.J.A. Wanders, and H.R. Waterham. 2012. A novel defect of peroxisome division due to a homozygous non-sense mutation in the *PEX11* gene. *J. Med. Genet.* 49:307–313.

Ebersberger, I., R. de Matos Simoes, A. Kupczok, M. Gube, E. Kothe, K. Voigt, and A. von Haeseler. 2012. A Consistent Phylogenetic Backbone for the Fungi. *Mol. Biol. Evol*. 29:1319–1334.

Eddy, S.R. 1998. Profile hidden Markov models. *Bioinformatics*. 14:755–763.

Eddy, S.R. 2009. A new generation of homology serch tools based on probablistic inference. *Genome Inform*. 23:1–7.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32:1792–1797.

Eichinger, L., J.A. Pachebat, G. Glockner, M.A. Rajandream, R. Sucgang, M. Berriman, J. Song, R. Olsen, K. Szafranski, Q. Xu, B. Tunggal, S. Kummerfeld, M. Madera, B.A. Konfortov, F. Rivero, A.T. Bankier, R. Lehmann, N. Hamlin, R. Davies, P. Gaudet, P. Fey, K. Pilcher, G. Chen, D. Saunders, E. Sodergren, P. Davis, A. Kerhornou, X. Nie, N. Hall, C. Anjard, L. Hemphill, N. Bason, P. Farbrother, B. Desany, E. Just, T. Morio, R. Rost, C. Churcher, J. Cooper, S. Haydock, N. van Driessche, A. Cronin, I. Goodhead, D. Muzny, T. Mourier, A. Pain, M. Lu, D. Harper, R. Lindsay, H. Hauser, K. James, M. Quiles, M.M. Babu, T. Saito, C. Buchrieser, A. Wardroper, M. Felder, M. Thangavelu, D. Johnson, A. Knights, H. Loulseged, K. Mungall, K. Oliver, C. Price, M.A. Quail, H. Urushihara, J. Hernandez, E. Rabbinowitsch, D. Steffen, M. Sanders, Y. Kohara, S. Sharp, M. Simmonds, S. Spiegler, A. Tivey, S. Sugano, B. White, D. Walker, J. Woodward, T. Winckler, Y. Tanaka, G. Shaulsky, M. Schleicher, G. Weinstock, A. Rosenthal, E.C. Cox, R.L. Chisholm, R. Gibbs, W.F. Loomis, M. Platzer, R.R. Kay, J. Williams, P.H. Dear, A.A. Noegel, B. Barrell, and A. Kuspa. 2005. The genome of the social amoeba *Dictyostelium discoideum*. *Nature*. 435:43–57.

Einerhand, A.W.C., W.T. Kos, Ben Distal, and H.F. Tabak. 1993. Characterization of a transcriptional control element involved in proliferation of peroxisomes in yeast in response to oleate. *Eur J Biochem*. 214:323–331.

Eisen, J.A., R.S. Coyne, M. Wu, D. Wu, M. Thiagarajan, J.R. Wortman, J.H. Badger, Q. Ren, P. Amedeo, K.M. Jones, L.J. Tallon, A.L. Delcher, S.L. Salzberg, J.C. Silva, B.J. Haas, W.H. Majoros, M. Farzad, J.M. Carlton, R.K. Smith, J. Garg, R.E. Pearlman, K.M. Karrer, L. Sun, G. Manning, N.C. Elde, A.P. Turkewitz, D.J. Asai, D.E. Wilkes, Y. Wang, H. Cai, K. Collins, B.A. Stewart, S.R. Lee, K. Wilamowska, Z. Weinberg, W.L. Ruzzo, D. Wloga, J. Gaertig, J. Frankel, C.-C. Tsao, M.A. Gorovsky, P.J. Keeling, R.F. Waller, N.J. Patron, J.M. Cherry, N.A. Stover, C.J. Krieger, C. del Toro, H.F. Ryder, S.C. Williamson, R.A. Barbeau, E.P. Hamilton, and E. Orias. 2006. Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. *PLOS Biol*.

Eitzen, G.A., R.K. Szilard, and R.A. Rachubinski. 1997. Enlarged peroxisomes are present in oleic acid–grown *Yarrowia lipolytica* overexpressing the *PEX16* gene encoding an intraperoxisomal peripheral membrane peroxin. *J. Cell Biol*. 137:1266–1278.

Eitzen, G.A., V.I. Titorenko, J.J. Smith, M. Veenhuis, R.K. Szilard, and R.A. Rachubinski. 1996. The *Yarrowia lipolytica* gene *PAY5* encodes a peroxisomal integral membrane protein homologous to the mammalian peroxisome assembly factor PAF-1*. *J. Biol. Chem*. 1–7.

Elgersma, Y., L. Kwast, M. van den Berg, W.B. Snyder, B. Distel, S. Subramani, and H.F. Tabak. 1997. Overexpression of Pex15p, a phosphorylated peroxisomal integral membrane protein required for peroxisome assembly in *S.cerevisiae*, causes proliferation of the endoplasmic reticulum membrane. *EMBO J*. 16:7326–7341.

El-Sayed, N.M., P.J. Myler, D.C. Bartholomeu, D. Nilsson, G. Aggarwal, A.-N. Tran, E. Ghedin, E.A. Worthey, A.L. Delcher, G. Blandin, S.J. Westenberger, E. Caler, G.C. Cerqueira, C. Branche, B. Haas, A. Anupama, E. Arner, L. Aslund, P. Attipoe, E. Bontempi, F. Bringaud, P. Burton, E. Cadag, D.A. Campbell, M. Carrington, J. Crabtree, H. Darban, J.F. da Silveira, P. de Jong, K.

Edwards, P.T. Englund, G. Fazelina, T. Feldblyum, M. Ferella, A.C. Frasch, K. Gull, D. Horn, L. Hou, Y. Huang, E. Kindlund, M. Klingbeil, S. Kluge, H. Koo, D. Lacerda, M.J. Levin, H. Lorenzi, T. Louie, C.R. Machado, R. McCulloch, A. McKenna, Y. Mizuno, J.C. Mottram, S. Nelson, S. Ochaya, K. Osoegawa, G. Pai, M. Parsons, M. Pentony, U. Pettersson, M. Pop, J.L. Ramirez, J. Rinta, L. Robertson, S.L. Salzberg, D.O. Sanchez, A. Seyler, R. Sharma, J. Shetty, A.J. Simpson, E. Sisk, M.T. Tammi, R. Tarleton, S. Teixeira, S. Van Aken, C. Vogt, P.N. Ward, B. Wickstead, J. Wortman, O. White, C.M. Fraser, K.D. Stuart, and B. Andersson. 2005. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science*. 309:409–415.

Emanuelsson, O., A. Elofsson, G. von Heijne, and S. Cristóbal. 2003. *In silico* prediction of the peroxisomal proteome in Fungi, plants and animals. *J. Mol. Biol*. 330:443–456.

Emanuelsson, O., H. Nielsen, S. Brunak, and G. von Heijne. 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol*. 300:1005-1016.

Erdmann, R., and G. Blobel. 1995. Giant peroxisomes in oleic acid-induced *Saccharomyces cerevisiae* lacking the peroxisomal membrane protein Pmp27p. *J. Cell. Biol*. 128:509–523.

Erdmann, R., M. Veenhuis, D. Mertens, and W.-H. Kunau. 1989. Isolation of peroxisome-deficient mutants of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci*. 86:5419–5423.

Fagarasanu, A., M. Fagarasanu, G.A. Eitzen, J.D. Aitchison, and R.A. Rachubinski. 2006. The peroxisomal membrane protein Inp2p is the peroxisome-specific receptor for the myosin V motor Myo2p of *Saccharomyces cerevisiae*. *Dev. Cell.* 10:587-600.

Fagarasanu, M., A. Fagarasanu, Y.Y.C. Tam, J.D. Aitchison, and R.A. Rachubinski. 2005. Inp1p is a peroxisomal membrane protein required for peroxisome inheritance in *Saccharomyces cerevisiae*. *J. Cell Biol.* 169:765-775.

Fagarasanu, A., F.D. Mast, B. Knoblach, and R.A. Rachubinski. 2010. Molecular mechanisms of organelle inheritance: lessons from peroxisomes in yeast. *Nat. Rev. Mol. Cell Biol*. 11:644–654.

Fagarasanu, A., F.D. Mast, B. Knoblach, Y. Jin, M.J. Brunner, M.R. Logan, J.N.M. Glover, G.A. Eitzen, J.D. Aitchison, L.S. Weisman, and R.A. Rachubinski. 2009. Myosin-driven peroxisome partitioning in *S. cerevisiae*. *J. Cell Biol*. 186:541–554.

Fagarasanu, A., M. Fagarasanu, and R.A. Rachubinski. 2007. Maintaining peroxisome populations: a story of division and inheritance. *Annu. Rev. Cell Dev. Biol.* 23:321–344.

Fang, Y., J.C. Morrell, J.M. Jones, and S.J. Gould. 2004. PEX3 functions as a PEX19 docking factor in the import of class I peroxisomal membrane proteins. *J. Cell. Biol*. 164:863-875.

Faust, J.E., A. Verma, C. Peng, and J.A. McNew. 2012. An inventory of peroxisomal proteins and pathways in *Drosophila melanogaster*. *Traffic*.

Filipits, M., M.M. Simon, W. Rapatz, B. Hamilton, and H. Ruis. 1993. A *Saccharomyces cerevisiae* upstream activating sequence mediates induction of peroxisome proliferation by fatty acids. *Gene*. 132:49–55.

Floudas, D., M. Binder, R. Riley, K. Barry, R.A. Blanchette, B. Henrissat, A.T. Martinez, R. Otillar, J.W. Spatafora, J.S. Yadav, A. Aerts, I. Benoit, A. Boyd, A. Carlson, A. Copeland, P.M. Coutinho, R.P. de Vries, P. Ferreira, K. Findley, B. Foster, J. Gaskell, D. Glotzer, P. Gorecki, J. Heitman, C. Hesse,

C. Hori, K. Igarashi, J.A. Jurgens, N. Kallen, P. Kersten, A. Kohler, U. Kues, T.K.A. Kumar, A. Kuo, K. LaButti, L.F. Larrondo, E. Lindquist, A. Ling, V. Lombard, S. Lucas, T. Lundell, R. Martin, D.J. McLaughlin, I. Morgenstern, E. Morin, C. Murat, L.G. Nagy, M. Nolan, R.A. Ohm, A. Patyshakuliyeva, A. Rokas, F.J. Ruiz-Duenas, G. Sabat, A. Salamov, M. Samejima, J. Schmutz, J.C. Slot, F. St John, J. Stenlid, H. Sun, S. Sun, K. Syed, A. Tsang, A. Wiebenga, D. Young, A. Pisabarro, D.C. Eastwood, F. Martin, D. Cullen, I.V. Grigoriev, and D.S. Hibbett. 2012. The paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science*. 336:1715–1719.

Freitag, J., J. Ast, and M. Bölker. 2012. Cryptic peroxisomal targeting via alternative splicing and stop codon read-through in fungi. *Nature*. 485:522–525.

Fritz-Laylin, L.K., S.E. Prochnik, M.L. Ginger, J.B. Dacks, M.L. Carpenter, M.C. Field, A. Kuo, A. Paredez, J. Chapman, J. Pham, S. Shu, R. Neupane, M. Cipriano, J. Mancuso, H. Tu, A. Salamov, E. Lindquist, H. Shapiro, S. Lucas, I.V. Grigoriev, W.Z. Cande, C. Fulton, D.S. Rokhsar, and S.C. Dawson. 2010. The genome of *Naegleria gruberi* illuminates early eukaryotic versatility. *Cell*. 140:631–642.

Gabaldón, T. 2010. Peroxisome diversity and evolution. *Philos. Trans. R. Soc. Lond., B., Biol. Sci.* 365:765–773.

Gabaldón, T. 2013. A metabolic scenario for the evolutionary origin of peroxisomes from the endomembranous system. *Cell. Mol. Life Sci.*

Gabaldón, T. 2014. Evolutionary considerations on the origin of peroxisomes from the endoplasmic reticulum, and their relationships with mitochondria. *Cell. Mol. Life Sci.*

Gabaldón, T., and S. Capella-Gutiérrez. 2010. Lack of phylogenetic support for a supposed actinobacterial origin of peroxisomes. *Gene*. 465:61–65.

Gabaldón, T., B. Snel, F.V. Zimmeren, W. Hemrika, H. Tabak, and M.A. Huynen. 2006. Origin and evolution of the peroxisomal proteome. *Biol. Direct.* 1:8.

Galagan, J.E., S.E. Calvo, C. Cuomo, L.-J. Ma, J.R. Wortman, S. Batzoglou, S.-I. Lee, M. Baştürkmen, C.C. Spevak, J. Clutterbuck, V. Kapitonov, J. Jurka, C. Scazzocchio, M. Farman, J. Butler, S. Purcell, S. Harris, G.H. Braus, O. Draht, S. Busch, C. D'Enfert, C. Bouchier, G.H. Goldman, D. Bell-Pedersen, S. Griffiths-Jones, J.H. Doonan, J. Yu, K. Vienken, A. Pain, M. Freitag, E.U. Selker, D.B. Archer, M.Á. Peñalva, B.R. Oakley, M. Momany, T. Tanaka, T. Kumagai, K. Asai, M. Machida, W.C. Nierman, D.W. Denning, M. Caddick, M. Hynes, M. Paoletti, R. Fischer, B. Miller, P. Dyer, M.S. Sachs, S.A. Osmani, and B.W. Birren. 2005. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature Cell Biol*. 438:1105–1115.

Galagan, J.E., S.E. Calvo, K.A. Borkovich, E.U. Selker, N.D. Read, D. Jaffe, W. FitzHugh, S. Smirnov, S. Purcell, B. Rehman, T. Elkins, R. Engels, S. Wang, C.B. Nielsen, J. Butler, M. Endrizzi, D. Qui, P. Ianakiev, D. Bell-Pedersen, M.A. Nelson, M. Werner-Washburne, C.P. Selitrennikoff, J.A. Kinsey, E.L. Braun, Alex Zelter, U. Schulte, G.O. Kothe, G. Jedd, Werner Mewes, C. Staben, E. Marcotte, D. Greenberg, A. Roy, K. Foley, J. Naylor, N. Stange-Thomann, R. Barrett, S. Gnerre, M. Kamal, M. Kamvysselis, E. Mauceli, C. Bielke, S. Rudd, D. Frishman, S. Krystofova, C. Rasmussen, R.L. Metzenberg, D.D. Perkins, S. Kroken, C. Cogoni, G. Macino, D. Catcheside, W. Li, R.J. Pratt, S.A. Osmani, Colin P. C. DeSouza, L. Glass, M.J. Orbach, J.A. Berglund, R. Voelker, O. Yarden, M. Plamann, S. Seiler, J. Dunlap, A. Radford, R. Aramayo, D.O. Natvig, L.A. Alex, G. Mannhaupt, D.J. Ebbole, M. Freitag, I. Paulsen, M.S. Sachs, E.S. Lander, C.N. Bruce, and B. Birren.

2003. The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature*. 422:859–868.

Gancedo, J.M. 1998. Yeast carbon catabolite repression. *Microbiol. Mol. Biol. Rev*. 62:334–361.

Gardner, M.J., M.J. Gardner, R. Bishop, T. Shah, E.P. de Villiers, J.M. Carlton, N. Hall, Q. Ren, I.T. Paulsen, A. Pain, M. Berriman, R.J.M. Wilson, S. Sato, S.A. Ralph, D.J. Mann, Z. Xiong, S.J. Shallom, J. Weidman, L. Jiang, J. Lynn, B. Weaver, A. Shoaibi, A.R. Domingo, D. Wasawo, J. Crabtree, J.R. Wortman, B. Haas, S.V. Angiuoli, T.H. Creasy, B.S.J.C.S. Charles Lu, T.R. Utterback, T.V. Feldblyum, M. Pertea, J. Allen, W.C. Nierman, E.L.N. Taracha, S.L. Salzberg, O.R. White, H.A. Fitzhugh, S. Morzaria, J.C. Venter, C.M. Fraser, and V. Nene. 2005. Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. *Science*. 309:134–137.

Gardner, M.J., N. Hall, E. Fung, O. White, M. Berriman, R.W. Hyman, J.M. Carlton, A. Pain, K.E. Nelson, S. Bowman, I.T. Paulsen, K. James, J.A. Eisen, K. Rutherford, S.L. Salzberg, A. Craig, S. Kyes, M.-S. Chan, V. Nene, S.J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M.W. Mather, A.B. Vaidya, D.M.A. Martin, A.H. Fairlamb, M.J. Fraunholz, D.S. Roos, S.A. Ralph, G.I. McFadden, L.M. Cummings, G.M. Subramanian, C. Mungall, J.C. Venter, D.J. Carucci, S.L. Hoffman, C. Newbold, R.W. Davis, C.M. Barrell, and B. Barrell. 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*. 419:498–511.

Gietz, D., and R.A. Woods. 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods in Enzymol*. 350:87–96.

Glover, J.H., D.W. Andrews, and R.A. Rachubinski. 1994. *Saccharomyces cerevisiae* peroxisomal thiolase is imported as a dimer. *Proc. Natl. Acad. Sci.* 91:10541–10545.

Goff, S.A., D. Ricke, T.-H. Lan, G. Presting, R. Wang, M. Dunn, J. Glazebrook, A. Sessions, P. Oeller, H. Varma, D. Hadley, D. Hutchison, C. Martin, F. Katagiri, B.M. Lange, T. Moughamer, Y. Xia, P. Budworth, J. Zhong, T. Miguel, U. Paszkowski, S. Zhang, M. Colbert, W.-L. Sun, L. Chen, B. Cooper, S. Park, T.C. Wood, L. Mao, P. Quail, R. Wing, R. Dean, Y. Yu, A. Zharkikh, R. Shen, S. Sahasrabudhe, A. Thomas, R. Cannings, A. Gutin, D. Pruss, J. Reid, S. Tavtigian, J. Mitchell, G. Eldredge, T. Scholl, R.M. Miller, S. Bhatnagar, N. Adey, T. Rubano, N. Tusneem, R. Robinson, J. Feldhaus, T. Macalma, A. Oliphant, and S. Briggs. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science*. 296:92–100.

Goffeau, A., B.G. Barrell, H. Bussey, R.W. Davis, B. Dujon, H. Feldmann, F. Galilbert, J.D. Hoheisel, C. Jacq, M. Johnston, E.J. Louis, W. Mewes, Y. Murakami, P.T. Philippsen, H. Tettelin, and S.G. Oliver. 1996. Life with 6000 genes. *Science*. 274:546–567.

Gonzalez, F.J. 1997. Recent update on the PPAR alpha-null mouse. *Biochimie*. 79:139–144.

Gonzalez, N.H., G. Felsner, F.D. Schramm, A. Klingl, U.-G. Maier, and K. Bolte. 2011. A single peroxisomal targeting signal mediates matrix protein import in diatoms. *PLoS ONE*. 9:e25316.

Goodman, J.M., S.B. Trapp, H. Hwang, and M. Veenhuis. 1990. Peroxisomes induced in *Candida boidinii* by methanol, oleic acid and D-alanine vary in metabolic function but share common integral membrane proteins. *J. Cell Sci*. 97:193–204.

Goodwin, S.B., S. Ben M'Barek, B. Dhillon, A.H.J. Wittenberg, C.F. Crane, J.K. Hane, A.J. Foster, T.A.J. Van der Lee, J. Grimwood, A. Aerts, J. Antoniw, A. Bailey, B. Bluhm, J. Bowler, J. Bristow, A. van

der Burgt, B. Canto-Canché, A.C.L. Churchill, L. Conde-Ferràez, H.J. Cools, P.M. Coutinho, M. Csukai, P. Dehal, P. De Wit, B. Donzelli, H.C. van de Geest, R.C.H.J. van Ham, K.E. Hammond-Kosack, B. Henrissat, A. Kilian, A.K. Kobayashi, E. Koopmann, Y. Kourmpetis, A. Kuzniar, E. Lindquist, V. Lombard, C. Maliepaard, N. Martins, R. Mehrabi, J.P.H. Nap, A. Ponomarenko, J.J. Rudd, A. Salamov, J. Schmutz, H.J. Schouten, H. Shapiro, I. Stergiopoulos, S.F.F. Torriani, H. Tu, R.P. de Vries, C. Waalwijk, S.B. Ware, A. Wiebenga, L.-H. Zwiers, R.P. Oliver, I.V. Grigoriev, and G.H.J. Kema. 2011. Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS Genet.* 7:e1002070.

Gootjes, J., F. Schmohl, P.A.W. Mooijer, C. Dekker, H. Mandel, M. Topcu, M. Huemer, M. von Schütz, T. Marquardt, J.A. Smeitink, H.R. Waterham, and R.J.A. Wanders. 2004. Identification of the molecular defect in patients with peroxisomal mosaicism using a novel method involving culturing of cells at 40°C: implications for other inborn errors of metabolism. *Hum. Mutat.* 24:130–139.

Goto, S., Mano, S., Nakamori, C., and M. Nishimura. 2011. Arabidopsis ABERRANT PEROXISOME MORPHOLOGY9 is a peroxin that recruits the PEX1-PEX6 complex to peroxisomes. *Plant Cell.* 23:1573–1587.

Gould, S.J., G.-A. Keller, N. Hosken, J. Wilkinson, and S. Subramani. 1989. A Conserved tripeptide sorts proteins to peroxisomes. *J. Cell Biol.* 108:1657–1664.

Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52:696–704.

Guo, T., Y.Y. Kit, J.-M. Nicaud, M.-T. Le Dall, S.K. Sears, H. Vali, H. Chan, R.A. Rachubinski, and V.I. Titorenko. 2003. Peroxisome division in the yeast *Yarrowia lipolytica* is regulated by a signal from inside the peroxisome. *J. Cell Biol.* 162:1255–1266.

Guo, T., C. Gregg, T. Boukh-Viner, P. Kyryakov, A. Goldberg, S. Bourque, F. Banu, S. Haile, S. Milijevic, K.H.Y. San, J. Solomon, V. Wong, and V.I. Titorenko. 2007. A signal from inside the peroxisome initiates its division by promoting the remodeling of the peroxisomal membrane. *J. Cell Biol.* 177:289–303.

Guther, M.L.S., M. Urbaniak, A. Tavendale, A. Prescott, and M.A.J. Ferguson. 2014. A high-confidence glycosome proteome for procyclic form *Trypanosoma brucei* by epitope-tag organelle enrichment and SILAC proteomics. *J. Proteome Res.* 1–36.

Gurvitz, A., J.K. Hiltunen, R. Erdmann, B. Hamilton, A. Hartig, H. Ruis, and H. Rottensteiner. 2001. *Saccharomyces cerevisiae* Adr1p governs fatty acid beta-oxidation and peroxisome proliferation by regulating *POX1* and *PEX11*. *J. Biol. Chem.* 276:31825–31830.

Halbach, A., S. Lorenzen, C. Landgraf, R. Volkmer-Engert, R. Erdmann, and H. Rottensteiner. 2005. Function of the PEX19-binding site of human adrenoleukodystrophy protein as targeting motif in man and yeast. PMP targeting is evolutionarily conserved. *J. Biol. Chem.* 284:3906-3916.

Hammond, A.T., and B.S. Glick. 2000. Raising the speed limits for 4D fluorescence microscopy. *Traffic.* 1:935–940.

Hannaert, V., E. Saavedra, F. Duffieux, J.-P. Szikora, D.J. Rigden, P.A.M. Michels, and F.R. Opperdoes. 2003. Plant-like traits associated with metabolism of *Trypanosoma* parasites. *PNAS.* 100:1–5.

Hayashi, M., M. Yagi, K. Nito, T. Kamada, and M. Nishimura. 2005. Differential contribution of two peroxisomal protein receptors to the maintenance of peroxisomal functions in *Arabidopsis*. *J. Biol. Chem.* 280:14829–14835.

Heiland, I., and R. Erdmann. 2005. Biogenesis of peroxisomes. Topogenesis of the peroxisomal membrane and matrix proteins. *FEBS J.* 272:2362-2372.

Herman, E.K., A.L. Greninger, G.S. Visvesvara, F. Marciano-Cabral, J.B. Dacks, and C.Y. Chiu. 2013. The Mitochondrial Genome and a 60-kb Nuclear DNA segment from *Naegleria fowleri*, the causative agent of primary amoebic meningoencephalitis. *J. Eukaryot. Microbiol.* 60:179–191.

Hibbett, D.S. 2006. A phylogenetic overview of the Agaricomycotina. *Mycologia*. 98:917–925.

Hiltunen, J.K., A.M. Mursula, H. Rottensteiner, R.K. Wierenga, A.J. Kastaniotis, and A. Gurvitz. 2003. The biochemistry of peroxisomal beta-oxidation in the yeast *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev.* 27:35–64.

Hoepfner, D., D. Schildknegt, I. Braakman, P. Philippsen, and H.F. Tabak. 2005. Contribution of the endoplasmic reticulum to peroxisome formation. *Cell.* 122:85–95.

Hoepfner, D., M. van den Berg, P. Philippsen, H.F. Tabak, and E.H. Hettema. 2001. A role for Vps1p, actin, and the Myo2p motor in peroxisome abundance and inheritance in *Saccharomyces cerevisiae*. *J. Cell Biol.* 155:979–990.

Houdou, S., H. Kuruta, M. Hasegawa, H. Konomi, S. Takashima, Y. Suzuki, and T. Hashimoto. 1991. Developmental immunohistochemistry of catalase in the human brain. *Brain Res.* 556:267-270.

Howe, K., M.D. Clark, C.F. Torroja, J. Torrance, C. Berthelot, M. Muffato, J.E. Collins, S. Humphray, K. McLaren, L. Matthews, S. McLaren, I. Sealy, M. Caccamo, C. Churcher, C. Scott, J.C. Barrett, R. Koch, G.-J. Rauch, S. White, W. Chow, B. Kilian, L.T. Quintais, J.A. Guerra-Assunção, Y. Zhou, Y. Gu, J. Yen, J.-H. Vogel, T. Eyre, S. Redmond, R. Banerjee, J. Chi, B. Fu, E. Langley, S.F. Maguire, G.K. Laird, D. Lloyd, E. Kenyon, S. Donaldson, H. Sehra, J. Almeida-King, J. Loveland, S. Trevanion, M. Jones, M. Quail, D. Willey, A. Hunt, J. Burton, S. Sims, K. McLay, B. Plumb, J. Davis, C. Clee, K. Oliver, R. Clark, C. Riddle, D. Eliott, G. Threadgold, G. Harden, D. Ware, B. Mortimer, G. Kerry, P. Heath, B. Phillimore, A. Tracey, N. Corby, M. Dunn, C. Johnson, J. Wood, S. Clark, S. Pelan, G. Griffiths, M. Smith, R. Glithero, P. Howden, N. Barker, C. Stevens, J. Harley, K. Holt, G. Panagiotidis, J. Lovell, H. Beasley, C. Henderson, D. Gordon, K. Auger, D. Wright, J. Collins, C. Raisen, L. Dyer, K. Leung, L. Robertson, K. Ambridge, D. Leongamornlert, S. McGuire, R. Gilderthorp, C. Griffiths, D. Manthravadi, S. Nichol, G. Barker, S. Whitehead, M. Kay, J. Brown, C. Murnane, E. Gray, M. Humphries, N. Sycamore, D. Barker, D. Saunders, J. Wallis, A. Babbage, S. Hammond, M. Mashreghi-Mohammadi, L. Barr, S. Martin, P. Wray, A. Ellington, N. Matthews, M. Ellwood, R. Woodmansey, G. Clark, J. Cooper, A. Tromans, D. Grafham, C. Skuce, R. Pandian, R. Andrews, E. Harrison, A. Kimberley, J. Garnett, N. Fosker, R. Hall, P. Garner, D. Kelly, C. Bird, S. Palmer, I. Gehring, A. Berger, C.M. Dooley, Z. Ersan-Ürün, C. Eser, H. Geiger, M. Geisler, L. Karotki, A. Kirn, J. Konantz, M. Konantz, M. Oberländer, S. Rudolph-Geiger, M. Teucke, K. Osoegawa, B. Zhu, A. Rapp, S. Widaa, C. Langford, F. Yang, N.P. Carter, J. Harrow, Z. Ning, J. Herrero, S.M.J. Searle, A. Enright, R. Geisler, R.H.A. Plasterk, C. Lee, M. Westerfield, P.J. de Jong, L.I. Zon, J.H. Postlethwait, C. Nüsslein-Volhard, T.J.P. Hubbard, H. Roest Crollius, J. Rogers, and D.L. Stemple. 2013. The zebrafish reference genome sequence and its relationship to the human genome. *Nature*. 1–6.

Huybrechts, S.J., P.P. Van Veldhoven, C. Brees, G.P. Mannaerts, G.V. Los, and M. Fransen. 2009.

Peroxisome dynamics in cultured mammalian cells. *Traffic*. 10:1722–1733.

Huynh, T.V., R.A. Young, and R.W. Davis. 1985. DNA cloning: a practical approach. IRL Press, Oxford.

Hynes, M.J., S.L. Murray, A. Duncan, G.S. Khew, and M.A. Davis. 2006. Regulatory genes controlling fatty acid catabolism and peroxisomal functions in the filamentous fungus *Aspergillus nidulans*. *Eukaryot. Cell*. 5:794–805.

Innis, M.A., and D.H. Gelfand. 1990. Optimization of PCRs. *In* PCR Protocols: A guide to Methods and Applications. M.A. Innis, D.H. Gelfand, J.J. Sninsky, and T.J. White, editors. Academic Press, San Diego. 3–12.

International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. *Nature*. 409:860–921.

Islinger, M., M.J.R. Cardoso, and M. Schrader. 2010. Be different-the diversity of peroxisomes in the animal kingdom. *Biophys. Biochim. Acta.* 1803:881-897.

Issemann, I., and S. Green. 1990. Activation of a member of the steroid hormone receptor superfamily by peroxisome proliferators. *Nature.* 347:645–650.

Ivens, A.C., C.S. Peacock, E.A. Worthey, L. Murphy, G. Aggarwal, M. Berriman, E. Sisk, M.-A. Rajandream, E. Adlem, R. Aert, A. Anupama, Z. Apostolou, P. Attipoe, N. Bason, C. Bauser, A. Beck, S.M. Beverley, G. Bianchettin, K. Borzym, G. Bothe, C.V. Bruschi, M. Collins, E. Cadag, L. Ciarloni, C. Clayton, R.M.R. Coulson, A. Cronin, A.K. Cruz, R.M. Davies, J. De Gaudenzi, D.E. Dobson, A. Duesterhoeft, G. Fazelina, N. Fosker, A.C. Frasch, A. Fraser, M. Fuchs, C. Gabel, A. Goble, A. Goffeau, D. Harris, C. Hertz-Fowler, H. Hilbert, D. Horn, Y. Huang, S. Klages, A. Knights, M. Kube, N. Larke, L. Litvin, A. Lord, T. Louie, M. Marra, D. Masuy, K. Matthews, S. Michaeli, J.C. Mottram, S. Muller-Auer, H. Munden, S. Nelson, H. Norbertczak, K. Oliver, S. ONeil, M. Pentony, T.M. Pohl, C. Price, B. Purnelle, M.A. Quail, E. Rabbinowitsch, R. Reinhardt, M. Rieger, J. Rinta, J. Robben, L. Robertson, J.C. Ruiz, S. Rutter, D. Saunders, M. Schafer, J. Schein, D.C. Schwartz, K. Seeger, A. Seyler, S. Sharp, H. Shin, D. Sivam, R. Squares, S. Squares, V. Tosato, C. Vogt, G. Volckaert, R. Wambutt, T. Warren, H. Wedler, J. Woodward, S. Zhou, W. Zimmermann, D.F. Smith, J.M. Blackwell, et al. 2005. The genome of the kinetoplastid parasite, *Leishmania major*. *Science*. 309:436–442.

Jackson, A.P., M.A. Quail, and M. Berriman. 2009. Insights into the genome sequence of a free-living kinetoplastid: *Bodo saltans* (Kinetoplastida: Euglenozoa). *BMC Genomics*. 9.

James, T.Y., A. Pelin, L. Bonen, S. Ahrendt, D. Sain, N. Corradi, and J.E. Stajich. 2013. Shared signatures of parasitism and phylogenomics unite Cryptomycota and Microsporidia. *Curr. Biol*. 23:1548–1553.

James, T.Y., F. Kauff, C.L. Schoch, P.B. Matheny, V. Hofstetter, C.J. Cox, G. Celio, C. Gueidan, E. Fraker, J. Miadlikowska, H.T. Lumbsch, A. Rauhut, V. Reeb, A.E. Arnold, A. Amtoft, J.E. Stajich, K. Hosaka, G.-H. Sung, D. Johnson, B. O'Rourke, M. Crockett, M. Binder, J.M. Curtis, J.C. Slot, Z. Wang, A.W. Wilson, A. Schüßler, J.E. Longcore, K. O'Donnell, S. Mozley-Standridge, D. Porter, P.M. Letcher, M.J. Powell, J.W. Taylor, M.M. White, G.W. Griffith, D.R. Davies, R.A. Humber, J.B. Morton, J. Sugiyama, A.Y. Rossman, J.D. Rogers, D.H. Pfister, D. Hewitt, K. Hansen, S. Hambleton, R.A. Shoemaker, J. Kohlmeyer, B. Volkmann-Kohlmeyer, R.A. Spotts, M. Serdani, P.W. Crous, K.W. Hughes, K. Matsuura, E. Langer, G. Langer, W.A. Untereiner, R. Lücking, B.

Büdel, D.M. Geiser, A. Aptroot, P. Diederich, I. Schmitt, M. Schultz, R. Yahr, D.S. Hibbett, F. Lutzoni, D.J. McLaughlin, J.W. Spatafora, and R. Vilgalys. 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature*. 443:818–822.

Jeffries, T.W., I.V. Grigoriev, J. Grimwood, J.M. Laplaza, A. Aerts, A. Salamov, J. Schmutz, E. Lindquist, P. Dehal, H. Shapiro, Y.-S. Jin, V. Passoth, and P.M. Richardson. 2007. Genome sequence of the lignocellulose-bioconverting and xylose-fermenting yeast *Pichia stipitis*. *Nat. Biotechnol*. 25:319–326.

Jones, T., N.A. Federspiel, H. Chibana, J. Dungan, S. Kalman, B.B. Magee, G. Newport, Y.R. Thorstenson, N. Agabian, P.T. Magee, R.W. Davis, and S. Scherer. 2004. The diploid genome sequence of *Candida albicans*. *PNAS*. 101:7329–7334.

Joshi, S., G. Agrawal, and S. Subramani. 2012. Phosphorylation-dependent Pex11p and Fis1p interaction regulates peroxisome division. *Mol. Biol. Cell*. 23:1307–1315.

Kamiryo, T., M. Abe, K. Okazaki, S. Kato, and N. Shimamoto. 1982. Absence of DNA in peroxisomes of *Candida tropicalis*. *J. Bacteriol*. 152:269–274.

Katinka, M.D., S. Duprat, E. Cornillot, G. Metenier, F. Thomarat, G. Prensier, V. Barbe, E. Peyretaillade, P. Brottier, P. Wincker, F. Delbac, H. El Alaoui, P. Peyret, W. Saurin, M. Gouy, J. Weissenbach, and C.P. Vivares. 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi. Nature*. 414:450–453.

Kämper, J., R. Kahmann, M. Bölker, L.-J. Ma, T. Brefort, B.J. Saville, F. Banuett, J.W. Kronstad, S.E. Gold, O. Müller, M.H. Perlin, H.A.B. Wösten, R. de Vries, J. Ruiz-Herrera, C.G. Reynaga-Peña, K. Snetselaar, M. McCann, J. Pérez-Martín, M. Feldbrügge, C.W. Basse, G. Steinberg, J.I. Ibeas, W. Holloman, P. Guzman, M. Farman, J.E. Stajich, R. Sentandreu, J.M. González-Prieto, J.C. Kennell, L. Molina, J. Schirawski, A. Mendoza-Mendoza, D. Greilinger, K. Münch, N. Rössel, M. Scherer, M. Vraneš, O. Ladendorf, V. Vincon, U. Fuchs, B. Sandrock, S. Meng, E.C.H. Ho, M.J. Cahill, K.J. Boyce, J. Klose, S.J. Klosterman, H.J. Deelstra, L. Ortiz-Castellanos, W. Li, P. Sanchez-Alonso, P.H. Schreier, I. Häuser-Hahn, M. Vaupel, E. Koopmann, G. Friedrich, H. Voss, T. Schlüter, J. Margolis, D. Platt, C. Swimmer, A. Gnirke, F. Chen, V. Vysotskaia, G. Mannhaupt, U. Güldener, M. Münsterkötter, D. Haase, M. Oesterheld, H.-W. Mewes, E.W. Mauceli, D. DeCaprio, C.M. Wade, J. Butler, S. Young, D.B. Jaffe, S. Calvo, C. Nusbaum, J. Galagan, and B.W. Birren. 2006. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature*. 444:97–101.

Kassmann, C.M., C. Lappe-Siefke, M. Baes, B. Brügger, A. Mildner, H.B. Werner, O. Natt, T. Michaelis, M. Prinz, J. Frahm, and K.-A. Nave. 2007. Axonal loss and neuroinflammation caused by peroxisome-deficient oligodendrocytes. *Nat. Genet*. 39:969–976.

Kelley, R.I. 1983. The cerebrohepatorenal syndrome of Zellweger, morphologic and metabolic aspects. *Am. J. Med. Genet.* 16:503-517.

Kellis, M., N. Patterson, M. Endrizzi, B. Birren, and E.S. Lander. 2003. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature*. 423:241–254.

Kerscher, S.J., A. Eschemann, P.M. Okun, and U. Brant. 2001. External alternative NADH:ubiquinone oxidoreductase redirected to the internal face of the mitochondrial inner membrane rescues complex I deficiency in *Yarrowia lipolytica*. *J. Cell Sci.* 114:3915–3921.

Kiel, J.A.K.W., M. Veenhuis, and I.J. van der Klei. 2006. *PEX* genes in fungal genomes: common, rare or redundant. *Traffic*. 7:1291–1303.

Kim, P. K., R.T. Mullen, U. Schumann, and J. Lippincott-Schwartz. 2006. The origin and maintenance of mammalian peroxisomes involves a *de novo PEX16*-dependent pathway from the ER. *J. Cell Biol.* 173:521–532.

King, N., M.J. Westbrook, S.L. Young, A. Kuo, M. Abedin, J. Chapman, S. Fairclough, U. Hellsten, Y. Isogai, I. Letunic, M. Marr, D. Pincus, N. Putnam, A. Rokas, K.J. Wright, R. Zuzow, W. Dirks, M. Good, D. Goodstein, D. Lemons, W. Li, J.B. Lyons, A. Morris, S. Nichols, D.J. Richter, A. Salamov, J. Sequencing, P. Bork, W.A. Lim, G. Manning, W.T. Miller, W. McGinnis, H. Shapiro, R. Tjian, I.V. Grigoriev, and D. Rokhsar. 2008. The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. *Nature*. 451:783–788.

Kissinger, J.C., B. Gajria, L. Li, I.T. Paulsen, and D.S. Roos. 2003. ToxoDB: accessing the *Toxoplasma gondii* genome. *Nuc. Acids Res*. 31:234–236.

Klein, A.T., M. van den Berg, G. Bottger, H.F. Tabak, and B. Distel. 2002. *Saccharomyces cerevisiae* acyl-CoA oxidase follows a novel, non-PTS1, import pathway into peroxisomes that is dependent on Pex5p. *J. Biol. Chem*. 277:25011–25019.

Klute, M.J., P. Melançon, and J.B. Dacks. 2011. Evolution and diversity of the Golgi. *In* The Golgi. G. Warren and J. Rothman, editors. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 283-299.

Knoblach, B., and R.A. Rachubinski. 2010. Phosphorylation-dependent activation of peroxisome proliferator protein *PEX11* controls peroxisome abundance. *J. Biol. Chem*. 285:6670–6680.

Kobayashi, S., A. Tanaka, and Y. Fujiki. 2007. Fis1, DLP1, and Pex11p coordinately regulate peroxisome morphogenesis. *Exp. Cell Res*. 313:1675–1686.

Koch, A., M. Thiemann, M. Grabenbauer, Y. Yoon, M.A. McNiven, and M. Schrader. 2003. Dynamin-like protein 1 is involved in peroxisomal fission. *J. Biol. Chem*. 278:8597–8605.

Koch, J., K. Pranjic, A. Huber, A. Ellinger, A. Hartig, F. Krager, and C. Brocard. 2010. *PEX11* family members are membrane elongation factors that coordinate peroxisome proliferation and maintenance. *J. Cell Sci*. 123:3389–3400.

Kuravi, K., S. Nagotu, A.M. Krikken, K. Sjollema, M. Deckers, R. Erdmann, M. Veenhuis, and I.J. van der Klei. 2006. Dynamin-related proteins Vps1p and Dnm1p control peroxisome abundance in *Saccharomyces cerevisiae*. *J. Cell. Sci.* 119:3994-4001.

Lambkin, G.R., and R.A. Rachubinski. 2001. *Yarrowia lipolytica* cells mutant for the peroxisomal peroxin Pex19p contain structures resembling wild-type peroxisomes. *Mol. Biol. Cell*. 12:3353–3364.

Lee, S.S, T. Pineau, J. Drago, E.J. Lee, J.W. Owens, D.L. Kroetz, P.M. Fernandez-Salguero, H. Westphal, and F.J. Gonzalez. 1995. Targeted disruption of the alpha isoform of the peroxisome proliferator-activated receptor gene in mice results in abolishment of the pleiotropic effects of peroxisome proliferators. *Mol. Cell. Biol.* 15:3012–3022.

Leon, S., and S. Subramani. 2007. A conserved cysteine residue of *Pichia pastoris* Pex20p is essential

for its recycling from the peroxisome to the cytosol. *J. Biol. Chem*. 282:7424–7430.

Li, X., E. Baumgart, J.C. Morrell, G. Jimenez-Sanchez, D. Valle, and S.J. Gould. 2002. *PEX11* deficiency is lethal and impairs neuronal migration but does not abrogate peroxisome function. *Mol. Cell. Biol.* 22:4358–4365.

Lin, Y., L. Sun, L.V. Nguygen, R.A. Rachubinski, and H.M. Goodman. 1999. The Pex16p Homolog SSE1 and storage organelle formation in *Arabidopsis* seeds. *Science*. 284:328–330.

Lindblad-Toh, K., C.M. Wade, T.S. Mikkelsen, E.K. Karlsson, D.B. Jaffe, M. Kamal, M. Clamp, J.L. Chang, E.J. Kulbokas, M.C. Zody, E. Mauceli, X. Xie, M. Breen, R.K. Wayne, E.A. Ostrander, C.P. Ponting, F. Galibert, D.R. Smith, P.J. deJong, E. Kirkness, P. Alvarez, T. Biagi, W. Brockman, J. Butler, C.-W. Chin, A. Cook, J. Cuff, M.J. Daly, D. DeCaprio, S. Gnerre, M. Grabherr, M. Kellis, M. Kleber, C. Bardeleben, L. Goodstadt, A. Heger, C. Hitte, L. Kim, K.-P. Koepfli, H.G. Parker, J.P. Pollinger, S.M.J. Searle, N.B. Sutter, R. Thomas, C. Webber, J. Baldwin, A. Abebe, A. Abouelleil, L. Aftuck, M. Ait-zahra, T. Aldredge, N. Allen, P. An, S. Anderson, C. Antoine, H. Arachchi, A. Aslam, L. Ayotte, P. Bachantsang, A. Barry, T. Bayul, M. Benamara, A. Berlin, D. Bessette, B. Blitshteyn, T. Bloom, J. Blye, L. Boguslavskiy, C. Bonnet, B. Boukhgalter, A. Brown, P. Cahill, N. Calixte, J. Camarata, Y. Cheshatsang, J. Chu, M. Citroen, A. Collymore, P. Cooke, T. Dawoe, R. Daza, K. Decktor, S. DeGray, N. Dhargay, K. Dooley, K. Dooley, P. Dorje, K. Dorjee, L. Dorris, N. Duffey, A. Dupes, O. Egbiremolen, R. Elong, J. Falk, A. Farina, S. Faro, D. Ferguson, P. Ferreira, et al. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature Cell Biol*. 438:803–819.

Lingard, M.J., and R.N. Trelease. 2006. Five *Arabidopsis* peroxin 11 homologs individually promote peroxisome elongation, duplication or aggregation. *J. Cell Sci.* 119:1961–1972.

Lingner, T., A.R. Kataya, G.E. Antonicelli, A. Benichou, K. Nilssen, X.Y. Chen, T. Siemsen, B. Morgenstern, P. Meinicke, and S. Reumann. 2011. Identification of novel plant peroxisomal targeting signals by a combination of machine learning methods and in vivo subcellular targeting analyses. *Plant Cell*. 23:1556–1572.

Liti, G., D.M. Carter, A.M. Moses, J. Warringer, L. Parts, S.A. James, R.P. Davey, I.N. Roberts, A. Burt, V. Koufopanou, I.J. Tsai, C.M. Bergman, D. Bensasson, M.J.T. O'Kelly, A. van Oudenaarden, D.B.H. Barton, E. Bailes, A.N. Nguyen, M. Jones, M.A. Quail, I. Goodhead, S. Sims, F. Smith, A. Blomberg, R. Durbin, and E.J. Louis. 2009. Population genomics of domestic and wild yeasts. *Nature*. 458:337–341.

Litwin, J.A., A. Volkl, J. Muller-Hocker, and F.D. Fahimi. 1988. Immunocytochemical demonstration of peroxisomal enzymes in human kidney biopsies. *Virchows Arch., B Cell. Pathol. Incl. Mol. Pathol.* 54:207-213.

Loftus, B., I. Anderson, R. Davies, U.C.M. Alsmark, J. Samuelson, P. Amedeo, P. Roncaglia, M. Berriman, R.P. Hirt, B.J. Mann, T. Nozaki, B. Suh, M. Pop, M. Duchene, J. Ackers, E. Tannich, M. Leippe, M. Hofer, I. Bruchhaus, U. Willhoeft, A. Bhattacharya, T. Chillingworth, C. Churcher, Z. Hance, B. Harris, D. Harris, K. Jagels, S. Moule, K. Mungall, D. Ormond, R. Squares, S. Whitehead, M.A. Quail, E. Rabbinowitsch, H. Norbertczak, C. Price, Z. Wang, N.G. n, C. Gilchrist, S.E. Stroup, S. Bhattacharya, A. Lohia, P.G. Foster, T. Sicheritz-Ponten, C. Weber, U. Singh, C. Mukherjee, N.M. El-Sayed, W.A. Petri Jr, C.G. Clark, T.M. Embley, B. Barrell, C.M. Fraser, and N. Hall. 2005. The genome of the protist parasite *Entamoeba histolytica*. *Nature*. 433:865–868.

Lorenz, P.1., A.G. Maier, E. Baumgart, R. Erdmann, and C. Clayton. 1998. Elongation and clustering of glycosomes in *Trypanosoma brucei* overexpressing the glycosomal Pex11p. *EMBO J.* 17:3542–3555.

Lowry, O.H., N.J. Rosebrough, A.L. Farr, and R.J. Randall. 1951. Protein measurement with the folin phenol reagent. *J. Biol. Chem.* 193:265–275.

Ma, L.-J., A.S. Ibrahim, C. Skory, M.G. Grabherr, G. Burger, M. Butler, M. Elias, A. Idnurm, B.F. Lang, T. Sone, A. Abe, S.E. Calvo, L.M. Corrochano, R. Engels, J. Fu, W. Hansberg, J.-M. Kim, C.D. Kodira, M.J. Koehrsen, B. Liu, D. Miranda-Saavedra, S. O'Leary, L. Ortiz-Castellanos, R. Poulter, J. Rodriguez-Romero, J. Ruiz-Herrera, Y.-Q. Shen, Q. Zeng, J. Galagan, B.W. Birren, C.A. Cuomo, and B.L. Wickes. 2009. Genomic analysis of the basal lineage fungus *Rhizopus oryzae* reveals a whole-genome duplication. *PLoS Genet.* 5:e1000549.

Maier, A., P. Lorenz, F. Voncken, and C. Clayton. 2001. An essential dimeric membrane protein of trypanosome glycosomes. *Mol. Microbiol.* 39:1443–1451.

Managadze, D., C. Würtz, S. Wiese, M. Schneider, W. Girzalsky, H.E. Meyer, R. Erdmann, B. Warscheid, and H. Rottensteiner. 2010. Identification of *PEX33*, a novel component of the peroxisomal docking complex in the filamentous fungus *Neurospora crassa*. *Eur. J. Cell Biol.* 89:955–964.

Mandel, H., M. Espeel, F. Roels, N. Sofer, A. Luder, T.C. Iancu, A. Aizin, M. Berant, R.J.A. Wanders, and R.B.H. Schutgens. 1994. A new type of peroxisomal disorder with variable expression in liver and fibroblasts. *J. Pediatr.* 125:549–555.

Maniatis, T., E. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor.

Manollis, K., B.W. Birren, and E.S. Lander. 2004. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature*. 428:617–624.

Marelli, M., J.J. Smith, S. Jung, E. Yi, A.I. Nesvizhskii, R.H. Christmas, R.A. Saleem, Y.Y.C. Tam, A. Fagarasanu, D.R. Goodlett, R. Aebersold, R.A. Rachubinski, and J.D. Aitchison. 2004. Quantitative mass spectrometry reveals a role for the GTPase Rho1p in actin reorganization on the peroxisome membrane. *J. Cell Biol.* 167:1099–1112.

Marshall, P.A., J.M. Dyer, M.E. Quick, and J.M. Goodman. 1996. Redox-sensitive homodimerization of Pex11p: a proposed mechanism to regulate peroxisome division. *J Cell Biol.* 135:123-137.

Martin, F., A. Aerts, D. Ahrén, A. Brun, E.G.J. Danchin, F. Duchaussoy, J. Gibon, A. Kohler, E. Lindquist, V. Pereda, A. Salamov, H.J. Shapiro, J. Wuyts, D. Blaudez, M. Buée, P. Brokstein, B. Canbäck, D. Cohen, P.E. Courty, P.M. Coutinho, C. Delaruelle, J.C. Detter, A. Deveau, S. DiFazio, S. Duplessis, L. Fraissinet-Tachet, E. Lucic, P. Frey-Klett, C. Fourrey, I. Feussner, G. Gay, J. Grimwood, P.J. Hoegger, P. Jain, S. Kilaru, J. Labbé, Y.C. Lin, V. Legué, F. Le Tacon, R. Marmeisse, D. Melayah, B. Montanini, M. Muratet, U. Nehls, H. Niculita-Hirzel, M.P.O.-L. Secq, M. Peter, H. Quesneville, B. Rajashekar, M. Reich, N. Rouhier, J. Schmutz, T. Yin, M. Chalot, B. Henrissat, U. Kues, S. Lucas, Y. Van de Peer, G.K. Podila, A. Polle, P.J. Pukkila, P.M. Richardson, P. Rouze, I.R. Sanders, J.E. Stajich, A. Tunlid, G. Tuskan, and I.V. Grigoriev. 2008. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature*. 452:88–92.

Martin, F., A. Kohler, C. Murat, R. Balestrini, P.M. Coutinho, O. Jaillon, B. Montanini, E. Morin, B.

Noel, R. Percudani, B. Porcel, A. Rubini, A. Amicucci, J. Amselem, V. Anthouard, S. Arcioni, F. Artiguenave, J.-M. Aury, P. Ballario, A. Bolchi, A. Brenna, A. Brun, M. Buée, B. Cantarel, G. Chevalier, A. Couloux, C. Da Silva, F. Denoeud, S. Duplessis, S. Ghignone, B. Hilselberger, M. Iotti, B. Marçais, A. Mello, M. Miranda, G. Pacioni, H. Quesneville, C. Riccioni, R. Ruotolo, R. Splivallo, V. Stocchi, E. Tisserant, A.R. Viscomi, A. Zambonelli, E. Zampieri, B. Henrissat, M.-H. Lebrun, F. Paolocci, P. Bonfante, S. Ottonello, and P. Wincker. 2010. Perigord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature*. 464:1033–1038.

Martinez, A.J., and G.S. Visvesvara. 1997. Free-living, amphizoic and opportunistic amebas. *Brain Pathol*. 7:583–598.

Martinez, D., L.F. Larrondo, N. Putnam, M.D.S. Gelpke, K. Huang, J. Chapman, K.G. Helfenbein, P. Ramaiya, J.C. Detter, F. Larimer, P.M. Coutinho, B. Henrissat, R. Berka, D. Cullen, and D. Rokhsar. 2004. Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nat. Biotechnol*. 22:695–700.

Martinez, D., R.M. Berka, B. Henrissat, M. Saloheimo, M. Arvas, S.E. Baker, J. Chapman, O. Chertkov, P.M. Coutinho, D. Cullen, E.G.J. Danchin, I.V. Grigoriev, P. Harris, M. Jackson, C.P. Kubicek, C.S. Han, I. Ho, L.F. Larrondo, A.L. de Leon, J.K. Magnuson, S. Merino, M. Misra, B. Nelson, N. Putnam, B. Robbertse, A.A. Salamov, M. Schmoll, A. Terry, N. Thayer, A. Westerholm-Parvinen, C.L. Schoch, J. Yao, R. Barabote, M.A. Nelson, C. Detter, D. Bruce, C.R. Kuske, G. Xie, P. Richardson, D.S. Rokhsar, S.M. Lucas, E.M. Rubin, N. Dunn-Coleman, M. Ward, and T.S. Brettin. 2008. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat. Biotechnol*. 26:553–560.

Mast, F.D. 2013. A comprehensive assessment of peroxisome biology: ER-dependent peroxisome proliferation control, evolution of organelle inheritance in yeast, and a *Drosophila* model of Zellweger syndrome. University of Alberta.

Mast, F. D., L.D. Barlow, R.A. Rachubinski, and J.B. Dacks. 2014. Evolutionary mechanisms for establishing eukaryotic cellular complexity. *Trends Cell Biol*. 24:435–442.

Mast, F.D., A. Fagarasanu, B. Knoblach, and R.A. Rachubinski. 2010. Peroxisome biogenesis: something old, something new, something borrowed. *Physiol.* 25:347–356.

Mast, F.D., J. Li, M.K. Virk, S.C. Hughes, A.J. Simmonds, and R.A. Rachubinski. 2011. A *Drosophila* model for the Zellweger spectrum of peroxisome biogenesis disorders. *Dis. Models Mech.* 4:659–672.

Matsuzaki, T., and Y. Fujiki. 2008. The peroxisomal membrane protein import receptor Pex3p is directly transported to peroxisomes by a novel Pex19p- and Pex16p-dependent pathway. *J. Cell. Biol.* 183:1275-1286.

Matsuzaki, M., O. Misumi, T. Shin-i, S. Maruyama, M. Takahara, S.-Y. Miyagishima, T. Mori, K. Nishida, F. Yagisawa, K. Nishida, Y. Yoshida, Y. Nishimura, S. Nakao, T. Kobayashi, Y. Momoyama, T. Higashiyama, A. Minoda, M. Sano, H. Nomoto, K. Oishi, H. Hayashi, F. Ohta, S. Nishizaka, S. Haga, S. Miura, T. Morishita, Y. Kabeya, K. Terasawa, Y. Suzuki, Y. Ishii, S. Asakawa, H. Takano, N. Ohta, H. Kuroiwa, K. Tanaka, N. Shimizu, S. Sugano, N. Sato, H. Nozaki, N. Ogasawara, Y. Kohara, and T. Kurolwa. 2004. Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature*. 428:653–657.

Medina, E.M., G.W. Jones, and D.A. Fitzpatrick. 2011. Reconstructing the fungal tree of life using

phylogenomics and a preliminary investigation of the distribution of yeast prion-like proteins in the fungal kingdom. *J. Mol. Evol.* 73:116–133.

Meinecke, M., C. Cizmowski, W. Schliebs, V. Kruger, S. Beck, R. Wagner, and R. Erdmann. 2010. The peroxisomal importer constitutes a large and dynamic pore. *Nature Cell Biol.* 12:273–277.

Merchant, S.S., S.E. Prochnik, O. Vallon, E.H. Harris, S.J. Karpowicz, G.B. Witman, A. Terry, A. Salamov, L.K. Fritz-Laylin, L. Marechal-Drouard, W.F. Marshall, L.H. Qu, D.R. Nelson, A.A. Sanderfoot, M.H. Spalding, V.V. Kapitonov, Q. Ren, P. Ferris, E. Lindquist, H. Shapiro, S.M. Lucas, J. Grimwood, J. Schmutz, P. Cardol, H. Cerutti, G. Chanfreau, C.L. Chen, V. Cognat, M.T. Croft, R. Dent, S. Dutcher, E. Fernandez, H. Fukuzawa, D. Gonzalez-Ballester, D. Gonzalez-Halphen, A. Hallmann, M. Hanikenne, M. Hippler, W. Inwood, K. Jabbari, M. Kalanon, R. Kuras, P.A. Lefebvre, S.D. Lemaire, A.V. Lobanov, M. Lohr, A. Manuell, I. Meier, L. Mets, M. Mittag, T. Mittelmeier, J.V. Moroney, J. Moseley, C. Napoli, A.M. Nedelcu, K. Niyogi, S.V. Novoselov, I.T. Paulsen, G. Pazour, S. Purton, J.P. Ral, D.M. Riano-Pachon, W. Riekhof, L. Rymarquis, M. Schroda, D. Stern, J. Umen, R. Willows, N. Wilson, S.L. Zimmer, J. Allmer, J. Balk, K. Bisova, C.J. Chen, M. Elias, K. Gendler, C. Hauser, M.R. Lamb, H. Ledford, J.C. Long, J. Minagawa, M.D. Page, J. Pan, W. Pootakham, S. Roje, A. Rose, E. Stahlberg, A.M. Terauchi, P. Yang, S. Ball, C. Bowler, C.L. Dieckmann, V.N. Gladyshev, P. Green, R. Jorgensen, S. Mayfield, B. Mueller-Roeber, S. Rajamani, R.T. Sayre, P. Brokstein, I. Dubchak, D. Goodstein, L. Hornick, Y.W. Huang, J. Jhaveri, Y. Luo, D. Martínez, W.C.A. Ngau, B. Otillar, A. Poliakov, A. Porter, L. Szajkowski, G. Werner, K. Zhou, I.V. Grigoriev, D.S. Rokhsar, and A.R. Grossman. 2007. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science*. 318:245–250.

Miyata, N., and Y. Fujiki. 2005. Shuttling mechanism of peroxisome targeting signal type 1 receptor Pex5: ATP-independent import and ATP-dependent export. *Mol. Cell. Biol.* 25:10822-10832.

Mohanty, A., and H.M. McBride. 2013. Emerging roles of mitochondria in the evolution, biogenesis, and function of peroxisomes. *Front. Physiol.* 4:1–12.

Morin, E., A. Kohler, A.R. Baker, M. Foulongne-Oriol, V. Lombard, L.G. Nagy, R.A. Ohm, A. Patyshakuliyeva, A. Brun, A.L. Aerts, A.M. Bailey, C. Billette, P.M. Coutinho, G. Deakin, H. Doddapaneni, D. Floudas, J. Grimwood, K. Hilden, U. Kües, K.M. LaButti, A. Lapidus, E.A. Lindquist, S.M. Lucas, C. Murat, R.W. Riley, A.A. Salamov, J. Schmutz, V. Subramanian, H.A.B. Wösten, J. Xu, D.C. Eastwood, G.D. Foster, A.S.M. Sonnenberg, D. Cullen, R.P. de Vries, T. Lundell, D.S. Hibbett, B. Henrissat, K.S. Burton, R.W. Kerrigan, M.P. Challen, I.V. Grigoriev, and F. Martin. 2013. Genome sequence of the button mushroom *Agaricus bisporus* reveals mechanisms governing adaptation to a humic-rich ecological niche. *PNAS*. 110:4146–4146.

Motley, A.M., and E.H. Hettema. 2007. Yeast peroxisomes multiply by growth and division. *J. Cell Biol.* 178:399–410.

Motley, A.M., E.H. Hettema, R. Ketting, R. Plasterk, and H.F. Tabak. 2000. *Caenorhabditis elegans* has a single pathway to target matrix proteins to peroxisomes. *EMBO Reports*. 1:40–46.

Motley, A., G. Ward, and E. Hettema. 2008. Dnm1p-dependent peroxisome fission requires Caf4p, Mdv1p and Fis1p. *J. Cell. Sci.* 121:1633-1640.

Nagotu, S., A.M. Krikken, M. Otzen, J.A.K.W. Kiel, M. Veenhuis, and I.J. van der Klei. 2008. Peroxisome fission in *Hansenula polymorpha* requires Mdv1 and Fis1, two proteins also Involved in mitochondrial fission. *Traffic*. 9:1471–1484.

Nakai, K., and P. Horton. 1999. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem. Sci.* 24:34–35.

Nayidu, N.K., L. Wang, W. Xie, C. Zhang, C. Fan, X. Lian, Q. Zhang, and L. Xiong. 2008. Comprehensive sequence and expression profile analysis of *PEX11* gene family in rice. *Gene*. 412:59–70.

Neuberger, G., S. Maurer-Stroh, B. Eisenhaber, A. Hartig, and F. Eisenhaber. 2003a. Motif refinement of the peroxisomal targeting signal 1 and evaluation of taxon-specific differences. *J. Mol. Biol.* 328:567–579.

Neuberger, G., S. Maurer-Stroh, B. Eisenhaber, A. Hartig, and F. Eisenhaber. 2003b. Prediction of peroxisomal targeting signal 1 containing proteins from amino acid sequence. *J. Mol. Biol.* 328:581–592.

Neuspiel, M., A.C. Schauss, E. Braschi, R. Zunino, P. Rippstein, R.A. Rachubinski, M.A. Andrade-Navarro, and H.M. McBride. 2008. Cargo-selected transport from the mitochondria to peroxisomes is mediated by vesicular carriers. *Curr. Biol.* 18:102–108.

Nicaud, J.-M. 2012. *Yarrowia lipolytica*. *Yeast*. 29:409–418.

Nierman, W.C., A. Pain, M.J. Anderson, J.R. Wortman, H.S. Kim, J. Arroyo, M. Berriman, K. Abe, D.B. Archer, C. Bermejo, J. Bennett, P. Bowyer, D. Chen, M. Collins, R. Coulsen, R. Davies, P.S. Dyer, M. Farman, N. Fedorova, N. Fedorova, T.V. Feldblyum, R. Fischer, N. Fosker, A. Fraser, J.L. García, M.J. García, A. Goble, G.H. Goldman, K. Gomi, S. Griffith-Jones, R. Gwilliam, B. Haas, H. Haas, D. Harris, H. Horiuchi, J. Huang, S. Humphray, J. Jiménez, N. Keller, H. Khouri, K. Kitamoto, T. Kobayashi, S. Konzack, R. Kulkarni, T. Kumagai, A. Lafton, J.-P. Latgé, W. Li, A. Lord, C. Lu, W.H. Majoros, G.S. May, B.L. Miller, Y. Mohamoud, M. Molina, M. Monod, I. Mouyna, S. Mulligan, L. Murphy, S. ONeil, I. Paulsen, M.Á. Peñalva, M. Pertea, C. Price, B.L. Pritchard, M.A. Quail, E. Rabbinowitsch, N. Rawlins, M.-A. Rajandream, U. Reichard, H. Renauld, G.D. Robson, S.R. de Córdoba, J.M. Rodríguez-Peña, C.M. Ronning, S. Rutter, S.L. Salzberg, M. Sanchez, J.C. Sánchez-Ferrero, D. Saunders, K. Seeger, R. Squares, S. Squares, M. Takeuchi, F. Tekaia, G. Turner, C.R.V. de Aldana, J. Weidman, O. White, J. Woodward, J.-H. Yu, C. Fraser, J.E. Galagan, K. Asai, M. Machida, N. Hall, B. Barrell, and D.W. Denning. 2005. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature Cell Biol.* 438:1151–1156.

Nito, K., A. Kamigaki, M. Kondo, M. Hayashi, and M. Nishimura. 2007. Functional classification of *Arabidopsis* peroxisome biogenesis factors proposed from analyses of knockdown mutants. *Plant Cell Physiol.* 48:763-774.

Nuttley, W.M., A.M. Brade, C. Gaillardin, G.A. Eitzen, J.R. Glover, J.D. Aitchison, and R.A. Rachubinski. 2003. Rapid identification and characterization of peroxisomal assembly mutants in *Yarrowia lipolytica*. *Yeast*. 9:507–517.

Oeljeklaus, S., B.S. Reinartz, J. Wolf, S. Wiese, J. Tonillo, K. Podwojski, K. Kuhlmann, C. Stephan, H.E. Meyer, W. Schliebs, C. Brocard, R. Erdmann, and B. Warscheid. 2012. Identification of core components and transient interactors of the peroxisomal importomer by dual-track stable isotope labeling with amino acids in cell culture analysis. *J. Proteome Res.* 11:2567–2580.

Ohm, R.A., J.F. de Jong, L.G. Lugones, A. Aerts, E. Kothe, J.E. Stajich, R.P. de Vries, E. Record, A. Levasseur, S.E. Baker, K.A. Bartholomew, P.M. Coutinho, S. Erdmann, T.J. Fowler, A.C. Gathman,

V. Lombard, B. Henrissat, N. Knabe, U. Kües, W.W. Lilly, E. Lindquist, S. Lucas, J.K. Magnuson, F. Piumi, M. Raudaskoski, A. Salamov, J. Schmutz, F.W.M.R. Schwarze, P.A. vanKuyk, J.S. Horton, I.V. Grigoriev, and H.A.B. Wösten. 2010. Genome sequence of the model mushroom *Schizophyllum commune. Nature Biotechnol.* 28:949–955.

Ohm, R.A., N. Feau, B. Henrissat, C.L. Schoch, B.A. Horwitz, K.W. Barry, B.J. Condon, A.C. Copeland, B. Dhillon, F. Glaser, C.N. Hesse, I. Kosti, K. LaButti, E.A. Lindquist, S. Lucas, A.A. Salamov, R.E. Bradshaw, L. Ciuffetti, R.C. Hamelin, G.H.J. Kema, C. Lawrence, J.A. Scott, J.W. Spatafora, B.G. Turgeon, P.J.G.M. de Wit, S. Zhong, S.B. Goodwin, and I.V. Grigoriev. 2012. Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen dothideomycetes Fungi. *PLoS Pathog.* 8:e1003037.

Opaliński, Ł., J.A.K.W. Kiel, C. Williams, M. Veenhuis, and I.J. van der Klei. 2011. Membrane curvature during peroxisome fission requires Pex11. *EMBO J.* 30:5–16.

Opperdoes, F.R., P. Borst, S. Bakker, and W. Leene. 1977. Localization of glycerol-3-phosphate oxidase in the mitochondrion and particulate NAD+-linked glycerol-3-phosphate dehydrogenase in the microbodies of the bloodstream form of *Trypanosoma brucei*. *Eur. J. Biochem*. 76:29–39.

Orellana, M., R. Rodrigo, and E. Valdés. 1998. Peroxisomal and microsomal fatty acid oxidation in liver of rats after chronic ethanol consumption. *Gen. Pharmacol.* 31:817–820.

Otera, H., T. Harano, M. Honsho, K. Ghaedi, S. Mukai, A. Tanaka, A. Kawai, N. Shimizu, and Y. Fujiki. 2000. The mammalian peroxin Pex5pL, the longer isoform of the mobile peroxisome targeting signal (PTS) type 1 transporter, translocates the Pex7p–PTS2 protein complex into peroxisomes via its initial docking site, Pex14p. *J. Biol. Chem*. 275:21703–21714.

Otera, H., C. Wang, M.M. Cleland, K. Setoguchi, S. Yokota, R.J. Youle, and K. Mihara. 2010. Mff is an essential factor for mitochondrial recruitment of Drp1 during mitochondrial fission in mammalian cells. *J. Cell Biol.* 191:1141–1158.

Otzen, M., D. Wang, M.G. Lunenborg, and I.J. van der Klei. 2005. *Hansenula polymorpha* Pex20p is an oligomer that binds the peroxisomal targeting signal 2 (PTS2). *J. Cell Sci.* 118: 3409–3418.

Palenik, B., J. Grimwood, A. Aerts, P. Rouze, A. Salamov, N. Putnam, C. Dupont, R. Jorgensen, E. Derelle, S. Rombauts, K. Zhou, R. Otillar, S.S. Merchant, S. Podell, T. Gaasterland, C. Napoli, K. Gendler, A. Manuell, V. Tai, O. Vallon, G. Piganeau, S. Jancek, M. Heijde, K. Jabbari, C. Bowler, M. Lohr, S. Robbens, G. Werner, I. Dubchak, G.J. Pazour, Q. Ren, I. Paulsen, C. Delwiche, J. Schmutz, D. Rokhsar, Y. Van de Peer, H. Moreau, and I.V. Grigoriev. 2007. The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *PNAS*. 104:7705–7710.

Pereira-Leal, J.B. 2008. The Ypt/Rab Family and the Evolution of Trafficking in Fungi. *Traffic*. 9:27–38.

Perry, R.J., F.D. Mast, and R.A. Rachubinski. 2009. Endoplasmic reticulum-associated secretory proteins Sec20p, Sec39p, and Dsl1p are involved in peroxisome biogenesis. *Eukaryot. Cell.* 8:830–843.

Peterson, G.L. 1979. Review of the folin phenol protein quantitation method of Lowry, Rosebrough, Farr and Randall. *Anal. Biochem*. 100:201–220.

Pineda, M., M. Giro, F. Roels, M. Espeel, M. Ruiz, A. Moser, H.W. Moser, R.J.A. Wanders, C. Pavia, J. Conill, A. Aracil, L. Amat and T. Pampols. 1999. Diagnosis and follow-up of a case of peroxisomal disorder with peroxisomal mosaicism. *J. Child Neurol*. 14:434–439.

Piškur, J., Z. Ling, M. Marcet-Houben, O.P. Ishchuk, A. Aerts, K. LaButti, A. Copeland, E. Lindquist, K. Barry, C. Compagno, L. Bisson, I.V. Grigoriev, T. Gabaldón, and T. Phister. 2012. The genome of wine yeast *Dekkera bruxellnesis* provides a tool to explore its food-related properites. *Int. J. Food Microbiol*. 157:202–209.

Platta, H.W., and R. Erdmann. 2007. Peroxisomal dynamics. *Trends Cell Biol*. 17:474–484.

Platta, H.W., and R. Erdmann. 2013. Import of proteins into the peroxisomal matrix. *Front. Physiol*. 4: 1–12.

Platta, H.W., S. Grunau, K. Rosenkranz, W. Girzalsky, and R. Erdmann. 2005. Functional role of the AAA peroxins in dislocation of the cycling PTS1 receptor back to the cytosol. *Nat. Cell Biol.* 7:817-822.

Platta, H.W., F. El Magraoui, B.E. Bäumer, D. Schlee, W. Girzalsky, and R. Erdmann. 2009. Pex2 and Pex12 function as protein-ubiquitin ligases in peroxisomal import. *Mol. Cell. Biol.* 29:5505-5516.

Platta, H.W., F. El. Magraoui, D. Schlee, S. Grunau, W. Girzalsky, and R. Erdmann. 2007. Ubiquitination of the peroxisomal import receptor Pex5p is required for its recycling. *J. Cell. Biol.* 117:197-204.

Poopanitpan, N., S. Kobayashi, R. Fukuda, H. Horiuchi, and A. Ohta. 2010. An ortholog of farA of *Aspergillus nidulans* is implicated in the transcriptional activation of genes involved in fatty acid utilization in the yeast *Yarrowia lipolytica*. *Biochem. Biophys. Res. Commun.* 402:731–735.

Praefcke, G.J.K., and H.T. McMahon. 2004. The dynamin superfamily: universal membrane tubulation and fission molecules? *Nature*. 5:133–147.

Pringle, J.R., A.E. Adams, D.G. Drubin, and B.K. Haarer. 1991. Immunofluorescence methods for yeast. *Methods in Enzymol*. 194:565–602.

Prochnik, S.E., J. Umen, A.M. Nedelcu, A. Hallmann, S.M. Miller, I. Nishii, P. Ferris, A. Kuo, T. Mitros, L.K. Fritz-Laylin, U. Hellsten, J. Chapman, O. Simakov, S.A. Rensing, A. Terry, J. Pangilinan, V. Kapitonov, J. Jurka, A. Salamov, H. Shapiro, J. Schmutz, J. Grimwood, E. Lindquist, S. Lucas, I.V. Grigoriev, R. Schmitt, D. Kirk, and D.S. Rokhsar. 2010. Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. *Science*. 329:223–226.

Purdue, P.E., X. Yang, and P.B. Lazarow. 1998. Pex18p and Pex21p, a novel pair of related peroxins essential for peroxisomal targeting by the PTS2 pathway. *J. Cell Biol*. 143:1859–1869.

Putnam, N.H., M. Srivastava, U. Hellsten, B. Dirks, J. Chapman, A. Salamov, A. Terry, H. Shapiro, E. Lindquist, V.V. Kapitonov, J. Jurka, G. Genikhovich, I.V. Grigoriev, S.M. Lucas, R.E. Steele, J.R. Finnerty, U. Technau, M.Q. Martindale, and D.S. Rokhsar. 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science*. 317:86–94.

Pyper, S.R., N. Viswakarma, S. Yu, and J.K. Reddy. 2010. PPARalpha: energy combustion, hypolipidemia, inflammation and cancer. *Nucl. Recept. Signal*. 8:e002.

Rakhshandehroo, M., B. Knoch, M. Muller, and S. Kersten. 2010. Peroxisome proliferator-activated receptor alpha target genes. *PPAR Res.* 2010.

Ramirez, M.A., and M.C. Lorenz. 2009. The transcription factor homolog CTF1 regulates beta-oxidation in *Candida albicans*. *Eukaryot. Cell.* 8:1604–1614.

Rapoport, T.A. 2007. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature*. 450:663-669.

Read, B.A., J. Kegel, M.J. Klute, A. Kuo, S.C. Lefebvre, F. Maumus, C. Mayer, J. Miller, A. Monier, A. Salamov, J. Young, M. Aguilar, J.-M. Claverie, S. Frickenhaus, K. Gonzalez, E.K. Herman, Y.-C. Lin, J. Napier, H. Ogata, A.F. Sarno, Jeremy Shmutz, D. Schroeder, C. de Vargas, F. Verret, P. von Dassow, KlausValentin, Y. de Peer, G. Wheeler, *Emiliania huxleyi* Annotation Consortium, J.B. Dacks, C.F. Delwiche, S.T. Dyhrman, G. Glockner, U. John, T. Richards, A.Z. Worden, Z. Xin, and I.V. Grigoriev. 2013. Pan genome of the phytoplankton *Emiliania* underpins its global distribution. *Nature*. 499:209–213.

Reumann, S., S. Quan, K. Aung, P. Yang, K. Manandhar-Shrestha, D. Holbrook, N. Linka, R. Switzenberg, C.G. Wilkerson, A.P. Weber, L.J. Olsen, and J. Hu. 2009. In-depth proteome analysis of *Arabidopsis* leaf peroxisomes combined with *in vivo* subcellular targeting verification indicates novel metabolic and regulatory functions of peroxisomes. *Physiol.* 150:125-143.

Roberg, K.J., N. Rowley, and C.A. Kaiser. 1997. Physiological regulation of membrane protein sorting late in the secretory pathway of *Saccharomyces cerevisiae*. *J. Cell Biol.* 137:1469-1482.

Ronquist, F., and J.P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. 19:1572–1574.

Rose, M.D., W. F, and P. Heiter. 1988. Laboratory Course Manual for Methods in Yeast Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor.

Rottensteiner, H., A. Hartig, B. Hamilton, H. Ruis, R. Erdmann, and A. Gurvitz. 2003. *Saccharomyces cerevisiae* Pip2p-Oaf1p regulates *PEX25* transcription through an adenine-less ORE. *Eur. J. Biochem*. 270:2013–2022.

Rottensteiner, H., A. Kramer, S. Lorenzen, K. Stein, C. Landgraf, R. Volkmer-Engert, and R. Erdmann. 2004. Peroxisomal membrane proteins contain common Pex19-binding sites that are an integral part of their targeting signals. *Mol. Biol. Cell.* 15:3406-3417.

Rottensteiner, H., L. Palmier, A. Hartig, B. Hamilton, H. Ruis, R. Erdmann, and A. Gurvitz. 2002. The peroxisomal transporter gene *ANT1* is regulated by a deviant oleate response element (ORE): characterization of the signal for fatty acid induction. *Biochem. J.* 365:109–117.

Rucktäschel, R., W. Girzalsky, and R. Erdmann. 2011. Protein import machineries of peroxisomes. *Biochim. Biophys. Acta*. 1808:892–900.

Saiki, R.K. 1990. Amplification of genomic DNA. *In* PCR Protocols: A guide to Methods and Applications. M.A. Innis, D.H. Gelfand, J.J. Sninsky, and T.J. White, editors. Academic Press, San Diego. 13–21.

Sacksteder, K.A., J.M. Jones, S.T. South, X. Li, Y. Liu, and S.J. Gould. 2000. PEX19 binds multiple peroxisomal membrane proteins, is predominantly cytoplasmic, and is required for peroxisome

membrane synthesis. *J. Cell Biol.* 148:931–944.

Salomons, F.A., I.J. van der Klei, A.M. Kram, W. Harder, and M. Veenhuis. 1997. Brefeldin A interferes with peroxisomal protein sorting in the yeast *Hansenula polymorpha*. *FEBS Lett.* 411:133–139.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* 74:5463–5467.

Santarella-Mellwig, R., J. Franke, A. Jaedicke, M. Gorjanacz, U. Bauer, A. Budd, I.W. Mattaj, and D.P. Devos. 2010. The compartmentalized bacteria of the Planctomycetes-Verrucomicrobia-Chlamydiae superphylum have membrane coat-like proteins. PLoS *Biol.* 8:e000281.

Scannell, D.R., A.C. Frank, G.C. Conant, K.P. Byrne, M. Woolfit, and K.H. Wolfe. 2007. Independent sorting-out of thousands of duplicated gene pairs in two yeast species descended froma whole-genome duplication. *PNAS*. 104:8397–8402.

Schekman, R. 2005. Peroxisomes: another branch of the secretory pathway? *Cell*. 122:1–2.

Schlüter, A., A. Real-Chicharro, T. Gabaldón, F. Sanchez-Jimenez, and A. Pujol. 2010. PeroxisomeDB 2.0: an integrative view of the global peroxisomal metabolome. *Nucleic Acids Res*. 38.

Schlüter, A., S. Fourcade, R. Ripp, J.L. Mandel, O. Poch, and A. Pujol. 2006. The evolutionary origin of peroxisomes: an ER-peroxisome connection. *Mol. Biol. Evol*. 23:838–845.

Schrader, M., B.E. Reuber, J.C. Morrell, G. Jimenez-Sanchez, C. Obie, T.A. Stroh, D. Valle, T.A. Schroer, and S.J. Gould. 1998. Expression of *PEX11* mediates peroxisome proliferation in the absence of extracellular stimuli. *J. Biol. Chem*. 273:29607–29614.

Schrader, M. 2006. Shared components of mitochondrial and peroxisomal division. *Biochim. Biophys. Acta*. 1763:531–541.

Schrader, M., and Y. Yoon. 2007. Mitochondria and peroxisomes: Are the 'big brother' and the 'little sister' closer than assumed? *Bioessays*. 29:1105–1114.

Schrader, M., N.A. Bonekamp, and M. Islinger. 2012. Fission and proliferation of peroxisomes. *Biochim. Biophys. Acta. Mol. Basis Dis*. 1822:1343–1357.

Schoch, C.L., P.W. Crous, J.Z. Groenewald, E.W.A. Boehm, T.I. Burgess, J. de Gruyter, G.S. de Hoog, L.J. Dixon, M. Grube, C. Gueidan, Y. Harada, S. Hatakeyama, K. Hirayama, T. Hosoya, S.M. Huhndorf, K.D. Hyde, E.B.G. Jones, J. Kohlmeyer, A. Kruys, Y.M. Li, R. Lucking, H.T. Lumbsch, L. Marvanova, J.S. Mbatchou, A.H. McVay, A.N. Miller, G.K. Mugambi, L. Muggia, M.P. Nelsen, P. Nelson, C.A. Owensby, A.J.L. Phillips, S. Phongpaichit, S.B. Pointing, V. Pujade-Renaud, H.A. Raja, E.R. Plata, B. Robbertse, C. Ruibal, J. Sakayaroj, T. Sano, L. Selbmann, C.A. Shearer, T. Shirouzu, B. Slippers, S. Suetrong, K. Tanaka, B. Volkmann-Kohlmeyer, M.J. Wingfield, A.R. Wood, J.H.C. Woudenberg, H. Yonezawa, Y. Zhang, and J.W. Spatafora. 2010. A class-wide phylogenetic assessment of Dothideomycetes. *Stud. Mycol*. 64:1–15.

Schuldiner, M., J. Metz, V. Schmid, V. Denic, M. Rakwalska, H.D. Schmitt, B. Schwappach, and J.S. Weissman. 2008. The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell*. 134:634-645.

Schuller, H.-J. 2003. Transcriptional control of nonfermentative metabolism in the yeast Saccharomyces cerevisiae. *Curr. Genet.* 43:139–160.

Shimizu, M., A. Takeshita, T. Tsukamoto, F.J. Gonzalez, and T. Osumi. 2004. Tissue-selective, bidirectional regulation of PEX11 alpha and perilipin genes through a common peroxisome proliferator response element. *Mol. Cell. Biol.* 24:1313–1323.

Shimozawa, N., A. Imamura, Z. Zhang, Y. Suzuki, T. Orii, T. Tsukamoto, T. Osumi, Y. Fujiki, R.J.A. Wanders, G. Besley, and N. Kondo. 1999. Defective *PEX* gene products correlate with the protein import, biochemical abnormalities, and phenotypic heterogeneity in peroxisome biogenesis disorders. *J. Med. Genet.* 36:779–781.

Shinozaki, A., N. Sato, and Y. Hayashi. 2009. Peroxisomal targeting signals in green algae. *Protoplasma*. 235:57–66.

Sichting, M., A. Schell-Steven, H. Prokisch, R. Erdmann, and H. Rottensteiner. 2003. Pex7p and Pex20p of *Neurospora crassa* function together in PTS2-dependent protein import into peroxisomes. *Mol. Biol. Cell.* 14:810–821.

Sigrist, C.J.A., L. Cerutti, N. Hulo, A. Gattiker, L. Falquet, M. Pagni, A. Bairoch, and P. Bucher. 2002. PROSITE: A documented database using patterns and profiles as motif descriptors. *Brief. Bioinform.* 3:265–274.

Smith, J.J., and J.D. Aitchison. 2013. Peroxisomes take shape. *Nature*. 14:803–817.

Smith, J.J., M. Marelli, R.H. Christmas, F.J. Vizeacoumar, D.J. Dilworth, T. Ideker, T. Galitski, K. Dimitrov, R.A. Rachubinski, and J.D. Aitchison. 2002. Transcriptome profiling to identify genes involved in peroxisome assembly and function. *J. Cell Biol.* 158:259–271.

Smith, D.W., J.M. Opitz, and S.L. Inhorn. 1965. A syndrome of multiple developmental defects including polycystic kidneys and intrahepatic biliary dysgenesis in two siblings. *J. Pediatr.* 67:617-629.

Souppart L., H. Moussa, A. Cian, G. Sanciu, P. Poirier, H. El Alaoui , F. Delbac, K. Boorom, L. Delhaes, E. Dei-Cas, and E. Viscogliosi. 2010. Subtype analysis of *Blastocystis* isolates from symptomatic patients in Egypt. *Parasitol. Res.* 106:505-511.

South, S.T., K.A. Sacksteder, X. Li, Y. Liu, and S.J. Gould. 2000. Inhibitors of COPI and COPII do not block PEX3-mediated peroxisome synthesis. *J. Cell Biol.* 149:1345–1360.

Speijer, D. 2013. Reconsidering ideas regarding the evolution of peroxisomes: the case for a mitochondrial connection. *Cell. Mol. Life Sci.*

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 22:2688–2690.

Steinberg, S., L. Chen, L. Wei, A. Moser, H. Moser, G. Cutting, and N.E. Braverman. 2004. The *PEX* gene screen: molecular diagnosis of peroxisome biogenesis disorders in the Zellweger syndrome spectrum. *Mol. Genet. Metab.* 83:252–263.

Steinberg, S.J., G. Dodt, G.V. Raymond, N.E. Braverman, A.B. Moser, and H.W. Moser. 2006. Peroxisome biogenesis disorders. *Biochim. Biophys. Acta.* 1763:1733–1748.

Subramani, S. 2002. Hitchhiking fads en route to peroxisomes. *J. Cell Biol*. 156:415–417.

Suh, S.-O., M. Blackwell, C.P. Kurtzman, and M.A. Lachance. 2006. Phylogenetics of Saccharomycetales, the ascomycete yeasts. *Mycologia*. 98:1006–1017.

Swinkels, B.W., S.J. Gould, A.G. Bodnar, R.A. Rachubinski, and S. Subramani. 1991. A novel, cleavable peroxisomal targeting signal at the amino-terminus of the rat 3-ketoacyl-CoA thiolase. *EMBO J*. 10:3255–3262.

Szilard, R.K. 2000. Identification and Characterization of *Yarrowia lipolytica* Pex5p, a Component of the Peroxisomal Translocation Machinery of *Yarrowia lipolytica*. University of Alberta.

Tam, Y.Y.C. 2005. Early to late events in peroxisome biogenesis. University of Alberta.

Tam, Y.Y.C., A. Fagarasanu, M. Fagarasanu, and R.A. Rachubinski. 2005. Pex3p Initiates the formation of a preperoxisomal compartment from a subdomain of the endoplasmic reticulum in *Saccharomyces cerevisiae*. *J. Biol. Chem*. 280:34933–34939.

Tam, Y.Y.C., and R.A. Rachubinski. 2002. *Yarrowia lipolytica* cells mutant for the *PEX24* gene encoding a peroxisomal membrane peroxin mislocalize peroxisomal proteins and accumulate membrane structures containing both peroxisomal matrix and membrane proteins. *Mol. Biol. Cell*. 13:2681–2691.

Tam, Y.Y.C., J.C. Torres-Guzman, F.J. Vizeacoumar, J.J. Smith, M. Marelli, J.D. Aitchison, and R.A. Rachubinski. 2003. Pex11-related proteins in peroxisome dynamics: a role for the novel peroxin Pex27p in controlling peroxisome size and number in *Saccharomyces cerevisiae*. *Mol. Biol. Cell*. 14:4089–4102.

Tan, K.S.W. 2008. New insights on classification, identification, and clinical relevance of *Blastocystis* spp. *Clin. Microbiol. Rev*. 21:639–665.

The *Arabidopsis* Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. 408:796–815.

The Genolevures Consortium, J.L. Souciet, B. Dujon, C. Gaillardin, M. Johnston, P.V. Baret, P. Cliften, D.J. Sherman, J. Weissenbach, E. Westhof, P. Wincker, C. Jubin, J. Poulain, V. Barbe, B. Segurens, F. Artiguenave, V. Anthouard, B. Vacherie, M.E. Val, R.S. Fulton, P. Minx, R. Wilson, P. Durrens, G. Jean, C. Marck, T. Martin, M. Nikolski, T. Rolland, M.L. Seret, S. Casaregola, L. Despons, C. Fairhead, G. Fischer, I. Lafontaine, V. Leh, M. Lemaire, J. de Montigny, C. Neuveglise, A. Thierry, I. Blanc-Lenfle, C. Bleykasten, J. Diffels, E. Fritsch, L. Frangeul, A. Goeffon, N. Jauniaux, R. Kachouri-Lafond, C. Payen, S. Potier, L. Pribylova, C. Ozanne, G.F. Richard, C. Sacerdot, M.L. Straub, and E. Talla. 2009. Comparative genomics of protoploid Saccharomycetaceae. *Genome Res*. 19:1696–1709.

Thurman, R.G., and W.R. McKenna. 1975. Pathways of ethanol metabolism in perfused rat liver. *Adv Exp Med Biol*. 56:57–76

Titorenko, V.I., H. Chan, and R.A. Rachubinski. 2000. Fusion of small peroxisomal vesicles in vitro reconstructs an early step in the *in vivo* multistep peroxisome assembly pathway of *Yarrowia lipolytica*. *J Cell Biol*. 148:29–43.

Titorenko, V.I., and R.A. Rachubinski. 1998. Mutants of the yeast *Yarrowia lipolytica* defective in

protein exit from the endoplasmic reticulum are also defective in peroxisome biogenesis. *Mol. Cell. Biol.* 18:2789–2803.

Titorenko, V.I., J.J. Smith, R.K. Szilard, and R.A. Rachubinski. 1998. Pex20p of the yeast *Yarrowia lipolytica* is required for the oligomerization of thiolase in the cytosol and for its targeting to the peroxisome. *J. Cell Biol.* 142:403–420.

Titorenko, V.I., J.M. Nicaud, H. Wang, H. Chan, and R.A. Rachubinski. 2002. Acyl- CoA oxidase is imported as a heteropentameric, cofactor-containing complex into peroxisomes of *Yarrowia lipolytica*. *J. Cell Biol.* 156:481–494.

Toome, M., R.A. Ohm, R.W. Riley, T.Y. James, K.L. Lazarus, B. Henrissat, S. Albu, A. Boyd, J. Chow, A. Clum, G. Heller, A. Lipzen, M. Nolan, L. Sandor, N. Zvenigorodsky, I.V. Grigoriev, J.W. Spatafora, and M.C. Aime. 2013. Genome sequencing provides insight into the reproductive biology, nutritional mode and ploidy of the fern pathogen *Mixia osmundae*. *New Phytol.*

Toro, A.A., C.A. Araya, G.J. Cordova, C.A. Arredondo, H.G. Cardenas, R.E. Moreno, A. Venegas, C.S. Koenig, J. Cancino, A. Gonzalez, and M.J. Santos. 2009. Pex3p-dependent peroxisomal biogenesis initiates in the endoplasmic reticulum of human fibroblasts. *J. Cell. Biochem.* 107:1083–1096.

Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci.* 76:4350–4354.

Tower, R.J., A. Fagarasanu, J.D. Aitchison, and R.A. Rachubinski. 2011. The peroxin Pex34p functions with the Pex11 family of peroxisomal divsional proteins to regulate the peroxisome population in yeast. *Mol. Biol. Cell.* 22:1727–1738.

Tower, R.J. 2013. The peroxin Pex34p functions with the Pex11 family of proteins to regulate the peroxisome population in the yeast *Saccharomyces cerevisiae*. University of Alberta.

Tuskan, G.A., S. DiFazio, S. Jansson, J. Bohlmann, I. Grigoriev, U. Hellsten, N. Putnam, S. Ralph, S. Rombauts, A. Salamov, J. Schein, L. Sterck, A. Aerts, R.R. Bhalerao, R.P. Bhalerao, D. Blaudez, W. Boerjan, A. Brun, A. Brunner, V. Busov, M. Campbell, J. Carlson, M. Chalot, J. Chapman, G.L. Chen, D. Cooper, P.M. Coutinho, J. Couturier, S. Covert, Q. Cronk, R. Cunningham, J. Davis, S. Degroeve, A. Dejardin, C. dePamphilis, J. Detter, B. Dirks, I. Dubchak, S. Duplessis, J. Ehlting, B. Ellis, K. Gendler, D. Goodstein, M. Gribskov, J. Grimwood, A. Groover, L. Gunter, B. Hamberger, B. Heinze, Y. Helariutta, B. Henrissat, D. Holligan, R. Holt, W. Huang, N. Islam-Faridi, S. Jones, M. Jones-Rhoades, R. Jorgensen, C. Joshi, J. Kangasjarvi, J. Karlsson, C. Kelleher, R. Kirkpatrick, M. Kirst, A. Kohler, U. Kalluri, F. Larimer, J. Leebens-Mack, J.C. Leple, P. Locascio, Y. Lou, S. Lucas, F. Martin, B. Montanini, C. Napoli, D.R. Nelson, C. Nelson, K. Nieminen, O. Nilsson, V. Pereda, P.J. Myler, R. Philippe, G. Pilate, A. Poliakov, J. Razumovskaya, P. Richardson, C. Rinaldi, K. Ritland, P. Rouze, D. Ryaboy, J. Schmutz, J. Schrader, B. Segerman, H. Shin, A. Siddiqui, F. Sterky, A. Terry, C.J. Tsai, E. Uberbacher, P. Unneberg, J. Vahala, K. Wall, S. Wessler, G. Yang, T. Yin, C. Douglas, M. Marra, G. Sandberg, Y. Van de Peer, and D. Rokhsar. 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*. 313:1596–1604.

Tyler, B.M., S. Tripathy, X. Zhang, P. Dehal, R.H.Y. Jiang, A. Aerts, F.D. Arredondo, L. Baxter, D. Bensasson, J.L. Beynon, J. Chapman, C.M.B. Damasceno, A.E. Dorrance, D. Dou, A.W. Dickerman, I.L. Dubchak, M. Garbelotto, M. Gijzen, S.G. Gordon, F. Govers, N.J. Grunwald, W. Huang, K.L. Ivors, R.W. Jones, S. Kamoun, K. Krampis, K.H. Lamour, M.-K. Lee, W.H. McDonald,

M. Medina, H.J.G. Meijer, E.K. Nordberg, D.J. Maclean, M.D. Ospina-Giraldo, P.F. Morris, V. Phuntumart, N.H. Putnam, S. Rash, J.K.C. Rose, Y. Sakihama, A.A. Salamov, A. Savidor, C.F. Scheuring, B.M. Smith, B.W.S. Sobral, A. Terry, T.A. Torto-Alalibo, J. Win, Z. Xu, H. Zhang, I.V. Grigoriev, D.S. Rokhsar, and J.L. Boore. 2006. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science*. 313:1261–1266.

van den Berg, M.A., R. Albang, K. Albermann, J.H. Badger, J.-M. Daran, A.J. M Driessen, C. Garcia-Estrada, N.D. Fedorova, D.M. Harris, W.H.M. Heijne, V. Joardar, J.A.K. W Kiel, A. Kovalchuk, J.F. Martín, W.C. Nierman, J.G. Nijland, J.T. Pronk, J.A. Roubos, I.J. van der Klei, N.N.M.E. van Peij, M. Veenhuis, H. von Döhren, C. Wagner, J. Wortman, and R.A.L. Bovenberg. 2008. Genome sequencing and analysis of the filamentous fungus *Penicillium chrysogenum*. *Nature Biotechnol*. 26:1161–1168.

van der Klei, I.J., and M. Veenhuis. 2006. PTS1-independent sorting of peroxisomal matrix proteins by Pex5p. *Biochim. Biophys. Acta. Mol. Cell Res*. 1763:1794–1800.

van der Zand, A., I. Braakman, and H.F. Tabak. 2010. Peroxisomal membrane proteins insert into the endoplasmic reticulum. *Mol. Biol. Cell*. 21:2057–2065.

van der Zand, A., J. Gent, I. Braakman, and H.F. Tabak. 2012. Biochemically distinct vesicles from the endoplasmic reticulum fuse to form peroxisomes. *Cell*. 149:397–409.

van Roermund, C.W.T., H.F. Tabak, M. van den Berg, R.J.A. Wanders, and E.H. Hettema. 2000. Pex11p plays a primary role in medium-chain fatty acid oxidation, a process that affects peroxisome number and size in *Saccharomyces cerevisiae*. *J. Cell Biol*. 150:489–497.

Veenhuis, M., and I.J. van der Klei. 2014. A critical reflection on the principles of peroxisome formation in yeast. *Front. Physiol*. 5:1–8.

Veenhuis, M., M. Mateblowski, W.H. Kunau, and W. Harder. 1987. Proliferation of Microbodies in *Saccharomyces cerevisiae*. *Yeast*. 3:77–84.

Venter, J.C., M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, J.D. Gocayne, P. Amanatides, R.M. Ballew, D.H. Huson, J.R. Wortman, Q. Zhang, C.D. Kodira, X.H. Zheng, L. Chen, M. Skupski, G. Subramanian, P.D. Thomas, J. Zhang, G.L.G. Miklos, C. Nelson, S. Broder, A.G. Clark, J. Nadeau, V.A. McKusick, N. Zinder, A.J. Levine, R.J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A.E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T.J. Heiman, M.E. Higgins, R.-R. Ji, Z. Ke, K.A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G.V. Merkulov, N. Milshina, H.M. Moore, A.K. Naik, V.A. Narayan, B. Neelam, D. Nusskern, D.B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z.Y. Wang, A. Wang, X. Wang, J. Wang, M.-H. Wei, R. Wides, C. Xiao, C. Yan, A. Yao, J. Ye, M. Zhan, W. Zhang, H. Zhang, Q. Zhao, L. Zheng, F. Zhong, W. Zhong, S.C. Zhu, S. Zhao, D. Gilbert, S. Baumhueter, G. Spier, C. Carter, A. Cravchik, T. Woodage, F. Ali, H. An, A. Awe, D. Baldwin, H. Baden, M. Barnstead, I. Barrow, K. Beeson, D. Busam, A. Carver, A. Center, M.L. Cheng, L. Curry, S. Danaher, L. Davenport, R. Desilets, S. Dietz, K. Dodson, L. Doup, S. Ferriera, N. Garg, A. Gluecksmann, B. Hart, J. Haynes, C. Haynes, C. Heiner, S. Hladun, D. Hostin, J. Houck, T. Howland, C. Ibegwam, J. Johnson, F. Kalush, L. Kline, S. Koduru, A. Love, F. Mann, D. May, S. McCawley, T. McIntosh, I. McMullen, M. Moy, L. Moy, B. Murphy, K. Nelson, C. Pfannkoch, E. Pratts, V. Puri, H. Qureshi, M. Reardon, R.

Rodriguez, Y.-H. Rogers, D. Romblad, B. Ruhfel, R. Scott, C. Sitter, M. Smallwood, E. Stewart, R. Strong, E. Suh, R. Thomas, N.N. Tint, S. Tse, C. Vech, G. Wang, J. Wetter, S. Williams, M. Williams, S. Windsor, E. Winn-Deen, K. Wolfe, J. Zaveri, K. Zaveri, J.F. Abril, R. Guigó, M.J. Campbell, K.V. Sjolander, B. Karlak, A. Kejariwal, H. Mi, B. Lazareva, T. Hatton, A. Narechania, K. Diemer, A. Muruganujan, N. Guo, S. Sato, V. Bafna, S. Istrail, R. Lippert, R. Schwartz, B. Walenz, S. Yooseph, D. Allen, A. Basu, J. Baxendale, L. Blick, M. Caminha, J. Carnes-Stine, P. Caulk, Y.-H. Chiang, M. Coyne,C. Dahlke, A.D. Mays, M. Dombroski, M. Donnelly, D. Ely, S. Esparham, C. Fosler, H. Gire, S. Glanowski, K. Glasser, A. Glodek, M. Gorokhov, K. Graham, B. Gropman, M. Harris, J. Heil, S. Henderson, J. Hoover, D. Jennings, C. Jordan, J. Jordan, J. Kasha, L. Kagan, C. Kraft, A. Levitsky, M. Lewis, X. Liu, J. Lopez, D. Ma, W. Majoros, J. McDaniel, S. Murphy, M. Newman, T. Nguyen, N. Nguyen, M. Nodell, S. Pan, J. Peck, M. Peterson, W. Rowe, R. Sanders, J. Scott, M. Simpson, T. Smith, A. Sprague, T. Stockwell, R. Turner, E. Venter, M. Wang, M. Wen, D. Wu, M. Wu, A. Xia, A. Zandieh, and X. Zhu. 2001. The sequence of the human genome. *Science*. 291:1304–1351.

Visvesvara, G.S., H. Moura, and F.L. Schuster. 2007. Pathogenic and opportunistic free-living amoebae: *Acanthamoeba* spp., *Balamuthia mandrillaris*, *Naegleria fowleri*, and *Sappinia diploidea*. *FEMS Immunol. Med. Mic*. 50:1–26.

Visvesvara, G.S., and J.K. Stehr-Green. 1990. Epidemiology of free-living ameba infections. *J. Protozool.* 37:25s–33s.

Vizeacoumar, F.J., J.C. Torres-Guzman, D. Bouard, J.D. Aitchison, and R.A. Rachubinski. 2004. Pex30p, Pex31p, and Pex32p form a family of peroxisomal integral membrane proteins regulating peroxisome size and number in *Saccharomyces cerevisiae*. *Mol. Biol. Cell*. 15:665–677.

Vizeacoumar, F.J., J.C. Torres-Guzman, Y.Y.C. Tam, J.D. Aitchison, and R.A. Rachubinski. 2003. YHR150w and YDR479c encode peroxisomal integral membrane proteins involved in the regulation of peroxisome number, size, and distribution in *Saccharomyces cerevisiae. J. Cell Biol*. 161:321–332.

Voncken, F., J.J. van Hellemond, I. Pfisterer, A. Maier, S. Hillmer, and C. Clayton. 2003. Depletion of *GIM5* causes cellular fragility, a decreased glycosome number, and reduced levels of ether-linked phospholipids in trypanosomes. *J. Biol. Chem*. 278:35299–35310.

Voorn-Brouwer, T., A. Kragt, H.F. Tabak, and B. Distel. 2001. Peroxisomal membrane proteins are properly targeted to peroxisomes in the absence of COPI- and COPII- mediated vesicular transport. *J. Cell Sci*. 114:2199–2204.

Walker, G., R. Dorrell, A. Schlacht, and J.B. Dacks. 2011. Eukaryotic systematics: a user's guide for cell biologists and parasitologists. *Parasitology*. 138:1638-1663.

Wanders, R.J.A., and H.R. Waterham. 2006. Biochemistry of mammalian peroxisomes revisited. *Ann. Rev. Biochem*. 75:295–323.

Wanders, R.J., P. Vreken, S. Ferdinandusse, G.A. Jansen, H.R. Waterham, C.W. van Roermund, and E.G. Van Grunsven. 2001. Peroxisomal fatty acid alpha- and beta-oxidation in humans: enzymology, peroxisomal metabolite transporters and peroxisomal diseases. *Biochem. Soc. Trans*. 29:250–267.

Wang, H., and X.W. Deng. 2003. Dissecting the phytochrome A-dependent signaling network in higher plants. *Trends Plant Sci*. 8:172-178.

Wang, J.-H., W. Hung, and S.-H. Tsai. 2011. High efficiency transformation by electroporation of *Yarrowia lipolytica*. *J. Microbiol.* 49:469–472.

Wapinski, I., A. Pfeffer, N. Friedman, and A. Regev. 2007. Natural history and evolutionary principles of gene duplication in fungi. *Nature*. 449:54–61.

Wardinsky, T.D., R.A. Pagon, B.R. Powell, B. McGillivray, M. Stephan, J. Zonana, and A. Moser. 1990. Rhizomelic chondrodysplasia punctata and survival beyond one year: a review of the literature and five case reports. *Clin. Genet.* 38:84–93.

Waterham, H.R., and M.S. Ebberink. 2012. Genetics and molecular basis of human peroxisome biogenesis disorders. *Biochim. Biophys. Acta. Mol. Basis Dis.* 1822:1430–1441.

Waterham, H.R., J. Koster, C.W.T. van Roermund, P.A.W. Mooyer, R.J.A. Wanders, and J.V. Leonard. 2007. A lethal defect of mitochondrial and peroxisomal fission. *N. Engl. J. Med.* 356:1736–1741.

Weibel, E.R., and Bolender, P. 1973. Stereological techniques for electron microscopic morphometry. *In* Principles and Techniques of Electron Microscopy, Vol. 3. M.A. Hayat, editor. Van Nostrand Reinhold, New York. 237–296.

Weller, S., S.J. Gould, and D. Valle. 2003. Peroxisome biogenesis disorders. *Annu. Rev. Genomics Hum. Genet.* 4:165-211.

Wheeler, T.J., and S.R. Eddy. 2013. nhmmer: DNA homology search with profile HMMs. *Bioinformatics*. 29:2487–2489.

Wilcke, M., K. Hultenby, and S.E.H. Alexon. 1995. Novel peroxisome populations in subcellular fractions from rat liver. *J. Biol. Chem.* 270:6949–6958.

Williams, C., M. van der Berg, E. Geers, and B. Distel. 2008. Pex10p functions as an E(3) ligase for the Ubc4p-dependent ubiquitination of Pex5p. *Biochem. Biophys. Res. Commun.* 374:620-624.

Wilson, G.N., R.G. Holmes, J. Custer, J.L. Lipkowitz, J. Stover, N. Datta, and A. Hajra. 1986. Zellweger syndrome: diagnostic assays, syndrome delineation, and potential therapy. *Am. J. Med. Genet.* 24:69–82.

Windsor, J.J. 2007. *Blastocystis hominis* and *Dientamoeba fragilis*: neglected human protozoa. *The Biomedical Scientist*. 64:524-527.

Wohlbach, D.J., A. Kuo, T.K. Sato, K.M. Potts, A.A. Salamov, K.M. LaButti, H. Sun, A. Clum, J.L. Pangilinan, E.A. Lindquist, S. Lucas, A. Lapidus, M. Jin, C. Gunawan, V. Balan, B.E. Dale, T.W. Jeffries, R. Zinkel, K.W. Barry, I.V. Grigoriev, and A.P. Gasch. 2011. Comparative genomics of xylose-fermenting fungi for enhanced biofuel production. *PNAS*. 1–6.

Woodward, A.W., and B. Bartel. 2005. The *Arabidopsis* peroxisomal targeting signal type 2 receptor PEX7 is necessary for peroxisome function and dependent on PEX5. *Mol. Biol. Cell.* 16:573–583.

Worden, A.Z., J.-H. Lee, T. Mock, P. Rouze, M.P. Simmons, A.L. Aerts, A.E. Allen, M.L. Cuvelier, E. Derelle, M.V. Everett, E. Foulon, J. Grimwood, H. Gundlach, B. Henrissat, C. Napoli, S.M. McDonald, M.S. Parker, S. Rombauts, A. Salamov, P. Von Dassow, J.H. Badger, P.M. Coutinho, E. Demir, I. Dubchak, C. Gentemann, W. Eikrem, J.E. Gready, U. John, W. Lanier, E.A. Lindquist,

S. Lucas, K.F.X. Mayer, H. Moreau, F. Not, R. Otillar, O. Panaud, J. Pangilinan, I. Paulsen, B. Piegu, A. Poliakov, S. Robbens, J. Schmutz, E. Toulza, T. Wyss, A. Zelensky, K. Zhou, E.V. Armbrust, D. Bhattacharya, U.W. Goodenough, Y. Van de Peer, and I.V. Grigoriev. 2009. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science*. 324:268–272.

Wu, M., S. Chatterji, and J.A. Eisen. 2012. Accounting for alignment uncertainty in phylogenomics. *PLoS ONE*. 7:e30288.

Yan, M., N. Rayapuram, and S. Subramani. 2005. The control of peroxisome number and size during division and proliferation. *Curr. Opin. Cell Biol.* 17:376–383.

Appendices

Appendix 1: Pex11p homologues identified in comparative genomic survey

Genome	Accession	Length	R. pHMMer (*H. sapiens*)	R. pHMMer (*S. cerevisiae*)	R. pHMMer (Other genome)
Fungi					
Acremonium alcalophilum	2103663	236	1.90E-09	3.30E-23	
Agaricus bisporus	189714	248	1.40E-13	1.80E-18	
Allomyces macrogynus	AMAG_02953	241	3.10E-12	1.20E-11	
Ascoidea rubescens	124207	242	0.019	1.20E-34	
Ashbya gossypii	NP_987066.1	226	0.00048	9.60E-85	
Aspergillus fumigatus	Afu6g07740	236	9.70E-11	9.30E-26	
Aspergillus nidulans	EAA65086.1	236	5.60E-09	7.70E-24	
Aureobasidium pullulans	239497	237	1.30E-17	9.00E-30	
Batrachochytrium dendrobatidis	BDEG_03726	252	1.00E-10	6.10E-05	
Botryotinia fuckeliana	XP_001555247.1	234	8.70E-15	2.40E-26	
Candida albicans	XP_712077.1	245	5.70E-09	5.00E-47	
Candida glabrata	XP_445529.1	248	0.0022	5.00E-96	
Catenaria anguillulae	216502	252	6.60E-19	1.50E-18	
Cladonia grayi	22819	235	3.10E-17	5.90E-25	
Coccidioides immitis	CIMG_01269	247	3.00E-09	2.00E-20	
Cochliobolus heterostrophus	1142909	236	1.10E-08	1.10E-19	
Coemansia reversa	79930	261	2.20E-10	3.20E-12	
Conidiobolus coronatus	77384	242	5.50E-08	3.10E-16	
Coprinopsis cinerea	CC1G_03626	250	5.70E-09	2.40E-21	
Cryphonectria parasitica	104004	237	4.80E-11	6.20E-24	
Cryptococcus neoformans	CNAG_07341	271	9.30E-05	1.60E-15	
Dacryopinax sp.	117454	251	1.10E-07	2.60E-19	
Debaryomyces hansenii	XP_456578.1	225	3.70E-09	1.10E-44	
Gibberella zeae	XP_383974.1	203	0.0021	-	
	XP_389457.1	235	8.00E-10	5.60E-19	
Gonapodya prolifera	129839	232	4.20E-12	1.40E-12	

Hansenula polymorpha	16132	260	4.60E-07	1.30E-33
Hysterium pulicare	6088	234	8.90E-10	8.80E-23
Kluyveromyces lactis	XP_454032.1	234	-	2.40E-85
Kluyveromyces waltii	Kwal_22162	226	3.50E-06	7.00E-92
Laccaria bicolor	XP_001873857.1	249	4.20E-11	1.80E-22
Lipomyces starkeyi	43134	267	3.60E-07	-
Lodderomyces elongisporus	LELG_04882	234	2.60E-08	8.80E-44
Magnaporthe oryzae	MGG_08896	235	2.50E-13	3.70E-26
Malassezia globosa	3416	245	2.60E-16	1.30E-15
Metschnikowia bicuspidata	32487	229	1.10E-10	2.80E-50
Mixia osmundae	56139	278	7.30E-12	9.20E-19
Mucor circinelloides	156615	243	2.50E-13	2.00E-18
	33896	235	7.00E-13	1.60E-18
Myceliophthora thermophila	2306949	235	6.10E-11	4.40E-21
Mycosphaerella graminicola	41212	237	4.00E-10	2.60E-21
Neurospora crassa	NCU04802	235	4.30E-10	1.10E-23
Pachysolen tannophilus	32773	242	3.90E-09	1.80E-63
Penicillium chrysogenum	XP_002557767.1	238	2.50E-10	2.90E-25
Phaenerochaete chrysosporium	76645	178	0.00061	2.70E-06
Phaeosphaeria nodorum	SNOG_12438	236	4.80E-10	2.80E-23
Phycomyces blakesleeanus	147772	244	1.60E-11	1.60E-21
Pichia membranifaciens	36155	254	5.40E-10	8.60E-18
Pichia pastoris	XP_002491415.1	249	8.00E-10	1.80E-40
Pichia stipitis	59339	227	2.50E-09	1.50E-46
Piromyces sp.	67233	232	5.60E-15	3.90E-13
Puccinia graminis	PGTG_05241	250	1.00E-17	1.40E-20
Rhizophagus irregularis	39806	242	5.30E-16	2.40E-21
Rhizopus oryzae	RO3G_04385	244	9.10E-16	5.00E-22
Rhodotorula graminis	12042	269	2.10E-15	2.90E-22
Rhytidhysteron rufulum	5683	266	1.10E-08	3.20E-19

Rozella allomycis	-			
Saccharomyces bayanus	ORFP22202	238	6.80E-05	4.60E-149
Saccharomyces castelli	ORFPScas_Contig716.43	239	8.40E-05	8.20E-87
	ORFPScas_Contig589.6	244	0.0012	8.50E-84
Saccharomyces cerevisiae[a]	NP_014494.1	237	7.60E-05	0
Saccharomyces kluyveri	SAKL0C13816g	230	0.0028	5.10E-100
Saccharomyces mikatae	ORFP19868	237	9.40E-05	4.10E-150
Saccharomyces paradoxus	ORFP19767	237	2.00E-05	9.10E-156
Saitoella complicata	16014	242	2.00E-12	1.20E-19
Schizophyllum commune	2489226	337	2.40E-08	8.80E-20
Schizosaccharomyces pombe	SPCB582.09	239	8.60E-09	5.70E-07
Sclerotinia sclerotiorum	SS1G_02409	235	1.80E-15	1.00E-27
Spathaspora passalidarum	63314	229	9.00E-10	2.50E-40
Sphaerobolus stellatus	271675	245	4.80E-08	1.30E-20
Spizellomyces punctatus	SPPG_01476	240	1.60E-13	5.90E-14
Thielavia terrestris	2110343	238	9.30E-09	3.90E-21
Tremella mesenterica	68170	271	1.80E-10	2.70E-19
Trichoderma reesei	50131	235	1.50E-14	2.20E-21
Tuber melanosporum	4209	235	3.20E-13	8.20E-28
Ustilago maydis	UM04294	249	1.30E-14	3.70E-17
Vanderwaltozyma polyspora	XP_001643159.1	247	-	1.00E-91
Wickerhamomyces anomalus	27583	232	3.40E-10	7.80E-58
Xanthoria parietina	54943	236	2.00E-14	1.40E-23
Xylona heveae	284103	235	2.90E-15	1.60E-25
Yarrowia lipolytica	XP_501425.1	233	2.70E-13	3.10E-29
Zygosaccharomyces rouxii	XP_002495813.1	229	-	1.50E-89
Holozoa				
Caenorhabditis elegans	C47B2.8	214	1.20E-12	-
Canis familiaris	XP_003639010.1	247	9.40E-149	3.00E-08
	XP_533034.1	259	4.50E-166	0.00022

	XP_542110.1[c]	241	1.20E-136	-
Capsaspora owczarzaki	CAOG_02737	332	4.40E-09	2.40E-06
	CAOG_06484	243	8.70E-43	4.20E-08
	CAOG_07019	218	3.70E-23	5.10E-08
Danio rerio	NP_001025343.1[c]	240	1.60E-79	-
	NP_001096590.1	246	7.80E-75	5.20E-08
	NP_001039319.1	266	8.90E-108	0.00034
Drosophila melanogaster	CG8315	241	1.10E-36	5.40E-08
	CG13827	233	1.40E-32	8.50E-06
	CG33474	201	4.20E-12	-
Fonticula alba	H696_03870	269	-	0.0035
Homo sapiens[a]	NP_003838.1	247	0	1.00E-05
	NP_003837.1	259	0	0.00016
	NP_542393.1[c]	241	0	-
Monosiga brevicollis	9946	265	3.50E-06	7.20E-13
Mus musculus	NP_035198.1	246	3.90E-192	4.80E-07
	NP_035199.3	259	6.40E-162	0.0027
	NP_081227.1[c]	241	2.40E-126	-
Nematostella vectensis	240392	228	5.70E-28	-
	188278[c]	230	5.10E-52	-
Thecamonas trahens	AMSG_00420	264	1.70E-10	2.40E-08
Amoebozoa				
Dictyostelium discoideum	DDB0238054	255	2.10E-11	0.0032
Archaeplastida				
Arabidopsis thaliana	AT1G47750.1	249	1.60E-07	-
	AT3G47430.1	228	1.20E-07	-
	AT1G01820.1	236	-	0.0002
	AT2G45740.1	237	1.40E-03	-
	AT3G61070.1	232	1.00E-02	-
Chlamydomonas reinhardtii	Cre06.g263300.t1.2	235	0.00098	1.60E-08

	Cre12.g540500.t1.2[b,d]	238	3.80E-05	-	
Cyanidioschyzon merolae	CMN269C	333	0.0011	-	
Micromonas sp.	61774[b]	241	0.04	3.10E-05	
	57000	282	0.0014	-	
Oryza sativa	LOC_Os03g02590	238	-	4.00E-06	
	LOC_Os03g19000	255	7.40E-05	-	
	LOC_Os03g19010	243	4.20E-05	-	
	LOC_Os04g45210	223	0.00027	-	
	LOC_Os06g03660	426	0.038	-	
Ostreococcus tauri	24199	277	-	-	1.4E-14 (*Arabidopsis thaliana*) 9.2E-15 (*Chlamydomonas reinhardtii*)
	37965	180	-	-	7.3E-8 (*Arabidopsis thaliana*) 6.5E-18 (*Chlamydomonas reinhardtii*) 4.3E-6 (*Cyanidioschyzon merolae*) 2.4E-38 (*Micromonas* sp.) 5.1E-8 (*Populus trichocarpa*) 9.5E-18 (*Volvox carteri*)
Populus trichocarpa	Potri.002G134000.1	284	7.60E-05	-	
	Potri.001G124400.1	228	3.40E-08	-	
	Potri.014G076400.1	236	6.50E-03	0.00012	
Volvox carteri	Vocar20006405m[b,d]	237	1.70E-05	-	
	Vocar20002964m	233	0.00015	6.50E-09	
SAR					
Bigelowiella natans	91994	218	5.40E-12	0.0026	
	145768	239	2.90E-07	-	
	90664	218	5.40E-12	-	
	142022	231	4.30E-03	-	
Ectocarpus siliculosus	Esi0000_0032	250	2.50E-06	-	
	Esi0137_0087	298	-	0.018	
Phytophthora ramourum	82744[c]	244	2.60E-09	-	

Tetrahymena thermophila	TTHERM_00355720[b]	384	0.016	-	
	TTHERM_00703430	250	0.00023	2.00E-06	
	TTHERM_00522160	232	3.20E-07	6.60E-13	
	TTHERM_00051990	369	-	9.70E-04	
	TTHERM_00145910	252	-	1.90E-02	
Toxoplasma gondii	TGGT1_046160	260	0.0019	3.60E-07	
CCTH					
Guillardia theta	163884	262	2.30E-05	-	
	154547[c]	240	1.10E-05	-	
Excavata					
Bodo saltans	BS66630	216	-	-	3.8E-57 (*Trypanosoma brucei*)
	BS50905[e]	707	0.0089	-	7.2E-39 (*Trypanosoma brucei*)
Crithidia fasciculata	KB217518-1-27940-28887	316	-	-	3E-13 (*Trypanosoma brucei*) 1.7E-27 (*Trypanosoma cruzi*) 3.9E-108 (*Leishmania major*)
	KB217503-1-116308-116991	228	-	-	1.7E-78 (*Trypanosoma brucei*) 1.4E-86 (*Trypanosoma cruzi*) 6.2E-119 (*Leishmania major*)
	KB217503-2-114356-115084	243	-	-	4.5E-20, (*Trypanosoma brucei*) 3.8E-16 (*Trypanosoma cruzi*) 4.4E-61 (*Leishmania major*)
	KB217487-1-565189-566205[e]	339	-	-	5.9E-82 (*Trypanosoma brucei*) 4.8E-81 (*Trypanosoma cruzi*) 6.2E-126 (*Leishmania major*)
Leishmania major	LmjF.28.2260	222	-	-	3.9E-17 (*Trypanosoma brucei*)
	LmjF.24.0150	220	-	-	2.7E-15 (*Trypanosoma brucei*)
	LmjF.28.2250	221	-	-	2.9E-26 (*Trypanosoma brucei*)
	Lmj35.3700[e]	225	-	-	1.1E-79 (*Trypanosoma brucei*)
Naegleria gruberi	45694[b]	256	0.0034	0.015	
Trypanosoma brucei	Tb927.11.11520	218	-	-	

	Tb927.9.11580[e]	243	-	-	
	Tb927.9.11600[e]	241	-	-	
Trypanosoma cruzi	TcCLB.509877.50	218	-	-	4.3E-5 (*Trypanosoma brucei*)
	TcCLB.511017.50	218	-	-	1.2E-83 (*Trypanosoma brucei*)
	TcCLB.508461.570[e]	244	-	-	2.9E-102 (*Trypanosoma brucei*)

[a]Query
[b]Pex11 homologues identified with both Pex11 and Pex11B queries
[c]Pex11 homologues identified with both Pex11 and Pex11C queries
[d]Pex11 homologues identified with both Pex11 and Pex11/25 queries
[e]GIM5 homologue

Appendix 2: Pex11Bp homologues identified in comparative genomic survey

Genome	Accession	Length	R. pHMMer	R. pHMMer next hit	R. pHMMer Genome
Fungi					
Acremonium alcalophilum	1082276	270	2.4E-12	Pex11 (1.7E-5)	*A. fumigatus*
			0.00000011	Pex11 (2.3E-6)	*A. nidulans*
			1.20E-40	-	*G. zeae*
			4.10E-34	Pex11 (6E-6)	*M. oryzae*
			1.00E-17	Pex11 (2.1E-8)	*N. crassa*
			6.80E-07	-	*P. chrysogenum*
Aspergillus fumigatus[a]	Afu6g10070	250	0	Pex11 (5.1E-9)	*A. fumigatus*
			1.30E-57	Pex11 (2.1E-8)	*A. nidulans*
			5.00E-11	Pex11 (2E-9)	*G. zeae*
			5.80E-16	Pex11 (6.9E-10)	*N. crassa*
			2.60E-59	Pex11 (8.4E-8)	*P. chrysogenum*
Aspergillus nidulans[a]	EAA62588.1	238	8.90E-58	Pex11 (4.2E-5)	*A. fumigatus*
			0	Pex11 (5.3E-5)	*A. nidulans*
			1.10E-07	Pex11 (0.0063)	*G. zeae*
			1.00E-16	Pex11 (6.7E-6)	*M. oryzae*
			4.70E-16	Pex11 (1.3E-5)	*N. crassa*
			3.40E-54	Pex11 (0.0023)	*P. chrysogenum*
Aureobasidium pullulans	297952	272	1E-27	Pex11 (1.1E-6)	*A. fumigatus*
			4.70E-25	Pex11 (1.2E-7)	*A. nidulans*
			1.20E-11	Pex11 (1.7E-6)	*G. zeae*
			8.30E-25	Pex11 (2.7E-5)	*P. chrysogenum*
Botryotinia fuckeliana	XP_001549681.1	239	2.7E-15	Pex11 (0.0017)	*A. fumigatus*
			6.80E-11	Pex11 (0.0019)	*A. nidulans*
			1.30E-21	-	*G. zeae*
			5.20E-16	Pex11 (5.7E-6)	*M. oryzae*
			9.50E-11	Pex11 (0.00085)	*N. crassa*

			1.90E-12	Pex11 (0.00065)	*P. chrysogenum*
Coccidioides immitis	CIMG_04752	264	5.7E-41	Pex11 (0.0028)	*A. fumigatus*
			1.00E-31	Pex11 (1.8E-5)	*A. nidulans*
			2.00E-18	Pex11 (2.4E-9)	*G. zeae*
			1.30E-15	Pex11 (5.1E-12)	*M. oryzae*
			3.10E-12	Pex11 (3.4E-8)	*N. crassa*
			1.30E-35	Pex11 (0.00039)	*P. chrysogenum*
Cryphonectria parasitica	323933	347	2.8E-15	Pex11 (0.0048)	*A. fumigatus*
			1.80E-12	Pex11 (0.00037)	*A. nidulans*
			4.00E-35	-	*G. zeae*
			1.10E-38	Pex11 (0.00019)	*M. oryzae*
			2.80E-45	Pex11 (7.1E-5)	*N. crassa*
			5.20E-08	-	*P. chrysogenum*
Gibberella zeae[a]	XP_389427.1	292	1.7E-10	Pex11 (2.9E-10)	*A. fumigatus*
			0	-	*G. zeae*
			6.00E-34	Pex11 (2.3E-8)	*M. oryzae*
			7.50E-28	Pex11 (2.9E-8)	*N. crassa*
			3.20E-13	Pex11 (3.6E-8)	*P. chrysogenum*
Magnaporthe oryzae[a]	MGG_00648	281	1.2E-14	Pex11 (9.1E-19)	*A. fumigatus*
			3.30E-16	Pex11 (1.5E-8)	*A. nidulans*
			2.90E-34	Pex11 (0.00079)	*G. zeae*
			0	Pex11 (2.9E-7)	*M. oryzae*
			7.80E-40	Pex11 (9.5E-10)	*N. crassa*
			3.90E-15	Pex11 (1.8E-9)	*P. chrysogenum*
Myceliophthora thermophila	2306713	332	5.1E-10	Pex11 (0.0054)	*A. fumigatus*
			2.40E-08	Pex11 (2.2E-5)	*A. nidulans*
			3.10E-30	-	*G. zeae*
			3.50E-43	Pex11 (2.3E-6)	*M. oryzae*
			7.40E-40	Pex11 (3.1E-7)	*N. crassa*
			1.40E-07	Pex11 (0.00074)	*P. chrysogenum*

Mycosphaerella graminicola	62362	248	4.3E-39	Pex11 (0.001)	*A. fumigatus*
			3.40E-37	Pex11 (0.012)	*A. nidulans*
			2.90E-08	Pex11 (7.8E-5)	*G. zeae*
			8.20E-10	Pex11 (2.4E-7)	*M. oryzae*
			0.00038	Pex11 (0.004)	*N. crassa*
			1.20E-27	Pex11 (0.025)	*P. chrysogenum*
Neurospora crassa[a]	NCU04045	368	1E-13	Pex11 (0.0011)	*A. fumigatus*
			1E-13	Pex11 (0.017)	*A. nidulans*
			1.90E-26	-	*G. zeae*
			6.20E-39	Pex11 (0.0013)	*M. oryzae*
			0	Pex11 (0.0073)	*N. crassa*
			4.70E-11	-	*P. chrysogenum*
Penicillium chrysogenum[a]	XP_002564016.1	224	4.1E-59	Pex11 (2.8E-10)	*A. fumigatus*
			3E-54	Pex11 (5.2E-10)	*A. nidulans*
			5.30E-15	Pex11 (8.7E-11)	*G. zeae*
			8.40E-16	Pex11 (9E-16)	*M. oryzae*
			7.20E-13	Pex11 (6.9E-12)	*N. crassa*
			0	Pex11 (1.6E-11)	*P. chrysogenum*
Trichoderma reesei	120524	312	2.8E-13	Pex11 (0.00012)	*A. fumigatus*
			0.000000078	Pex11 (8.5E-5)	*A. nidulans*
			4.10E-45	-	*G. zeae*
			7.10E-27	Pex11 (0.0087)	*M. oryzae*
			3.60E-15	-	*N. crasssa*
			1.50E-09	Pex11 (0.00011)	*P. chrysogenum*
Archaeplastida					
Chlamydomonas reinhardtii	Cre12.g540500[b]	238	5.60E-06	Pex11 (8.8E-5)	*P. chrysogenum*
Micromonas sp.	61774[b]	241	0.011	-	*P. chrysogenum*
Volvox carteri	Vocar20006405m[b]	312	2.00E-05	Pex11 (2.9E-5)	*P. chrysogenum*
SAR					
Bigelowiella natans	89634	242	0.0017	-	*A. fumigatus*

Tetrahymena thermophila	TTHERM_00355720[b]	384	1.40E-150	Pex11 (1.6E-11)	*P. chrysogenum*
Excavata					
Naegleria gruberi	45694[b]	256	6.40E-05	Pex11 (0.0071)	*P. chrysogenum*

[a]Query
[b]Pex11 homologues identified with both Pex11 and Pex11B queries

Appendix 3: Pex11Cp homologues identified in comparative genomic survey

Genome	Accession	Length	R. pHMMer (*A. fumigatus*)	R. pHMMer next hit (*Aspergillus fumigatus*)	R. pHMMer Genome
Fungi					
Acremonium alcalophilum	1065377	347	1.40E-15	-	*A. fumigatus*
			4.30E-22	-	*A. nidulans*
			0.027	-	*C. albicans*
			1.10E-20	-	*C. neoformans*
			0.0075	-	*D. hansenii*
			9.20E-49	-	*G. zeae*
			4.70E-11	-	*H. polymorpha*
			1.90E-61	-	*M. oryzae*
			6.10E-53	-	*N. crassa*
			1.90E-13	-	*P. chrysogenum*
			0.0012	-	*S. pombe*
			3.30E-13	-	*U. maydis*
			6.40E-20	-	*Y. lipolytica*
Agaricus bisporus	191439	309	1.90E-12	-	*A. fumigatus*
			2.30E-13	-	*A. nidulans*
			9.60E-26	-	*C. neoformans*
			0.002	Pex11 (0.013)	*H. polymorpha*
			0.031	-	*M. oryzae*
			7.70E-07	-	*N. crassa*
			5.20E-15	-	*P. chrysogenum*
			0.0034	-	*S. pombe*
			4.40E-67	Pex11 (0.032)	*U. maydis*
			7.40E-19	-	*Y. lipolytica*
Allomyces macrogynus	AMAG_11374	239	8.50E-05		*A. fumigatus*
			1.60E-06	-	*A. nidulans*
			0.045	-	*C. albicans*

			1.20E-08	Pex11 (0.0078)	*C. neoformans*
			6.70E-09	-	*H. polymorpha*
			8.50E-16	Pex11 (0.0019)	*U. maydis*
			3.30E-05	Pex11 (0.00031)	*Y. lipolytica*
Ascoidea rubescens	75790	338	3.40E-09	-	*C. albicans*
			9.40E-11	-	*D. hansenii*
			2.70E-11	-	*H. polymorpha*
			4.60E-07	-	*S. pombe*
			0.002	-	*U. maydis*
			1.60E-70	Pex11 (0.0078)	*Y. lipolytica*
Aspergillus fumigatus[a]	Afu2g17100	317	0	-	*A. fumigatus*
			6.30E-104	-	*A. nidulans*
			0.0012	-	*C. albicans*
			4.10E-07	-	*C. neoformans*
			0.00037	-	*G. zeae*
			0.022	-	*H. polymorpha*
			2.40E-08	Pex11 (0.00043)	*M. oryzae*
			1.10E-15	-	*N. crassa*
			4.20E-116	-	*P. chrysogenum*
			6.00E-06	-	*S. pombe*
			2.30E-17	-	*U. maydis*
			1.30E-08	-	*Y. lipolytica*
Aspergillus nidulans[a]	EAA61162.1	255	1.70E-70	-	*A. fumigatus*
			0	-	*A. nidulans*
			5.00E-06	-	*C. neoformans*
			0.047	-	*D. hansenii*
			0.0063	-	*G. zeae*
			4.40E-07	-	*M. oryzae*
			3.60E-13	-	*N. crassa*
			6.40E-62	-	*P. chrysogenum*

			7.60E-07	-	*U. maydis*
			6.70E-08	-	*Y. lipolytica*
Aureobasidium pullulans	4763	283	1.50E-31	-	*A. fumigatus*
			4.40E-24	-	*A. nidulans*
			7.40E-09	-	*C. neoformans*
			5.60E-05	-	*D. hansenii*
			0.0001	-	*G. zeae*
			1.20E-09	-	*H. polymorpha*
			5.60E-17	-	*M. oryzae*
			2.40E-18	-	*N. crassa*
			5.60E-21	-	*P. chrysogenum*
			3.60E-06	-	*S. pombe*
			1.70E-18	-	*U. maydis*
			5.20E-14	Pex11 (0.02)	*Y. lipolytica*
Batrachochytrium dendrobatidis	BDEG_03726	252	3.50E-13	Pex11 (0.00039)	*A. fumigatus*
			1.80E-13	Pex11 (0.0055)	*A. nidulans*
			1.10E-06	Pex11 (0.00012)	*C. albicans*
			1.40E-30	Pex11 (0.00066)	*C. neoformans*
			1.30E-09	Pex11 (0.00078)	*G. zeae*
			1.40E-08	Pex11 (0.025)	*H. polymorpha*
			6.50E-09	Pex11 (0.00035)	*M. oryzae*
			1.40E-08	Pex11 (0.00011)	*N. crassa*
			7.00E-19	Pex11 (0.0028)	*P. chrysogenum*
			5.40E-14	Pex11 (0.021)	*S. pombe*
			3.00E-38	Pex11 (7.9E-8)	*U. maydis*
			3.30E-21	Pex11 (0.0039)	*Y. lipolytica*
Botryotinia fuckeliana	XP_001551936.1	306	1.00E-32	-	*A. fumigatus*
			4.70E-33	-	*A. nidulans*
			1.40E-08	-	*C. albicans*
			2.00E-17	-	*C. neoformans*

			6.40E-11	Pex11 (0.0086)	*D. hansenii*
			2.70E-29	-	*G. zeae*
			4.50E-18	-	*H. polymorpha*
			3.30E-35	-	*M. oryzae*
			4.80E-40	-	*N. crassa*
			4.20E-27	-	*P. chrysogenum*
			1.50E-08	-	*S. pombe*
			1.40E-28	-	*U. maydis*
			2.40E-29	-	*Y. lipolytica*
Candida albicans[a]	EAK96575.1	330	3.20E-06	-	*A. fumigatus*
			1.10E-05	-	*A. nidulans*
			0	-	*C. albicans*
			0.011	-	*C. neoformans*
			1.60E-124	-	*D. hansenii*
			1.30E-22	-	*H. polymorpha*
			5.90E-10	-	*M. oryzae*
			1.00E-15	-	*N. crassa*
			0.001	-	*P. chrysogenum*
			5.30E-05	-	*S. pombe*
			8.70E-07	-	*U. maydis*
			3.70E-11	-	*Y. lipolytica*
Catenaria anguillulae	125157	232	7.40E-08	-	*A. fumigatus*
			4.70E-05	-	*A. nidulans*
			7.20E-13	Pex11 (3.8E-5)	*C. neoformans*
			0.0005	-	*H. polymorpha*
			0.0048	-	*P. chrysogenum*
			0.033	-	*S. pombe*
			4.40E-12	Pex11 (0.00012)	*U. maydis*
			0.0007	-	*Y. lipolytica*
Cladonia grayi	118861	270	9.80E-24	-	*A. fumigatus*

			1.40E-19	-	*A. nidulans*
			3.10E-05	-	*C. albicans*
			5.40E-18	-	*C. neoformans*
			6.80E-06	-	*D. hansenii*
			4.80E-09	-	*G. zeae*
			1.50E-07	-	*H. polymorpha*
			2.10E-23	-	*M. oryzae*
			5.80E-27	-	*N. crassa*
			3.50E-18	-	*P. chrysogenum*
			3.40E-06	-	*S. pombe*
			7.00E-20	Pex11 (0.038)	*U. maydis*
			2.20E-18	-	*Y. lipolytica*
Coccidioides immitis	CIMG_09403	314	1.00E-81	-	*A. fumigatus*
			6.40E-63	-	*A. nidulans*
			0.00037	-	*C. albicans*
			7.10E-09	-	*C. neoformans*
			0.001	-	*D. hansenii*
			2.00E-07	-	*G. zeae*
			1.10E-08	-	*H. polymoprha*
			8.80E-12	-	*M. oryzae*
			4.70E-15	-	*N. crassa*
			1.40E-75	-	*P. chrysogenum*
			8.60E-11	-	*S. pombe*
			1.70E-15	-	*U. maydis*
			9.70E-16	-	*Y. lipolytica*
Cochliobolus heterostrophus	1227209	340	1.10E-21	-	*A. fumigatus*
			7.30E-19	-	*A. nidulans*
			1.80E-07	-	*C. albicans*
			1.50E-09	-	*C. neoformans*
			2.80E-08	-	*D. hansenii*

			3.10E-14	-	*G. zeae*
			7.40E-16	-	*H. polymorpha*
			7.60E-25	-	*M. oryzae*
			7.60E-34	-	*N. crassa*
			7.10E-23	-	*P. chrysogenum*
			1.10E-06	-	*S. pombe*
			6.20E-16	-	*U. maydis*
			5.90E-23	-	*Y. lipolytica*
	1133585	328	8.90E-29	Pex11 (0.00066)	*A. fumigatus*
			2.80E-21	Pex11 (0.03)	*A. nidulans*
			0.0002	-	*M. oryzae*
			0.00068	-	*N. crassa*
			8.70E-34	-	*P. chrysogenum*
			5.60E-08	-	*U. maydis*
			8.50E-06	-	*Y. lipolytica*
Coemansia reversa	37739	308	2.20E-16	Pex11 (8.8E-7)	*A. fumigatus*
			5.30E-12	Pex11 (5.3E-5)	*A. nidulans*
			2.10E-06	Pex11 (1.7E-5)	*C. albicans*
			1.00E-21	Pex11 (0.0019)	*C. neoformans*
			7.60E-12	Pex11 (0.0024)	*H. polymorpha*
			2.70E-07	Pex11 (3.5E-7)	*N. crassa*
			2.20E-12	Pex11 (6.8E-7)	*P. chrysogenum*
			2.10E-06	-	*S. pombe*
			7.80E-38	Pex11 (6.3E-5)	*U. maydis*
			6.60E-17	Pex11 (2.4E-8)	*Y. lipolytica*
Conidiobolus coronatus	78627	259	0.0027	-	*A. fumigatus*
			2.80E-13	-	*C. neoformans*
			6.50E-05	-	*M. oryzae*
			0.023	-	*N. crassa*
			5.10E-06	Pex11 (0.0014)	*P. chrysogenum*

			7.90E-13	Pex11 (0.0082)	*U. maydis*
Coprinopsis cinerea	CC1G_11632	304	3.30E-12	-	*A. fumigatus*
			4.00E-16	-	*A. nidulans*
			7.90E-08	-	*C. neoformans*
			0.00013	-	*G. zeae*
			1.20E-06	Pex11 (0.043)	*H. polymorpha*
			1.10E-07	-	*M. oryzae*
			1.60E-09	-	*N. crassa*
			2.90E-17	-	*P. chrysogenum*
			0.00011	-	*S. pombe*
			2.10E-75	-	*U. maydis*
			3.20E-19	-	*Y. lipolytica*
Cryphonectria parasitica	283973	327	6.30E-06	-	*A. fumigatus*
			2.60E-11	-	*A. nidulans*
			6.60E-07	-	*C. albicans*
			1.90E-08	-	*C. neoformans*
			0.0007	-	*D. hansenii*
			8.60E-33	-	*G. zeae*
			2.70E-11	-	*H. polymorpha*
			6.90E-07	-	*M. oryzae*
			2.50E-09	-	*N. crassa*
			6.30E-07	-	*P. chrysogenum*
			4.10E-07	-	*S. pombe*
			4.20E-13	-	*U. maydis*
			1.40E-11	-	*Y. lipolytica*
Cryptococcus neoformans[a]	CNAG_03394	316	2.10E-06	-	*A. fumigatus*
			2.80E-09	-	*A. nidulans*
			0	-	*C. neoformans*
			0.0027	-	*D. hansenii*
			1.10E-05	-	*G. zeae*

			1.70E-06	-	*H. polymorpha*
			2.20E-11	-	*M. oryzae*
			1.20E-14	-	*N. crassa*
			1.10E-05	-	*P. chrysogenum*
			1.70E-07	-	*S. pombe*
			2.00E-24	-	*U. maydis*
			3.30E-22	-	*Y. lipolytica*
Dacryopinax sp.	20433	285	7.50E-12	-	*A. fumigatus*
			1.50E-09	-	*A. nidulans*
			2.30E-15	-	*C. neoformans*
			0.00018	-	*D. hansenii*
			5.90E-09	-	*G. zeae*
			0.00047	-	*H. polymorpha*
			0.00016	-	*M. oryzae*
			1.10E-08	-	*P. chrysogenum*
			2.40E-09	-	*S. pombe*
			1.20E-15	-	*U. maydis*
			3.00E-13	-	*Y. lipolytica*
Debaryomyces hansenii[a]	XP_457126.1	313	0.0033	-	*A. fumigatus*
			0.0002	-	*A. nidulans*
			4.30E-121	-	*C. albicans*
			0.00021	-	*C. neoformans*
			0	-	*D. hansenii*
			5.40E-22	-	*H. polymorpha*
			1.50E-08	-	*M. oryzae*
			4.30E-15	-	*N. crassa*
			0.0047	-	*P. chrysogenum*
			0.0042	-	*S. pombe*
			4.00E-07	-	*U. maydis*
			1.20E-09	-	*Y. lipolytica*

Dekkera bruxellensis	25980	318	1.60E-13	-	*C. albicans*
			1.90E-15	-	*D. hansenii*
			0.027	-	*G. zeae*
			3.90E-46	-	*H. polymorpha*
			0.0018	-	*N. crassa*
			0.00017	-	*S. pombe*
			0.00014	-	*U. maydis*
			1.50E-09	-	*Y. ipolytica*
Gibberella zeae[a]	XP_382722.1	331	0.04	-	*A. fumigatus*
			1.20E-05	-	*A. nidulans*
			0.00011	-	*C. neoformans*
			0	-	*G. zeae*
			0.0012	-	*H. polymorpha*
			2.10E-37	-	*M. oryzae*
			1.00E-33	-	*N. crassa*
			0.027	-	*P. chrysogenum*
			4.80E-05	-	*U. maydis*
			7.00E-07	-	*Y. lipolytica*
	XP_383527.1	390	0.0018	-	*A. fumigatus*
			0.004	-	*A. nidulans*
			2.50E-06	-	*C. neoformans*
			0	-	*G. zeae*
			8.20E-07	-	*H. polymorpha*
			2.00E-21	-	*M. oryzae*
			1.90E-13	-	*N. crassa*
			5.70E-05	-	*U. maydis*
			1.60E-09	-	*Y. lipolytica*
Gonapodya prolifera	50392	256	9.90E-16	Pex11 (3.7E-5)	*A. fumigatus*
			2.20E-14	Pex11 (0.0025)	*A. nidulans*
			0.0077	-	*C. albicans*

			3.60E-27	Pex11 (9.1E-8)	*C. neoformans*
			9.60E-12	Pex11 (0.029)	*G. zeae*
			7.90E-09	0.0006	*H. polymoprha*
			5.50E-08	Pex11 (3.5E-5)	*M. oryzae*
			4.90E-12	Pex11 (1.9E-5)	*N. crassa*
			4.40E-18	Pex11 (1.1E-5)	*P. chrysogenum*
			5.00E-07	-	*S. pombe*
			4.20E-37	Pex11 (1.4E-6)	*U. maydis*
			2.20E-11	Pex11 (1.2E-8)	*Y. lipolytica*
	63211	251	1.40E-07	Pex11 (3.8E-7)	*A. fumigatus*
			7.80E-05	Pex11 (0.00013)	*A. nidulans*
			6.40E-12	-	*C. neoformans*
			9.60E-06	Pex11 (0.00012)	*M. oryzae*
			0.0015	Pex11 (0.0056)	*N. crassa*
			1.80E-05	Pex11 (0.0004)	*P. chrysogenum*
			3.50E-17	Pex11 (1.9E-5)	*U. maydis*
Hansenula polymorpha[a]	83008	309	0.00021	-	*A. fumigatus*
			0.00025	-	*A. nidulans*
			2.10E-23	-	*C. albicans*
			1.50E-06	-	*C. neoformans*
			7.30E-22	-	*D. hansenii*
			1.50E-06	-	*G. zeae*
			0	-	*H. polymorpha*
			3.30E-12	-	*M. oryzae*
			1.10E-09	-	*N. crassa*
			0.00013	-	*P. chrysogenum*
			3.10E-05	-	*S. pombe*
			5.60E-15	-	*U. maydis*
			3.60E-25	-	*Y. lipolytica*
Hysterium pulicare	7801	314	6.50E-26	Pex11 (7.6E-7)	*A. fumigatus*

			5.30E-21	Pex11 (7.1E-5)	*A. nidulans*
			3.20E-05	-	*C. neoformans*
			4.80E-05	Pex11 (0.032)	*M. oryzae*
			0.0026	Pex11 (0.0039)	*N. crassa*
			1.10E-30	Pex11 (0.0025)	*P. chrysogenum*
			1.00E-07	-	*U. maydis*
			1.80E-06	Pex11 (0.002)	*Y. lipolytica*
	11214	335	8.10E-24	-	*A. fumigatus*
			6.60E-25	-	*A. nidulans*
			0.00039	-	*C. albicans*
			5.30E-08	-	*C. neoformans*
			1.30E-08	-	*D. hansenii*
			1.20E-10	-	*G. zeae*
			1.90E-13	-	*H. polymorpha*
			8.40E-26	-	*N. crassa*
			2.30E-22	-	*P. chrysogenum*
			0.00055	-	*S. pombe*
			8.50E-19	-	*U. maydis*
			2.00E-23	-	*Y. lipolytica*
Lipomyces starkeyi	43134	267	9.90E-29	-	*A. fumigatus*
			3.80E-24	-	*A. nidulans*
			9.80E-11	-	*C. albicans*
			6.90E-17	-	*C. neoformans*
			1.70E-14	-	*D. hansenii*
			8.00E-11	-	*G. zeae*
			1.30E-28	-	*H. polymorpha*
			4.00E-18	-	*M. oryzae*
			4.30E-18	-	*N. crassa*
			4.40E-25	-	*P. chrysogenum*
			5.40E-11	-	*S. pombe*

			1.40E-27	-	*U. maydis*
			1.30E-37	-	*Y. lipolytica*
Lodderomyces elongisporus	LELG_05664	304	2.70E-06	-	*A. fumigatus*
			4.80E-06	-	*A. nidulans*
			2.10E-139	-	*C. albicans*
			0.00085	-	*C. neoformans*
			1.30E-124	-	*D. hansenii*
			3.00E-20	-	*H. polymorpha*
			7.40E-11	-	*M. oryzae*
			1.10E-16	-	*N. crassa*
			2.90E-05	-	*P. chrysogenum*
			0.0039	-	*S. pombe*
			4.90E-06	-	*U. maydis*
			3.90E-09	-	*Y. lipolytica*
Magnaporthe oryzae[a]	MGG_05271	320	8.20E-07	-	*A. fumigatus*
			2.80E-11	-	*A. nidulans*
			7.70E-09	-	*C. albicans*
			2.10E-11	-	*C. neoformans*
			7.70E-08	Pex11 (0.034)	*D. hansenii*
			4.10E-37	-	*G. zeae*
			1.70E-11	-	*H. polymorpha*
			0	-	*M. oryzae*
			5.60E-72	-	*N. crassa*
			1.30E-08	-	*P. chrysogenum*
			1.80E-09	-	*S. pombe*
			2.30E-13	-	*U. maydis*
			2.80E-13	-	*Y. lipolytica*
Metschnikowia bicuspidata	37511	306	2.90E-05	-	*A. fumigatus*
			0.00015	-	*A. nidulans*
			1.20E-107	-	*C. albicans*

			0.00018	-	*C. neoformans*
			9.40E-122	-	*D. hansenii*
			2.10E-21	-	*H. polymorpha*
			2.10E-11	-	*M. oryzae*
			8.30E-14	-	*N. crassa*
			0.0094	-	*P. chrysogenum*
			0.0025	-	*S. pombe*
			7.00E-09	-	*U. mayids*
			1.40E-11	-	*Y. lipolytica*
Mixia osmundae	44932	258	3.50E-14	Pex11 (2.1E-5)	*A. fumigatus*
			6.40E-09	Pex11 (0.0027)	*A. nidulans*
			2.50E-05	-	*C. albicans*
			4.80E-28	Pex11 (1.6E-5)	*C. neoformans*
			2.50E-05	Pex11 (0.0014)	*D. hansenii*
			0.0001	Pex11 (0.0012)	*G. zeae*
			1.20E-10	Pex11 (1.8E-5)	*H. polymorpha*
			2.70E-12	Pex11 (1.8E-6)	*M. oryzae*
			9.30E-14	Pex11 (0.011)	*N. crassa*
			6.00E-22	Pex11 (0.00017)	*P. chrysogenum*
			1.40E-08	-	*S. pombe*
			1.10E-64	Pex11 (5.7E-10)	*U. maydis*
			6.90E-23	Pex11 (6.5E-9)	*Y. lipolytica*
Mucor circinelloides	76316	274	5.40E-13	Pex11 (1.4E-7)	*A. fumigatus*
			1.60E-13	Pex11 (2.3E-6)	*A. nidulans*
			0.00049	-	*C. albicans*
			3.70E-27	Pex11 (0.025)	*C. neoformans*
			6.80E-09	Pex11 (0.033)	*H. polymorpha*
			2.40E-12	Pex11 (9.8E-9)	*M. oryzae*
			3.70E-13	Pex11 (2.9E-7)	*N. crassa*
			5.60E-16	Pex11 (8E-6)	*P. chrysogenum*

			1.90E-10	Pex11 (0.0016)	*S. pombe*
			2.70E-43	Pex11 (8.6E-6)	*U. maydis*
			6.90E-22	Pex11 (9.5E-6)	*Y. lipolytica*
Myceliophthora thermophila	2304628	341	8.50E-13	-	*A. fumigatus*
			1.20E-16	-	*A. nidulans*
			9.30E-08	-	*C. albicans*
			8.10E-11	-	*C. neoformans*
			5.70E-09	-	*D. hansenii*
			3.40E-41	-	*G. zeae*
			1.30E-09	-	*H. polymorpha*
			4.50E-87	-	*M. oryzae*
			2.80E-95	-	*N. crassa*
			2.20E-13	-	*P. chrysogenum*
			2.30E-05	-	*S. pombe*
			6.70E-15	-	*U. maydis*
			1.10E-12	-	*Y. lipolytica*
Mycosphaerella graminicola	103337	312	3.10E-26	-	*A. fumigatus*
			2.70E-19	-	*A. nidulans*
			2.30E-05	-	*M. oryzae*
			9.10E-08	-	*N. crassa*
			3.50E-23	-	*P. chrysogenum*
			7.60E-05	-	*U. maydis*
			5.80E-05	-	*Y. lipolytica*
	91228	287	1.20E-10	-	*A. fumigatus*
			2.20E-15	-	*A. nidulans*
			9.80E-05	-	*G. zeae*
			8.50E-11	-	*M. oryzae*
			1.50E-06	-	*N. crassa*
			4.10E-23	-	*P. chrysogenum*
			9.50E-05	-	*S. pombe*

			4.10E-06	-	*U. maydis*
			0.00074	-	*Y. lipolytica*
Neurospora crassa[a]	NCU06275	330	8.20E-14	-	*A. fumigatus*
			5.90E-17	-	*A. nidulans*
			1.60E-15	-	*C. albicans*
			9.00E-15	-	*C. neoformans*
			5.10E-15	-	*D. hansenii*
			1.90E-35	-	*G. zeae*
			1.50E-10	-	*H. polymorpha*
			2.30E-12	-	*M. oryzae*
			0	-	*N. crassa*
			1.10E-11	-	*P. chrysogenum*
			3.40E-11	-	*S. pombe*
			7.30E-15	-	*U. maydis*
			2.10E-15	-	*Y. lipolytica*
Pachysolen tannophilus	33324	317	0.00022	-	*A. fumigatus*
			3.30E-31	-	*C. albicans*
			5.10E-05	-	*C. neoformans*
			1.10E-34	-	*D. hansenii*
			0.00057	-	*G. zeae*
			3.80E-49	-	*H. polymorpha*
			8.80E-09	-	*M. oryzae*
			9.00E-10	-	*N. crassa*
			3.50E-05	-	*P. chrysogenum*
			0.0024	-	*S. pombe*
			1.50E-07	-	*U. maydis*
			6.90E-21	-	*Y. lipolytica*
Penicillium chrysogenum[a]	XP_002568919.1	303	6.70E-116	-	*A. fumigatus*
			1.50E-86	-	*A. nidulans*
			0.026	-	*C. albicans*

			1.30E-06	-	*C. neoformans*
			0.0085	-	*G. zeae*
			8.10E-15	-	*H. polymorpha*
			3.30E-10	Pex11 (0.029)	*M. oryzae*
			1.30E-11	-	*N. crassa*
			0.00E+00	-	*P. chrysogenum*
			1.10E-06	-	*S. pombe*
			1.40E-11	-	*U. maydis*
			2.70E-11	Pex11 (0.0075)	*Y. lipolytica*
Phaenerochaete chrysosporium	65313	326	0.0024	-	*C. neoformans*
			3.90E-16	-	*U. maydis*
			0.00079	-	*Y. lipolytica*
Phaeosphaeria nodorum	SNOG_10242	338	5.00E-22	-	*A. fumigatus*
			2.90E-17	-	*A. nidulans*
			6.50E-06	-	*C. albicans*
			2.90E-11	-	*C. neoformans*
			4.10E-08	-	*D. hansenii*
			1.40E-15	-	*G. zeae*
			8.10E-15	-	*H. polymorpha*
			5.20E-30	-	*M. oryzae*
			5.80E-37	-	*N. crassa*
			3.30E-20	-	*P. chrysogenum*
			1.20E-07	-	*S. pombe*
			1.50E-08	-	*U. maydis*
			2.20E-24	-	*Y. lipolytica*
	SNOG_12948	396	9.80E-23	0.0022	*A. fumigatus*
			2.40E-19	Pex11 (0.019)	*A. nidulans*
			0.0075	-	*M. oryzae*
			0.016	-	*N. crassa*
			2.10E-25	Pex11B (0.0057)	*P. chrysogenum*

			6.90E-07	-	*U. maydis*
			0.00034	-	*Y. lipolytica*
Phycomyces blakesleeanus	126109	265	2.20E-12	Pex11 (5E-7)	*A. fumigatus*
			1.10E-12	Pex11 (3E-6)	*A. nidulans*
			0.0011	-	*C. albicans*
			1.60E-24	-	*C. neoformans*
			1.90E-08	Pex11 (2.6E-5)	*G. zeae*
			7.30E-10	Pex11 (0.0011)	*H. polymorpha*
			1.30E-09	Pex11 (6.6E-8)	*M. oryzae*
			3.30E-11	Pex11 (2.2E-7)	*N. crassa*
			8.40E-17	Pex11 (4.3E-6)	*P. chrysogenum*
			8.40E-09	Pex11 (0.00017)	*S. pombe*
			2.80E-40	Pex11 (2.1E-6)	*U. maydis*
			4.00E-18	Pex11 (4.8E-7)	*Y. lipolytica*
Pichia membranifaciens	71684	305	0.00043	-	*A. fumigatus*
			1.00E-19	-	*C. albicans*
			2.20E-25	-	*D. hansenii*
			0.017	-	*G. zeae*
			2.00E-35	-	*H. polymorpha*
			3.00E-09	-	*M. orzyae*
			7.60E-09	-	*N. crassa*
			0.00011	-	*P. chrysogenum*
			1.50E-06	-	*S. pombe*
			4.10E-06	-	*U. maydis*
			6.50E-11	-	*Y. lipolytica*
Pichia pastoris	XP_002490664.1	256	2.70E-06	-	*A. fumigatus*
			0.00061	-	*A. nidulans*
			5.60E-07	-	*C. albicans*
			0.028	-	*C. neoformans*
			3.50E-05	-	*D. hansenii*

			0.02	-	*G. zeae*
			5.10E-19	Pex11 (0.045)	*H. polymoprha*
			0.0034	-	*M. oryzae*
			0.0015	-	*N. crassa*
			0.00015	-	*P. chrysogenum*
			8.70E-13	-	*U. maydis*
			9.40E-10	-	*Y. lipolytica*
Pichia stipitis	86597	315	5.60E-06	-	*A. fumigatus*
			4.70E-139	-	*C. albicans*
			0.0059	-	*C. neoformans*
			2.50E-137	-	*D. hansenii*
			1.10E-20	Pex11 (0.023)	*H. polymorpha*
			1.70E-10	-	*M. oryzae*
			8.70E-15	-	*N. crassa*
			0.0042	-	*P. chrysogenum*
			0.011	-	*S. pombe*
			6.60E-08	-	*U. maydis*
			3.20E-10	-	*Y. lipolytica*
Piromyces sp.	66086	240	5.50E-12	Pex11 (4.8E-6)	*A. fumigatus*
			1.20E-06	Pex11 (4.2E-5)	*A. nidulans*
			1.70E-16	Pex11 (0.00012)	*D. hansenii*
			3.90E-05	Pex11 (0.0051)	*H. polymorpha*
			4.40E-09	Pex11 (2.4E-5)	*P. chrysogenum*
			2.00E-31	Pex11 (0.00028)	*U. maydis*
			1.40E-10	-	*Y. lipolytica*
Puccinia graminis	PGTG_14615	246	2.50E-17	-	*A. fumigatus*
			5.20E-21	-	*A. nidulans*
			0.01	-	*C. albicans*
			7.60E-30	-	*C. neoformans*
			1.70E-02	-	*D. hansenii*

			0.0048	-	*G. zeae*
			2.40E-09	-	*H. polymoprha*
			2.50E-12	-	*M. oryzae*
			1.30E-16	-	*N. crassa*
			1.40E-20	Pex11B (0.0023)	*P. chrysogenum*
			1.40E-07	-	*S. pombe*
			4.70E-43	Pex11 (0.0012)	*U. maydis*
			5.40E-20	-	*Y. lipolytica*
Rhizophagus irregularis	349385	253	4.30E-15	Pex11 (0.00025)	*A. fumigatus*
			2.30E-25	Pex11 (8.6E-6)	*C. neoformans*
			1.70E-06	Pex11 (0.0083)	*G. zeae*
			1.50E-14	Pex11 (3.6E-5)	*H. polymorpha*
			2.80E-10	Pex11 (1.8E-7)	*M. oryzae*
			1.00E-07	Pex11 (2.2E-7)	*N. crassa*
			1.30E-15	Pex11 (7.5E-5)	*P. chrysogenum*
			1.20E-09	Pex11 (0.0066)	*S. pombe*
			2.50E-45	Pex11 (1.8E-6)	*U. maydis*
			3.00E-23	Pex11 (3E-7)	*Y. lipolytica*
Rhizopus oryzae	RO3G_04077	269	9.90E-12	Pex11 (1.4E-7)	*A. fumigatus*
			4.30E-14	Pex11 (4E-6)	*A. nidulans*
			1.70E-06	-	*C. albicans*
			1.20E-25	Pex11 (0.01)	*C. neoformans*
			6.70E-08	Pex11 (7.3E-5)	*G. zeae*
			6.20E-11	Pex11 (0.00015)	*H. polymorpha*
			1.40E-12	Pex11 (4.5E-8)	*M. oryzae*
			4.10E-14	Pex11 (1.5E-7)	*N. crassa*
			2.30E-14	Pex11 (1.6E-5)	*P. chrysogenum*
			1.70E-11	Pex11 (0.008)	*S. pombe*
			4.20E-42	Pex11 (3.8E-5)	*U. maydis*
			7.10E-21	Pex11 (6.7E-7)	*Y. lipolytica*

Rhodotorula graminis	44831	267	4.50E-09	-	*A. fumigatus*
			1.10E-12	-	*A. nidulans*
			2.00E-07	-	*C. albicans*
			7.50E-38	-	*C. neoformans*
			5.50E-07	-	*D. hansenii*
			9.70E-05	Pex11 (0.014	*G. zeae*
			2.50E-07	-	*H. polymorpha*
			2.90E-13	-	*M. oryzae*
			1.60E-12	-	*N. crassa*
			5.30E-15	Pex11B (0.039)	*P. chrysogenum*
			3.50E-07	-	*S. pombe*
			2.20E-43	-	*U. maydis*
			3.10E-16	Pex11 (0.00016)	*Y. lipolytica*
Rhytidhysteron rufulum	6588	329	3.40E-14	-	*A. fumigatus*
			2.30E-14	-	*A. nidulans*
			6.40E-14	-	*C. neoformans*
			0.00015	-	*D. hansenii*
			5.90E-12	-	*G. zeae*
			2.60E-14	-	*H. polymorpha*
			7.40E-17	-	*M. oryzae*
			9.80E-16	-	*N. crassa*
			8.60E-17	-	*P. chrysogenum*
			1.10E-11	-	*U. maydis*
			1.40E-20	-	*Y. lipolytica*
	10149	315	3.60E-20	Pex11 (1.6E-5)	*A. fumigatus*
			6.70E-20	Pex11 (6.6E-5)	*A. nidulans*
			1.40E-18	-	*C. neoformans*
			2.90E-06	-	*H. polymorpha*
			6.00E-05	-	*M. oryzae*
			9.40E-06	-	*N. crassa*

			2.60E-24	-	*P. chrysogenum*
			1.40E-10	-	*U. maydis*
			3.20E-05	-	*Y. lipolytica*
Rozella allomycis	714	218	7.20E-07	-	*A. fumigatus*
			0.015	-	*A. nidulans*
			8.30E-08	-	*C. albicans*
			1.40E-11	-	*C. neoformans*
			5.20E-07	Pex11 (0.00027)	*D. hansenii*
			0.04	-	*G. zeae*
			6.30E-10	-	*H. polymorpha*
			4.40E-05	-	*M. oryzae*
			8.60E-06	-	*P. chrysogenum*
			0.00017	-	*S. pombe*
			3.00E-18	-	*U. maydis*
			9.00E-15	-	*Y. lipolytica*
Saitoella complicata	5664	315	2.00E-08	Pex11 (1.6E-5)	*A. fumigatus*
			3.90E-16	Pex11 (0.013)	*A. nidulans*
			6.80E-11	-	*C. albicans*
			3.30E-28	-	*C. neoformans*
			2.00E-12	-	*D. hansenii*
			6.30E-10	-	*G. zeae*
			1.80E-15	-	*H. polymorpha*
			1.10E-16	-	*M. oryzae*
			9.20E-17	-	*N. crassa*
			6.80E-09	Pex11 (0.015)	*P. chrysogenum*
			1.40E-18	-	*S. pombe*
			8.60E-33	-	*U. maydis*
			1.80E-21	-	*Y. lipolytica*
Schizophyllum commune	2642982	319	6.90E-08	-	*A. fumigatus*
			4.30E-07	-	*A. nidulans*

			5.80E-15	-	*C. neoformans*
			0.0083	-	*G. zeae*
			0.0025	-	*H. polymorpha*
			0.0078	-	*M. oryzae*
			0.0037	-	*N. crassa*
			5.20E-07	-	*P. chrysogenum*
			4.60E-64	-	*U. maydis*
			5.80E-11	-	*Y. lipolytica*
Schizosaccharomyces pombe[a]	SPAC1F12.04C	192	1.00E-05	-	*A. fumigatus*
			4.90E-07	-	*A. nidulans*
			0.00022	-	*C. albicans*
			3.80E-08	-	*C. neoformans*
			0.0034	-	*D. hansenii*
			1.30E-05	-	*H. polymorpha*
			1.90E-11	-	*M. oryzae*
			8.60E-11	-	*N.crassa*
			6.10E-07	-	*P. chrysogenum*
			0	-	*S. pombe*
			8.80E-12	-	*U. maydis*
			1.60E-10	-	*Y. lipolytica*
Spathaspora passalidarum	132131	303	1.30E-05	-	*A. fumigatus*
			1.20E-06	-	*A. nidulans*
			2.30E-129	-	*C. albicans*
			2.00E-136	-	*D. hansenii*
			0.00033	-	*G. zeae*
			5.00E-21	-	*H. polymorpha*
			2.60E-12	-	*M. oryzae*
			3.10E-13	-	*N. crassa*
			0.00021	-	*P. chrysogenum*
			0.0036	-	*S. pombe*

			3.60E-07	-	*U. maydis*
			2.00E-12	-	*Y. lipolytica*
Sphaerobolus stellatus	26914	283	3.90E-18	-	*A. fumigatus*
			1.80E-18	-	*A. nidulans*
			9.30E-07	-	*C. albicans*
			4.60E-29	-	*C. neoformans*
			2.30E-06	-	*D. hansenii*
			5.10E-10	-	*H. polymorpha*
			2.90E-11	-	*M. oryzae*
			4.30E-09	-	*N. crassa*
			5.80E-23	-	*P. chrysogenum*
			0.00024	-	*S. pombe*
			9.10E-80	-	*U. maydis*
			1.10E-19	Pex11 (0.002)	*Y. lipolytica*
Spizellomyces punctatus	SPPG_04635	298	4.80E-13	Pex11 (5.4E-7)	*A. fumigatus*
			3.10E-16	Pex11 (2E-6)	*A. nidulans*
			2.50E-29	Pex11 (6.2E-5)	*C. neoformans*
			5.30E-07	Pex11 (2E-6)	*G. zeae*
			9.90E-09	Pex11 (2.1E-5)	*H. polymorpha*
			3.60E-09	Pex11 (1.4E-7)	*N. crassa*
			1.50E-22	-	*P. chrysogenum*
			1.60E-09	Pex11 (5.6E-9)	*S. pombe*
			4.40E-38	Pex11 (2.4E-10)	*U. maydis*
			5.80E-20	Pex11 (1.7E-5)	*Y. lipolytica*
Thielavia terrestris	2119143	368	1.30E-17	-	*A. fumigatus*
			5.50E-19	-	*A. nidulans*
			1.80E-06	-	*C. albicans*
			3.60E-09	-	*C. neoformans*
			3.90E-06	-	*D. hansenii*
			3.80E-31	-	*G. zeae*

			2.20E-08	-	*H. polymorpha*
			2.70E-81	-	*M. oryzae*
			2.30E-83	-	*N. crassa*
			5.70E-17	-	*P. chrysogenum*
			0.00079	-	*S. pombe*
			2.30E-13	-	*U. maydis*
			2.80E-11	-	*Y. lipolytica*
Tremella mesenterica	32090	255	0.013	-	*C. albicans*
			2.20E-73	-	*C. neoformans*
			0.0007	-	*G. zeae*
			2.10E-05	-	*H. polymorpha*
			8.60E-08	-	*M. oryzae*
			3.70E-08	-	*N. crassa*
			0.038	-	*P. chrysogenum*
			2.70E-05	-	*S. pombe*
			7.20E-19	-	*U. maydis*
			1.10E-09	-	*Y. lipolytica*
Trichoderma reesei	3987	298	4.20E-17	-	*A. fumigatus*
			7.20E-10	-	*A. nidulans*
			2.20E-06	-	*C. albicans*
			1.00E-11	-	*C. neoformans*
			4.10E-05	-	*D. hansenii*
			6.50E-62	-	*G. zeae*
			8.90E-13	-	*H. polymorpha*
			1.70E-61	-	*M. oryzae*
			5.80E-46	-	*N. crassa*
			8.80E-10	-	*P. chrysogenum*
			4.90E-06	-	*S. pombe*
			2.90E-07	-	*U. maydis*
			9.80E-14	-	*Y. lipolytica*

Tuber melanosporum	6118	210	8.00E-18	-	*A. fumigatus*
			8.00E-17	-	*A. nidulans*
			7.40E-10	-	*C. albicans*
			1.10E-16	-	*C. neoformans*
			1.00E-11	-	*D. hansenii*
			1.10E-11	-	*G. zeae*
			8.20E-14	-	*H. polymorpha*
			2.00E-22	-	*M. orzyae*
			5.00E-29	-	*N. crassa*
			1.90E-17	-	*P. chrysogenum*
			3.20E-18	-	*S. pombe*
			2.30E-25	-	*U. maydis*
			6.30E-28	-	*Y. lipolytica*
Ustilago maydis[a]	UM00440	313	7.40E-09	Pex11 (0.0044)	*A. fumigatus*
			0.00017	-	*C. albicans*
			6.50E-25	-	*C. neoformans*
			9.40E-06	Pex11 (0.043)	*D. hansenii*
			0.00013	Pex11 (0.0091)	*G. zeae*
			1.90E-13	-	*H. polymorpha*
			4.10E-13	Pex11 (0.034)	*M. oryzae*
			8.90E-15	Pex11 (0.017)	*N. crassa*
			1.80E-12	Pex11 (0.00065)	*P. chrysogenum*
			1.60E-10	-	*S. pombe*
			2.90E-215	-	*U. maydis*
			2.40E-29	-	*Y. lipolytica*
Xanthoria parietina	7113	348	8.40E-22	-	*A. fumigatus*
			9.50E-16	Pex11 (0.0075)	*A. nidulans*
			6.60E-12	-	*C. albicans*
			1.40E-15	Pex11 (0.00077)	*C. neoformans*
			9.60E-06	-	*D. hansenii*

			0.00075	-	*G. zeae*
			3.70E-10	-	*H. polymorpha*
			1.20E-17	-	*M. oryzae*
			6.20E-18	-	*N. crassa*
			2.20E-23	-	*P. chrysogenum*
			5.80E-07	-	*S. pombe*
			1.00E-21	-	*U. maydis*
			2.10E-23	-	*Y. lipolytica*
Xylona heveae	243506	293	1.50E-37	-	*A. fumigatus*
			2.80E-16	-	*C. neoformans*
			2.00E-18	-	*G. zeae*
			2.50E-16	-	*H. polymorpha*
			5.90E-25	Pex11 (0.011)	*M. oryzae*
			1.00E-32	Pex11 (0.024)	*N. crassa*
			8.70E-40	-	*P. chrysogenum*
			1.50E-09	-	*S. pombe*
			1.50E-27	-	*U. maydis*
			1.70E-26	Pex11 (4.8E-6)	*Y. lipolytica*
Yarrowia lipolytica[a]	XP_501447.1	305	9.00E-07	-	*A. fumigatus*
			8.40E-11	-	*A. nidulans*
			2.50E-09	-	*C. albicans*
			2.20E-22	-	*C. neoformans*
			2.10E-08	-	*D. hansenii*
			8.10E-10	-	*G. zeae*
			1.40E-25	-	*H. polymorpha*
			9.60E-13	-	*M. oryzae*
			1.90E-14	-	*N. crassa*
			1.10E-11	-	*P. chrysogenum*
			5.80E-10	-	*S. pombe*
Holozoa			5.40E-30	-	*U. maydis*

Canis familiaris	XP_542110.1[c]	241	0.022	-	*A. fumigatus*
			0.0019	-	*A. nidulans*
Danio rerio	NP_001025343.1[c]	240	0.00039	Pex11 (0.00064)	*A. fumigatus*
			2.10E-05	Pex11 (0.029)	*H. polymorpha*
			0.00016	Pex11 (0.0013)	*P. chrysogenum*
Homo sapiens	NP_542393.1[c]	241	1.40E-05	-	*H. polymorpha*
			0.019	-	*Y. lipolytica*
Mus musculus	NP_081227.1[c]	241	0.042	-	*C. neoformans*
			9.40E-05	-	*H. polymorpha*
Nematostella vectensis	188278[c]	230	5.00E-05	-	*C. neoformans*
			4.20E-05	-	*H. polymorpha*
SAR					
Bigelowiella natans	89634	255	0.0017	-	*A. fumigatus*
			0.00068	-	*Y. lipolytica*
Ectocarpus siliculosus	Esi0000_0032[c]	250	0.0022	-	*C. neoformans*
			0.011	-	*P. chrysogenum*
			1.50E-06	-	*U. maydis*
			3.40E-06	-	*Y. lipolytica*
Phytophthora ramourum	82744[c]	244	9.90E-07	-	*A. nidualans*
			2.30E-08	-	*C. neoformans*
			1.60E-07	-	*P. chrysogenum*
			4.10E-06	-	*U. maydis*
			2.10E-06	-	*Y. lipolytica*
CCTH					
Guillardia theta	154547[c]	240	1.00E-06	-	*C. neoformans*

[a]Query

[c]Pex11 homologues identified with both Pex11 and Pex11C queries

Appendix 4: Pex11/25p homologues identified in comparative genomic survey

Genome	Accession	Length	R. pHMMer (*Y. lipolytica*)	R. pHMMer next hit (*Y. lipolytica*)
Fungi				
Agaricus bisporus	191941	362	1.30E-10	-
Allomyces macrogynus	AMAG_04379	268	1.00E-09	Pex11 (8.4E-6)
Ascoidea rubescens	100646	469	0.0015	-
Batrachochytrium dendrobatidis	BDEG_04173	379	2.30E-19	Pex11 (4.7E-7)
Catenaria anguillulae	341146	302	1.30E-09	-
Conidiobolus coronatus	80070	259	1.40E-13	Pex11 (0.0035)
Coprinopsis cinerea	CC1G_01265	374	9.80E-09	-
Cryptococcus neoformans	CNAG_04087	160	1.40E-05	-
Dekkera bruxellensis	8274	423	0.00031	-
Gonapodya prolifera	144677	285	0.00044	0.0014
	333839	271	1.20E-11	0.00018
Laccaria bicolor	XP_001876771.1	370	8.60E-09	-
Metschnikowia bicuspidata	39441	339	0.014	-
Mixia osmundae	96190	331	0.0028	-
Mucor circinelloides	190295	261	3.20E-25	Pex11 (1.9E-10)
	113603	216	4.50E-12	Pex11 (0.011)
Phycomyces blakesleeanus	160165	261	1.30E-27	Pex11 (2.2E-6)
	175339	248	2.20E-17	Pex11 (0.00083)
Pichia stipitis	56779	402	0.0016	-
Piromyces sp.	52850	225	1.20E-10	Pex11 (1.9E-5)
Rhizophagus irregularis	22141	259	8.50E-18	-
Rhizopus oryzae	RO3G_06403	248	8.00E-17	Pex11 (6.2E-8)
Rhodotorula graminis	53927	334	-	0.015
Rozella allomycis	5833	225	1.90E-16	-
Saitoella complicata	19474	260	1.30E-15	Pex11 (5.4E-5)
Schizophyllum commune	2730985	378	3.30E-07	-

Spathaspora passalidarum	157687	510	0.00017	-
Sphaerobolus stellatus	22728	396	6.20E-14	-
Spizellomyces punctatus	SPPG_03698		2.50E-14	Pex11 (1.2E-6)
Tremella mesenterica	65653	286	7.90E-09	-
Yarrowia lipolytica[a]	XP_503276.1	299	0	-
Archaeplastida				
Chlamydomonas reinhardtii	Cre12.g540500.t1.2[d]	238	0.00014	-
Volvox carteri	Vocar20006405m[d]	237	0.0052	-

[a]Query
[d]Pex11 homologues identified with both Pex11 and Pex11/25 queries

Appendix 5: Pex25p homologues identified in comparative genomic survey

Genome	Accession	Length	R. pHMMer (E-value) (*S. cerevisiae*)	R. pHMMer next hit (*S. cerevisiae*)
Fungi				
Ascoidea rubescens	100646	469	1.50E-19	-
Ashbya gossypii	NP_986892.1	323	1.20E-71	Pex27 (3.6E-5)
Candida albicans	XP_719254.1	539	3.40E-38	-
Candida glabrata	XP_447712.1	363	3.40E-93	Pex27 (9E-7)
Debaryomyces hansenii	XP_457136.1	497	3.30E-32	-
Dekkera bruxellensis	8274	423	4.50E-18	-
Hanseniaspora valbyensis	6447	380	1.30E-27	-
Hansenula polymorpha	15076	385	1.70E-21	Pex27 (0.046)
Kluyveromyces lactis	XP_455112.1	431	3.10E-71	Pex27 (0.00011)
Kluyveromyces waltii	Kwal_8479	398	2.80E-95	Pex27 (1.5E-10)
Metschnikowia bicuspidata	39441	339	3.00E-24	-
Pachysolen tannophilus	1661	422	5.80E-35	Pex27 (0.017)
Pichia membranifaciens	73142	447	9.60E-12	Pex27 (1.3E-5)
Pichia pastoris	XP_002492398.1	511	1.70E-28	-
Pichia stipitis	56779	402	1.10E-30	-
Rhizophagus irregularis	22141	259	0.0049	-
Saccharomyces bayanus	ORFP25324	401	1.40E-213	Pex27 (4.6E-11)
Saccharomyces castelli	ORFPScas_Contig605.4	350	3.10E-19	Pex27 (2.1E-6)
	ORFPScas_Contig700.15	376	4.10E-80	Pex27 (5.4E-8)
Saccharomyces cerevisiae[a]	NP_015213.1	395	0	Pex27 (5.3E-10)
Saccharomyces kluyveri	SAKL0H08492g	393	2.00E-104	Pex27 (9.7E-8)
Saccharomyces mikatae	ORFP24466	203	3.60E-109	Pex27 (2.7E-7)
	ORFP21858	394	2.30E-225	Pex27 (3.3E-10)
Saccharomyces paradoxus	ORFP22482	395	1.10E-247	Pex27 (4.8E-11)
Spathaspora passalidarum	157687	510	8.80E-35	-
Vanderwaltozyma polyspora	XP_001646454.1	418	8.60E-76	-

Zygosaccharomyces rouxii	XP_002497660.1	327	2.30E-73	Pex27 (2.6E-9)

[a]Query

Appendix 6: Pex27p homologues identified in comparative genomic survey

Genome	Accession	Length (aa)	R. pHMMer (*S. cerevisiae*)	R. pHMMer next hit (*S. cerevisiae*)
Fungi				
Saccharomyces bayanus	ORFP23699	374	8.20E-111	Pex25 (2.6E-9)
Saccharomyces cerevisiae[a]	NP_014836.3	377	0	Pex25 (5.6E-12)
Saccharomyces paradoxus	ORFP21062	230	2.10E-92	Pex25 (1.3E-7)

[a]Query

Appendix 7: *Homo sapiens*, *Saccharomyces cerevisiae* and *Neurospora crassa PEX* gene queries

Protein	*Homo sapiens* Queries	*Saccharomyces cerevisiae* Queries	*Neurospora crassa* Queries
Pex1	NP_000457.1	NP_012724.1 YKL197C (3627)	NCU08118
Pex2	NP_001165558.1	NP_012325.1 YJL210W (3220)	NCU02070
Pex3	NP_003621.1	NP_010616.3 YDR29C (1362)	NCU06175
Pex4	-	NP_011649.1 YGR133W (2709)	NCU02636
Pex5	XP_005253509.1 (Pex5L)	NP_010530.1 YCR244W (1273)	NCU02960
	XP_005253512.1 (Pex5S)	YMR018W (4693) (Pex5C)	
Pex6	NP_000278.1	NP_014070.1 YNL329C (5044)	NCU08373
Pex7	NP_000279.1	NP_010426.1 YDR142C (1166)	NCU07662
Pex8	-	NP_011591.1 YGR077C (2651)	NCU00032
Pex10	NP_722540.1	NP_010551.1 YDR265W (1295)	NCU03277
Pex11	NP_003838.1 (Pex11α)	NP_014494.1 YOL147C (5489)	NCU04802 (Pex11)
	NP_003837.1 (Pex11β)		NCU04045 (Pex11B)
	NP_542393.1 (Pex11γ)		NCU06275 (Pex11C)
Pex12	NP_000277.1	NP_013739.1 YMR026C (4701)	NCU05245
Pex13	NP_002609.1	NP_013292.1 YLR191W (4228)	NCU02618
Pex14	NP_004556.1	NP_011362.3 YGL153W (2413)	NCU03901
Pex15	-	NP_014598.1 YOL044W (5597)	-
Pex16	NP_004804.1	-	NCU01850
Pex17	-	NP_014185.1 YNL214W (5160)	-
Pex18	-	NP_012030.1 YHR160C (3128)	-
Pex19	NP_002848.1	YDL065C (951)	NCU04301
Pex20	-	-	NCU04062
Pex21	-	NP_011755.3 YGR239C (2820)	-
Pex22	-	NP_009346.1 YAL055W (21)	NCU03549
Pex23	-	NP_0101965.1 YER046W (1979) (Spo73)	NCU05564.7
Pex23L	-	-	NCU04604.7
Pex24	-	-	NCU06637

Pex25	-	NP_015213.1 YPL112C (6244)	-
Pex26	NP_00112112.1	-	NCU00347
Pex27	-	NP_014836.1 YOR193W (5851)	-
Pex28	-	NP_012020.1 YHR150W (3118)	-
Pex29	-	NP_010767.1 YDR479C (1520)	-
Pex30	-	NP_013428.1 YLR324W (4370)	-
Pex31	-	NP_011518.1 YGR004W (2575)	-
Pex32	-	NP_009727.3 YBR168W (426)	-
Pex33	-	-	NCU01535
Pex34	-	NP_009875.1 YCL056C (589)	-

Appendix 8: *Arabidopsis thaliana PEX* genes (adapted from Nito et al., 2007)

Protein	ID	R. pHMMer (*H. sapiens*)	R. pHMMer (*S. cerevisiae*)	Notes
Pex1	AT5G08470	4.20E-135	-	
Pex2	AT1G79810	1.20E-41	2.30E-08	
Pex3-1	AT3G18160	9.30E-19	4.20E-05	
Pex3-2	AT1G48635	5.20E-22	0.00015	
Pex4	AT5G25760	-	1.70E-32	
Pex5	AT5G56290	1.20E-66	7.60E-40	
Pex6	AT1G03000	-	-	Validated in NR
Pex7	AT1G29260	8.00E-87	5.30E-61	
Pex8	-			
Pex10	AT2G26350	1.20E-42	1.90E-36	
Pex11A	AT1G47750	1.60E-07	-	
Pex11B	AT3G47430	1.20E-07	-	
Pex11C	AT1G01820	-	0.0002	
Pex11D	AT2G45740	0.0014	0.00096	
Pex11E	AT3G61070	0.01	0.0025	
Pex12	AT3G04460	2.00E-42	4.30E-16	
Pex13	AT3G07560	-	0.0065	
Pex14	AT5G62810	1.30E-09	-	
Pex15	-			
Pex16	AT2G45690	1.80E-15	-	
Pex17	-	-	-	Omit AT4G18197 (Nito et al. 2007)
Pex18	-			
Pex19-1	AT3G03490	1.10E-16	4.60E-19	
Pex19-2	AT5G17550	1.50E-18	3.50E-09	
Pex20	-			
Pex21	-			
Pex22	AT3G21865	-	-	Validatd in NR

Pex23/Pex30	-		
Pex24/Pex28	-		
Pex25	-		
Pex26	-		
Pex27	-		
Pex29	-		
Pex31	-		
Pex32	-		
Pex33	-		
Pex34	-		
Apem9 (Pex15/Pex26)	AT3G10572	-	-

Appendix 9: *Phaeodactylum tricornutum PEX* genes (adpated from Gonzalez et al. 2011)

Protein	ID	R. pHMMer (*H. sapiens*)	R. pHMMer (*S. cerevisiae*)	R. pHMMer (*N. crassa*)	R. pHMMer (*A. thaliana*)	Notes
Pex1	-					Omit 14397 (Gonzalez et al. 2011)
Pex2	-					Omit 49301 (Gonzalez et al. 2011)
Pex3	50623	9.30E-09	2.50E-05	3.20E-09	2.50E-05	
Pex4	47555	-	-	-	1.20E-52	
Pex5	32173	1.10E-42	-	1.50E-28	1.10E-63	
Pex10	47516	8.50E-13	3.20E-18	5.90E-10	2.20E-11	
Pex11	44128	-	5.40E-08	3.80E-25	5.40E-08	
Pex12	49405	2.30E-19	1.10E-06	-	-	
Pex16	47406	4.00E-05	-	4.60E-03	2.10E-05	
Pex19	31927	1.50E-14	3.40E-05	2.30E-13	1.60E-29	

Appendix 10: *Phytophthora ramorum PEX* genes

Protein	ID	R. pHMMer (*H. sapiens*)	R. pHMMer (*S. cerevisiae*)	R. pHMMer (*N. crassa*)	R. pHMMer (*A. thaliana*)
Pex2	93459	2.50E-30	4.50E-07	2.90E-19	5.90E-37
Pex3	80388	9.70E-09	-	1.00E-06	1.20E-05
Pex4	49028	-	-	-	1.70E-52
Pex5	74660	-	8.60E-39	9.10E-46	1.60E-82
Pex7	94244	1.00E-71	2.10E-54	7.40E-64	9.10E-83
Pex10	85290	1.60E-07	-	1.40E-09	6.20E-16
Pex11	82744	2.60E-09	-	-	-
Pex12	93734	6.20E-32	4.40E-09	1.50E-20	1.10E-28
Pex14	74639	8.50E-09	1.90E-02	1.70E-09	1.20E-15
Pex16	76428	3.70E-13	-	9.70E-14	2.00E-15
Pex19	77483	1.20E-19	7.30E-07	5.70E-22	4.80E-37
Pex22	77773	-	-	-	1.00E-09

Appendix 11: *Thalassiosira pseudonana PEX* genes

Protein	ID	R. pHMMer (H. sapiens)	R. pHMMer (S. cerevisiae)	R. pHMMer (N. crassa)	R. pHMMer (A. thaliana)
Pex2	4404	8.30E-07	-	-	8.20E-07
Pex3	263695	1.10E-07	3.80E-02	3.50E-09	-
Pex4	23095	-	-	-	1.90E-54
Pex5	264438	9.50E-45	3.90E-29	7.60E-28	1.60E-63
Pex10	260886	6.40E-14	2.40E-15	3.10E-15	1.10E-12
Pex19	37854	9.80E-13	9.60E-06	6.40E-12	6.30E-25

Appendix 12: *Blastocystis hominis PEX* genes

Protein	Accession	R. BLAST (Genome)	Notes
Pex4	blasto.m3763	6E-44 (*A. thaliana*) 3E-42 (*P. tricornutum*) 4E-42 (*P. ramorum*) 1E-41 (*T. pseudonana*)	Possible Pex4, annotaed as ubiquitin-conjugating enzyme

Appendix 13A: *Naegleria gruberi PEX* genes

Protein	ID	Length	R. pHMMer (*H. sapiens*)	R. pHMMer (*S. cerevisiae*)	R. pHMMer (*N. crassa*)	R. pHMMer (*A. thaliana*)
Pex1	77892	293	4.3E-10	0.00046	0.0041	5.3E-10
Pex3	68465	381	3.9E-24	-	9.1E-12	4.6E-15
Pex4	63289	162	-	-	-	3.7E-58
Pex5	64046	470	-	4.4E-19	8.7E-17	5.5E-44
Pex7	76903	247	6.6E-50	-	2.2E-24	4.2E-63
Pex10	65721	430	4.5E-32	5.1E-21	1.5E-34	1.5E-37
Pex11	45694	256	0.0034	0.015	-	-
Pex12	31793	370	7.9E-47	1.7E-15	1.1E-26	1.2E-50
Pex19	75824	284	8.8E-21	0.00000001	7.1E-74	4.7E-17

Appendix 13B: *Naegleria gruberi* glycolysis enzymes (adapted from Fritz-Laylin et al., 2010)

Protein	*N. gruberi* ID	R. pHMMer (*T. brucei*)	PTS
fructose-bisphosphate aldolase (ALD)	56383	5.80E-95	-
glucose-6-phosphate isomerase (GPI)	30686	4.70E-114	-
glyceraldehyde 3-phosphate dehydrogenase (GAPDH)	53883	7.50E-137	-
glycerol kinase (GK)	38161	1.20E-146	-
glycerol-3-phosphate dehydrogenase [NAD+] (G3PDH)	29597	1.30E-100	-
	34539	1.40E-58	-
	80825	6.70E-109	-
hexokinase (HK) (glucokinase)	81163	-	-
phosphofructokinase (PFK)	35679	2.00E-31	-
phosphoglycerate kinase (PGK)	81218	1.20E-108	-
triosephosphate isomerase (TPI)	29287	1.50E-68	-

Appendix 14A: *Bodo saltans PEX* genes

Protein	ID	Length	R. pHMMer (*H. sapiens*)	R. pHMMer (*S. cerevisiae*)	R. pHMMer (*N. crassa*)	R. pHMMer (*A. thaliana*)	R. pHMMer (*N. gruberi*)	R. pHMMer (*T. brucei*)
Pex2	BS79395	332	2.70E-20	1.40E-06	7.30E-21	1.70E-30	4.40E-02	-
Pex4	BS87965c	180	-	-	-	5.70E-52	7.60E-43	-
Pex5	BS07445	631	1.40E-69	5.60E-43	1.10E-44	1.00E-57	2.00E-17	-
Pex6	BS66885	360	-	-	-	-	-	1.20E-144
Pex7	BS21885	404	1.50E-45	6.50E-43	1.80E-41	5.70E-62	1.50E-46	-
Pex10	BS79455	-	6.20E-23	-	-	6.80E-26	-	-
Pex11	BS66630	216	-	-	4.00E-06	1.50E-06	6.40E-04	-
Pex12	BS71105	321	2.40E-12	-	4.00E-06	2.50E-19	-	-
Pex14	BS68470	491	4.70E-05	1.20E-04	-	4.60E-07	-	-
Pex16	BS28750	997	-	-	-	-	-	7.20E-32
Pex19	BS22070	324	1.90E-03	-	1.80E-04	3.20E-03	2.60E-04	-

Appendix 14B: *Bodo saltans* glycolysis enzymes

Protein	*B. saltans* ID	R. pHMMer (*T. brucei*)	PTS
fructose-bisphosphate aldolase (ALD)	BS75865	1E-193	-
glucose-6-phosphate isomerase (GPI)	-		
glyceraldehyde 3-phosphate dehydrogenase (GAPDH)	BS66180	5.2E-203	PTS1 (AKL)
glycerol kinase (GK)	BS38775	7.60E-198	PTS1 (AKL)
glycerol-3-phosphate dehydrogenase [NAD+] (G3PDH)	BS28440	1.8E-36	-
hexokinase (HK)	BS73805	1.6E-155	-
phosphofructokinase (PFK)	BS33550	0.00089	PTS1 (SRL)
	BS59050	2.4E-196	PTS1 (SKL)
phosphoglycerate kinase (PGK)	BS70390	1E-139	-
	BS14730	8.8E-205	-
triosephosphate isomerase (TPI)	BS85785	2.9E-113	-

Appendix 15A: *Bodo saltans* proteins with named *T. brucei* glycosomal homologues

B. saltans ID	R. pHMMer (*T. brucei*)	Top *T. brucei* ID	*T. brucei* annotation	*B. saltans* PTS	[Leader]PTS
BS14730	8.80E-205	Tb927.1.720	Phosphoglycerate kinase	-	-
BS19010	6.00E-37	Tb927.1.3220	GTPase activating protein of Rab-like GTPase	-	-
BS19015	1.70E-39	Tb927.1.3220	GTPase activating protein of Rab-like GTPase	-	-
BS00795	3.40E-20	Tb927.1.4490	acetyltransferase	-	-
BS09660	2.00E-24	Tb927.2.4130	enoyl-CoA hydratase/Enoyl-CoA isomerase/3-hydroxyacyl-CoA dehydrogenase	-	-
BS26355	1.10E-49	Tb927.2.4130	enoyl-CoA hydratase/Enoyl-CoA isomerase/3-hydroxyacyl-CoA dehydrogenase	-	-
BS71405	7.50E-54	Tb927.2.4130	enoyl-CoA hydratase/Enoyl-CoA isomerase/3-hydroxyacyl-CoA dehydrogenase	-	-
BS36970	3.70E-266	Tb927.2.4210	aldehyde dehydrogenase	-	-
BS36975	3.70E-266	Tb927.2.4210	aldehyde dehydrogenase	-	-
BS22930	2.40E-94	Tb927.2.5800	sedoheptulose-1,7-bisphosphatase	-	-
BS51820	1.10E-114	Tb927.3.1840	3-oxo-5-alpha-steroid 4-dehydrogenase	-	-
BS79395	3.40E-58	Tb927.3.2340	PEX2	-	-
BS33550	0.00089	Tb927.3.3270	phosphofructokinase	PTS1	SRL
BS59050	2.40E-196	Tb927.3.3270	phosphofructokinase	PTS1	SKL
BS21885	2.30E-130	Tb927.3.3610	PEX7	-	-
BS26075	2.60E-23	Tb927.3.3760	tryparedoxin	-	-
BS37855c	1.50E-41	Tb927.3.4820	acyltransferase	-	-
BS54515	1.80E-56	Tb927.3.4820	acyltransferase	-	-
BS58630	6.20E-41	Tb927.3.4820	acyltransferase	-	-
BS00325	1.70E-115	Tb927.4.1250	PEX1	-	-
BS44545	8.80E-17	Tb927.4.2010	acyl-CoA binding protein	-	-
BS71650	5.20E-159	Tb927.4.3160	dihydroxyacetone phosphate acyltransferase	-	-
BS00240	5.00E-55	Tb927.4.4050	ABC transporter	-	-
BS00245	5.00E-55	Tb927.4.4050	ABC transporter	-	-

BS84040	3.00E-93	Tb927.4.4050	ABC transporter	-	-
BS51740	1.80E-121	Tb927.4.4070	mevalonate kinase	PTS1	AKL
BS59170	1.80E-97	Tb927.4.4360	monoglyceride lipase	-	-
BS58430	0.00E+00	Tb927.5.930	NADH-dependent fumarate reductase	-	-
BS76915	0.00E+00	Tb927.5.930	NADH-dependent fumarate reductase	PTS1	SKL
BS07445	6.70E-110	Tb927.5.1100	PEX5	-	-
BS16935	1.70E-72	Tb927.5.1210	short-chain dehydrogenase	-	-
BS25075	2.40E-04	Tb927.5.1210	short-chain dehydrogenase	-	-
BS29735	5.00E-17	Tb927.5.1210	short-chain dehydrogenase	-	-
BS41100	9.20E-28	Tb927.5.1210	short-chain dehydrogenase	-	-
BS61710	2.70E-86	Tb927.5.1210	short-chain dehydrogenase	-	-
BS61735	3.30E-113	Tb927.5.1210	short-chain dehydrogenase	-	-
BS90305	3.70E-19	Tb927.5.1210	short-chain dehydrogenase	-	-
BS16145	7.20E-216	Tb927.5.2080	guanosine monophosphate reductase	-	-
BS67110	4.70E-04	Tb927.5.2370	hydrolase, alpha/beta fold family	-	-
BS52325	1.10E-07	Tb927.5.2650	L-galactonolactone oxidase	-	-
BS55925	3.30E-04	Tb927.5.2650	L-galactonolactone oxidase	-	-
BS72040	9.40E-10	Tb927.5.2650	L-galactonolactone oxidase	-	-
BS89870	5.00E-152	Tb927.5.2650	L-galactonolactone oxidase	-	-
BS89875	5.00E-152	Tb927.5.2650	L-galactonolactone oxidase	-	-
BS66885	1.20E-144	Tb927.5.3920	PEX6	-	-
BS76790	9.30E-36	Tb927.5.4350	NUDIX hydrolase	-	-
BS15545	1.60E-13	Tb927.6.710	dephospho-CoA kinase	-	-
BS92540	9.10E-26	Tb927.6.710	dephospho-CoA kinase	-	-
BS36665	2.00E-69	Tb927.6.1080	glyoxalase II	-	-
BS21955	1.90E-20	Tb927.6.1500	alkyl-dihydroxyacetone phosphate synthase	-	-
BS46785	6.90E-271	Tb927.6.1501	alkyl-dihydroxyacetone phosphate synthase	PTS1	AKL
BS55370	3.90E-20	Tb927.6.1500	alkyl-dihydroxyacetone phosphate synthase	-	-
BS17560	7.50E-133	Tb927.6.1800	protein phosphatase 2C	-	-
BS73255	2.90E-74	Tb927.6.2540	DREV methyltransferase	-	-

BS48145	5.30E-153	Tb927.6.3050	aldehyde dehydrogenase family	-	-
BS25870	1.50E-32	Tb927.6.3840	reticulon domain protein	-	-
BS24955	4.80E-02	Tb927.6.4210	aldehyde dehydrogenase	-	-
BS66265	3.00E-267	Tb927.6.4210	aldehyde dehydrogenase	-	-
BS73870c	6.30E-39	Tb927.6.4210	aldehyde dehydrogenase	-	-
BS73875c	3.70E-58	Tb927.6.4210	aldehyde dehydrogenase	-	-
BS66180	5.20E-203	Tb927.6.4280	glyceraldehyde 3-phosphate dehydrogenase	PTS1	AKL
BS42355	1.50E-23	Tb927.7.540	chaperone protein DNAj	-	-
BS59885	8.10E-43	Tb927.7.1130	trypanothione/tryparedoxin dependent peroxidase 2	-	-
BS59890	2.90E-82	Tb927.7.1130	trypanothione/tryparedoxin dependent peroxidase 2	-	-
BS59945	2.20E-82	Tb927.7.1130	trypanothione/tryparedoxin dependent peroxidase 2	-	-
BS77340	1.20E-81	Tb927.7.1130	trypanothione/tryparedoxin dependent peroxidase 2	-	-
BS59885	3.00E-42	Tb927.7.1140	trypanothione/tryparedoxin dependent peroxidase 3	-	-
BS59890	6.70E-82	Tb927.7.1140	trypanothione/tryparedoxin dependent peroxidase 3	-	-
BS59945	2.50E-81	Tb927.7.1140	trypanothione/tryparedoxin dependent peroxidase 3	-	-
BS77340	3.20E-81	Tb927.7.1140	trypanothione/tryparedoxin dependent peroxidase 3	-	-
BS73410	2.20E-112	Tb927.7.3500	glutathione-S-transferase/glutaredoxin	-	-
BS88975	7.50E-53	Tb927.7.3940	mitochondrial carrier protein	-	-
BS80950	1.30E-84	Tb927.7.4770	cyclophilin-type peptidyl-prolyl cis-trans isomerase	-	-
BS47420	5.30E-95	Tb927.7.5680	deoxyribose-phosphate aldolase	-	-
BS84590	5.20E-72	Tb927.7.6200	chaperone protein DNAj	-	-
BS41425	2.10E-05	Tb927.8.920	ubiquitin-conjugating enzyme e2	-	-
BS87965c	7.80E-69	Tb927.8.920	ubiquitin-conjugating enzyme e2	-	-
BS93740	1.70E-70	Tb927.8.1910	acetylornithine deacetylase, metallo-peptidase, Clan MH, Family M18	-	-
BS28440	1.80E-36	Tb927.8.3530	glycerol-3-phosphate dehydrogenase [NAD+]	-	-
BS54700	2.30E-58	Tb927.8.6390	lysophospholipase, alpha/beta hydrolase	-	-
BS69965	2.30E-30	Tb927.8.6390	lysophospholipase, alpha/beta hydrolase	PTS1p	[LSGDRTAAK]AGL
BS77675	2.20E-60	Tb927.8.6390	lysophospholipase, alpha/beta hydrolase	-	-
BS77685	5.70E-81	Tb927.8.6390	lysophospholipase, alpha/beta hydrolase	-	-

BS77690	5.70E-81	Tb927.8.6390	lysophospholipase, alpha/beta hydrolase	-	-
BS71310	2.20E-41	Tb927.8.7170	inositol polyphosphate 1-phosphatase	-	-
BS68470	8.50E-35	Tb927.10.240	PEX14	-	-
BS19810	0	Tb927.10.1170	intraflagellar transport protein IFT172	-	-
BS29675	5.10E-32	Tb927.10.1390	hypoxanthine-guanine phosphoribosyltransferase	-	-
BS29690	5.50E-47	Tb927.10.1400	glycyl-tRNA synthetase	PTS1	SKL
BS73805	1.60E-155	Tb927.10.2010	hexokinase	-	-
BS73805	1.60E-155	Tb927.10.2020	hexokinase	-	-
BS03280	4.60E-201	Tb927.10.2490	glucose-6-phosphate 1-dehydrogenase	-	-
BS29750	1.10E-06	Tb927.10.2490	glucose-6-phosphate 1-dehydrogenase	-	-
BS29755	9.40E-197	Tb927.10.2490	glucose-6-phosphate 1-dehydrogenase	-	-
BS22320	0.035	Tb927.10.4610	dolicholphosphate-mannose synthase	-	-
BS33450	1.50E-144	Tb927.10.4610	dolicholphosphate-mannose synthase	-	-
BS75865	1.00E-193	Tb927.10.5620	fructose-bisphosphate aldolase	-	-
BS19220	1.20E-52	Tb927.10.5760	adenylate kinase	PTS1p	[GLGEPTAPK]GKL
BS52165	3.00E-65	Tb927.10.6150	neurobeachin/beige protein	-	-
BS71910	0	Tb927.10.6150	neurobeachin/beige protein	-	-
BS43185	3.50E-51	Tb927.10.6610	chaperone protein DNAj	-	-
BS30800	7.60E-38	Tb927.10.6810	guanylate kinase	-	-
BS38015	3.60E-10	Tb927.10.7100	delta-4 fatty acid desaturase	-	-
BS75450	1.70E-142	Tb927.10.7100	delta-4 fatty acid desaturase	-	-
BS00490	1.20E-25	Tb927.10.10390	trypanothione reductase	-	-
BS00500	2.20E-13	Tb927.10.10390	trypanothione reductase	-	-
BS06235	7.40E-239	Tb927.10.10390	trypanothione reductase	-	-
BS06155	5.20E-72	Tb927.10.10610	protein tyrosine phosphatase	-	-
BS27085	4.60E-30	Tb927.10.10610	protein tyrosine phosphatase	-	-
BS24685	5.30E-143	Tb927.10.12240	short-chain dehydrogenase	-	-
BS17295	3.40E-173	Tb927.10.13130	UTP-glucose-1-phosphate uridylyltransferase 2	-	-
BS69630	1.60E-23	Tb927.10.14720	PEX13	-	-
BS65465	1.10E-125	Tb927.10.15410	glycosomal malate dehydrogenase	-	-

BS71105	1.80E-25	Tb927.10.15850	peroxisome assembly protein (PEX2/PEX12)	-	-
BS71070	2.80E-233	Tb927.10.16120	inosine-5-monophosphate dehydrogenase	-	-
BS50905	7.20E-39	Tb09.211.2730	GIM5	-	-
BS22170	0.00055	Tb927.11.10020	short-chain dehydrogenase	-	-
BS23565	1.10E-05	Tb927.11.10020	short-chain dehydrogenase	-	-
BS23570	6.80E-41	Tb927.11.10020	short-chain dehydrogenase	-	-
BS85785	2.90E-113	Tb927.11.5520	triosephosphate isomerase	-	-
BS66630	3.80E-57	Tb927.11.11520	PEX11	-	-
BS50905	7.20E-39	Tb927.9.11600	GIM5	-	-
BS64125	7.90E-53	Tb927.11.14780	phosphomannose isomerase	-	-
BS41170	4.10E-104	Tb927.9.8720	fructose-1,6-bisphosphatase	-	-
BS52270	2.90E-216	Tb927.11.900	isocitrate dehydrogenase	-	-
BS11805	6.20E-65	Tb927.11.3130	ABC transporter, adrenoleukodystrophy protein homolog	-	-
BS54900	3.50E-65	Tb927.11.3130	ABC transporter, adrenoleukodystrophy protein homolog	-	-
BS57535	0	Tb927.11.6280	pyruvate phosphate dikinase	PTS1	SKL
BS34125	1.30E-24	Tb927.11.3850	AMP deaminase	-	-
BS34130	2.20E-23	Tb927.11.3850	AMP deaminase	-	-
BS34135	1.80E-211	Tb927.11.3850	AMP deaminase	-	-
BS12755	2.90E-44	Tb927.9.6100	TFIIF-stimulated CTD phosphatase	-	-
BS12780	6.50E-49	Tb927.9.6100	TFIIF-stimulated CTD phosphatase	-	-
BS22075	5.10E-72	Tb927.9.6100	TFIIF-stimulated CTD phosphatase	-	-
BS25185	5.90E-42	Tb927.9.6100	TFIIF-stimulated CTD phosphatase	-	-
BS44750	5.70E-48	Tb927.9.6100	TFIIF-stimulated CTD phosphatase	-	-
BS79620	4.80E-37	Tb927.9.6100	TFIIF-stimulated CTD phosphatase	-	-
BS89460	1.70E-42	Tb927.9.6100	TFIIF-stimulated CTD phosphatase	PTS1p	[DEQDKATAF]SKY
BS89470	5.10E-12	Tb927.9.6100	TFIIF-stimulated CTD phosphatase	-	-
BS26845	0.0027	Tb927.11.1020	ribokinase	-	-
BS46460	5.60E-120	Tb927.11.1020	ribokinase	-	-
BS08690	0.0032	Tb927.11.15020	iron superoxide dismutase	-	-
BS63740	4.00E-105	Tb927.11.15020	iron superoxide dismutase	-	-

BS46545	5.80E-132	Tb927.11.1070	ABC transporter	PTS1p	[DWSVISRER]APL
BS70390	1.00E-139	Tb927.11.2380	phosphoglycerate kinase	-	-
BS38775	7.60E-198	Tb927.9.12570	glycerol kinase	PTS1	AKL
BS37440	4.40E-173	Tb927.9.9000	isopentenyl-diphosphate delta-isomerase	-	-
BS55675	1.50E-06	Tb927.11.2520	UTP-glucose-1-phosphate uridylyltransferase 2	-	-
BS70450	2.40E-97	Tb927.11.2520	UTP-glucose-1-phosphate uridylyltransferase 2	-	-
BS83135	8.90E-45	Tb927.11.6330	splicing factor Prp31, 6-phosphogluconolactonase (6PGL)	-	-
BS11525	8.20E-22	Tb927.9.14070	short-chain dehydrogenase	-	-
BS13460	1.00E-76	Tb927.9.14070	short-chain dehydrogenase	-	-
BS36770	5.90E-07	Tb927.9.14070	short-chain dehydrogenase	PTS1	SKL
BS65560	3.10E-10	Tb927.9.14070	short-chain dehydrogenase	-	-
BS70800	1.90E-07	Tb927.9.14070	short-chain dehydrogenase	-	-
BS18805	2.00E-07	Tb927.9.4190	fatty acyl CoA syntetase 1	-	-
BS26105	1.60E-08	Tb927.9.4190	fatty acyl CoA syntetase 1	-	-
BS50005	5.50E-113	Tb927.9.4190	fatty acyl CoA syntetase 1	-	-
BS26175	4.20E-58	Tb927.9.12700	phospholipase A1	-	-
BS35735	3.10E-165	Tb927.9.10310	mitochondrial carrier protein (phosphate transporter)	-	-
BS01700	5.70E-239	Tb927.11.9920	polyubiquitin	-	-
BS05340	0	Tb927.11.9920	polyubiquitin	-	-
BS07810	1.10E-31	Tb927.11.9920	polyubiquitin	-	-
BS15305	5.40E-296	Tb927.11.9920	polyubiquitin	-	-
BS18300	2.40E-117	Tb927.11.9920	polyubiquitin	-	-
BS55365	3.20E-52	Tb927.11.9920	polyubiquitin	-	-
BS77240	6.90E-81	Tb927.11.9920	polyubiquitin	-	-
BS82160	2.40E-52	Tb927.11.9920	polyubiquitin	-	-
BS44870	8.50E-30	Tb927.11.9810	NUDIX hydrolase	-	-
BS05665	9.00E-64	Tb927.9.3280	acidocalcisomal exopolyphosphatase	-	-
BS22070	1.90E-06	Tb927.9.12290	PEX19	-	-
BS76225	2.00E-75	Tb927.11.12240	ubiquitin carboxyl-terminal hydrolase, cysteine peptidase, Clan CA, family C19	-	-

BS94640	7.10E-127	Tb927.9.8950	CAAX prenyl protease 1, metallo- peptidase, Clan M-Family M48	-	-
BS93205	5.50E-71	Tb927.11.9420	ATP synthase	-	-
BS72150	2.40E-169	Tb927.11.760	protein phosphatase 2C	-	-
BS26575	5.50E-125	Tb927.11.2730	UDP-galactose 4-epimerase	-	-
BS78840	1.70E-06	Tb927.11.2730	UDP-galactose 4-epimerase	-	-
BS73975	2.00E-28	Tb927.11.5220	chaperone protein DNAj	-	-
BS73980	2.50E-17	Tb927.11.5220	chaperone protein DNAj	-	-
BS81680	2.10E-39	Tb927.11.4700	prostaglandin f synthase	-	-
BS05850	2.10E-58	Tb927.11.14080	cyclin 2, G1 cyclin	-	-
BS59590	0.0001	Tb927.11.14080	cyclin 2, G1 cyclin	-	-
BS59785	8.00E-35	Tb927.11.14080	cyclin 2, G1 cyclin	-	-
BS00445	6.60E-66	Tb927.11.540	ABC transporter	-	-
BS09565	1.40E-71	Tb927.11.540	ABC transporter	-	-
BS28085	1.60E-63	Tb927.11.540	ABC transporter	-	-
BS29865	4.70E-09	Tb927.11.540	ABC transporter	-	-
BS59625	8.20E-275	Tb927.11.540	ABC transporter	-	-
BS58265	3.50E-192	Tb927.11.3270	squalene monooxygenase	-	-
BS05220	3.40E-191	Tb927.9.11030	protein kinase A catalytic subunit	-	-
BS86295	0.0016	Tb927.11.9590	S-adenosylhomocysteine hydrolase	-	-
BS16090	1.00E-05	Tb927.11.880	cyclophilin a, cyclophilin type peptidyl-prolyl cis-trans isomerase	-	-
BS66570	3.00E-10	Tb927.11.880	cyclophilin a, cyclophilin type peptidyl-prolyl cis-trans isomerase	-	-
BS85385	3.80E-84	Tb927.11.880	cyclophilin a, cyclophilin type peptidyl-prolyl cis-trans isomerase	-	-
BS85390	2.70E-99	Tb927.11.880	cyclophilin a, cyclophilin type peptidyl-prolyl cis-trans isomerase	-	-
BS68990	2.00E-263	Tb927.11.7290	pantothenate kinase subunit	-	-
BS24585	3.20E-132	Tb927.11.9530	14-3-3-1 protein	-	-

BS54965	2.30E-171	Tb927.11.3030	phosphoribosylpyrophosphate synthetase	-	-
BS57745	3.40E-116	Tb927.11.3030	phosphoribosylpyrophosphate synthetase	-	-
BS02820	1.60E-16	Tb927.11.11330	HSP70	-	-
BS04270	0	Tb927.11.11330	HSP70	-	-
BS34010	3.10E-228	Tb927.11.11330	HSP70	-	-
BS62755	3.70E-168	Tb927.11.11330	HSP70	-	-
BS78260	3.80E-95	Tb927.11.11330	HSP70	-	-

Appendix 15B: *Bodo saltans* proteins with unnamed *T. brucei* glycosomal homologues

B. saltans ID	R. pHMMer (*T. brucei*)	Top *T. brucei* ID	*T. brucei* annotation	*B. saltans* PTS	[Leader]PTS
BS91045	7.70E-15	Tb927.2.5200	hypothetical protein	-	-
BS64635	2.30E-02	Tb927.3.640	hypothetical protein	-	-
BS10960	7.10E-96	Tb927.3.1080	hypothetical protein	-	-
BS54070	3.40E-138	Tb927.3.4420	hypothetical protein	PTS1	ATL
BS53545	3.10E-46	Tb927.4.1360	hypothetical protein	-	-
BS48565	5.90E-37	Tb927.4.1540	hypothetical protein	-	-
BS53335	1.80E-157	Tb927.4.1540	hypothetical protein	-	-
BS54830	2.50E-18	Tb927.4.1670	hypothetical protein	-	-
BS33020	1.50E-19	Tb927.4.2320	hypothetical protein	-	-
BS07750	2.20E-43	Tb927.5.1630	hypothetical protein	-	-
BS09130	4.60E-81	Tb927.5.2590	hypothetical protein	-	-
BS23960	3.40E-46	Tb927.5.2590	hypothetical protein	-	-
BS23965	2.00E-06	Tb927.5.2590	hypothetical protein	-	-
BS23970	6.40E-12	Tb927.5.2590	hypothetical protein	-	-
BS23975	7.30E-43	Tb927.5.2590	hypothetical protein	-	-
BS10265	1.90E-42	Tb927.6.2200	hypothetical protein	PTS1p	[LVPPPSLLK]SKV
BS90235	5.70E-48	Tb927.6.2600	hypothetical protein	-	-
BS42465	2.10E-09	Tb927.7.330	hypothetical protein	PTS1	AKL
BS40610	1.90E-13	Tb927.7.900	hypothetical protein	-	-
BS03600	1.20E-06	Tb927.7.1630	hypothetical protein	-	-
BS59935	1.40E-04	Tb927.7.2160	hypothetical protein	-	-
BS92535	3.20E-02	Tb927.7.3770	hypothetical protein	-	-
BS06045	2.60E-30	Tb927.7.6770	hypothetical protein	-	-
BS74530	5.60E-64	Tb927.8.1370	hypothetical protein	-	-
BS21635	4.50E-55	Tb927.8.3280	hypothetical protein	-	-
BS29235	4.30E-75	Tb927.8.4370	hypothetical protein	-	-
BS33430	1.70E-133	Tb927.8.6080	hypothetical protein	-	-

BS93155	3.20E-09	Tb927.8.6080	hypothetical protein	-	-
BS93160	5.80E-12	Tb927.8.6080	hypothetical protein	-	-
BS93165	4.50E-09	Tb927.8.6080	hypothetical protein	-	-
BS29005	3.70E-12	Tb927.8.6620	(no annotation)	-	-
BS53820	1.20E-148	Tb927.8.6640	hypothetical protein	-	-
BS72250	5.20E-19	Tb927.8.6780	hypothetical protein	PTS1	SKL
BS77425	5.30E-07	Tb927.10.1280	hypothetical protein	-	-
BS77430	1.80E-08	Tb927.10.1280	hypothetical protein	-	-
BS21135	2.80E-80	Tb927.10.3080	hypothetical protein	-	-
BS58475	5.40E-08	Tb927.10.3880	hypothetical protein	-	-
BS77830	2.00E-18	Tb927.10.3880	hypothetical protein	-	-
BS78410	2.00E-47	Tb927.10.5740	hypothetical protein	-	-
BS33135	7.20E-19	Tb927.10.7480	hypothetical protein	-	-
BS47145	6.00E-20	Tb927.10.8410	hypothetical protein	-	-
BS47320	3.50E-68	Tb927.10.8630	hypothetical protein	-	-
BS84045	1.90E-25	Tb927.10.9890	hypothetical protein	-	-
BS08845	0.0012	Tb927.10.11810	hypothetical protein	-	-
BS21690	6.30E-55	Tb927.10.12360	hypothetical protein	-	-
BS75135	9.70E-72	Tb927.10.14020	hypothetical protein	-	-
BS49690	3.60E-163	Tb927.10.14510	hypothetical protein	-	-
BS06615	2.00E-30	Tb927.9.9430	hypothetical protein	-	-
BS06635	3.60E-30	Tb927.9.9430	hypothetical protein	-	-
BS06550	1.80E-40	Tb927.9.9640	hypothetical protein	-	-
BS53150	3.20E-70	Tb927.11.2620	hypothetical protein	PTS1p	[DAATPKNRL]KRL
BS63335	7.60E-16	Tb927.11.2620	hypothetical protein	-	-
BS50905	7.20E-39	Tb927.9.11640	hypothetical protein	-	-
BS18580	1.50E-35	Tb927.9.12980	hypothetical protein	-	-
BS10205	1.30E-10	Tb927.11.5180	hypothetical protein	-	-
BS55185	8.60E-08	Tb927.11.15080	hypothetical protein	-	-
BS24045	3.30E-34	Tb927.9.15260	hypothetical protein	-	-

BS44780	1.90E-100	Tb927.11.12060	hypothetical protein	-	-
BS47680	1.70E-29	Tb927.11.14770	hypothetical protein	-	-
BS93310	0.035	Tb927.11.14770	hypothetical protein	-	-
BS10055	0.00035	Tb927.11.5110	hypothetical protein	-	-
BS10060	3.70E-08	Tb927.11.5110	hypothetical protein	-	-
BS66970	2.00E-41	Tb927.11.10140	hypothetical protein	-	-
BS12405	1.60E-16	Tb927.11.13510	hypothetical protein	-	-
BS48505	1.50E-52	Tb927.11.2580	hypothetical protein	-	-
BS67860	1.60E-30	Tb927.9.6530	hypothetical protein	-	-
BS42170	7.40E-13	Tb927.11.1970	hypothetical protein	-	-
BS22215	3.10E-14	Tb927.9.12500	hypothetical protein	-	-

Appendix 15C: *T. brucei* glycosomal proteins with no *Bodo saltans* homologues

T. brucei ID	*T. brucei* annnotation
Tb927.10.1400	hypoxanthine-guanine phosphoribosyltransferase
Tb927.2.4210	glycosomal phosphoenolpyruvate carboxykinase (PEPCK)
Tb927.4.5730	variant surface glycoprotein (VSG, pseudogene), putative
Tb927.1.3830	glucose-6-phosphate isomerase, glycosomal
Tb927.10.3080	methionine biosynthetic protein, putative (AdoMetMT)
Tb927.3.5090	tryparedoxin, putative
Tb927.8.590	carnitine O-palmitoyltransferase, putative (ACS)
Tb927.7.7500	thymine-7-hydroxylase, putative (TDO)
Tb927.3.4650	C-8 sterol isomerase, putative
Tb927.7.1790	Adenine phosphoribosyltransferase, putative (APRT)
Tb927.5.300	thymine-7-hydroxylase, putative (TDO)
Tb927.6.160	retrotransposon hot spot protein 1 (RHS1), putative
Tb927.8.2310	(H)-ATPase G subunit, putative
Tb927.4.2540	ubiquitin
Tb927.10.9080	pteridine transporter, putative
Tb927.9.6090	TFIIF-stimulated CTD phosphatase
Tb927.11.10260	hypothetical protein
Tb927.11.3100	hypothetical protein
Tb927.11.14170	hypothetical protein
Tb927.9.2320	hypothetical protein
Tb927.11.13610	hypothetical protein
Tb927.1.5000	hypothetical protein
Tb927.4.1600	hypothetical protein
Tb927.6.4150	hypothetical protein
Tb927.5.960	hypothetical protein
Tb927.8.3270	hypothetical protein
Tb927.10.1860	hypothetical protein
Tb927.8.2460	hypothetical protein

Tb927.10.13240 hypothetical protein
Tb927.11.12280 hypothetical protein

Appendix 16: *Bodo saltans* proteins with predicted PTS

Bodo saltans ID	PTS1	PTS2	PTS1p	PTS1p Leader	R. pHMMer top hit (*T. brucei*)
BS01250	SRL	-	-	-	carnitine O-palmitoyltransferase II
BS04195	CHA	-	-	-	-
BS05360	SKA	-	-	-	nucleoside diphosphate kinase
BS05580	SKL	-	-	-	hypothetical protein
BS07630	SKL	-	-	-	orotidine-5-phosphate decarboxylase/orotate phosphoribosyltransferase
BS08155	AHA	-	-	-	-
BS08845	AKL	-	-	-	hypothetical protein
BS09805	AKL	-	-	-	DNA repair protein
BS12060	SHA	-	-	-	-
BS12780	AKA	-	-	-	TFIIF-stimulated CTD phosphatase
BS13080	SRL	-	-	-	Mu-adaptin 1
BS14320	SKL	-	-	-	glucosamine-6-phosphate isomerase
BS15085	AKL	-	-	-	-
BS15625	ARL	-	-	-	-
BS16015	ARL	-	-	-	-
BS16100	ARL	-	-	-	hypothetical protein
BS18035	-	TEVVATRVFHQAAQLILTSS	-	-	-
BS18100	SKA	-	-	-	-
BS21955	AKL	-	-	-	alkyl-dihydroxyacetone phosphate synthase (DHAP)
BS22170	AKL	-	-	-	short-chain dehydrogenase
BS22280	AKL	-	-	-	glycosyl transferase
BS23865	ARL	-	-	-	beta-D-hydroxybutyrate dehydrogenase
BS24115	SKL	-	-	-	-
BS24660	SHL	-	-	-	ribulose-5-phosphate 3-epimerase
BS26750	ARL	-	-	-	leucine-rich repeat protein
BS27275	SKL	-	-	-	ribulose-5-phosphate 3-epimerase
BS27400	SKA	-	-	-	hypothetical protein

BS29225	AHL	-	-	-	guanine nucleotide exchange factor subunit GCD7, translation initiation factor 2b
BS29690	SKL	-	-	-	hypoxanthine-guanine phosphoribosyltransferase
BS29930	SRL	-	-	-	phosphoglycerate kinase
BS30300	AKL	-	-	-	homoserine kinase
BS30360	ARL	-	-	-	small GTPase
BS31135	SHL	-	-	-	-
BS31925	CKA	-	-	-	-
BS32195	SKL	-	-	-	aldehyde dehydrogenase
BS32600	ARA	-	-	-	-
BS32680	SKL	-	-	-	-
BS33550	SRL	-	-	-	ATP-dependent phosphofructokinase (TbPFK)
BS33770	-	MAIVRLRHVSDQAMTSCTDP	-	-	hypothetical protein
BS34135	SKL	-	-	-	AMP deaminase
BS36310	-	MLSKRFSMSRLSTIASQLGI	-	-	3-ketoacyl-CoA thiolase
BS36770	SKL	-	-	-	short-chain dehydrogenase
BS36970	SKL	-	-	-	glycosomal phosphoenolpyruvate carboxykinase (PEPCK)
BS36975	SKL	-	-	-	glycosomal phosphoenolpyruvate carboxykinase (PEPCK)
BS37535	SKA	-	-	-	hypothetical protein
BS37585	AKL	-	-	-	-
BS38775	AKL	-	-	-	glycerol kinase
BS41555	SKL	-	-	-	-
BS41740	ARA	-	-	-	-
BS42465	AKL	-	-	-	hypothetical protein
BS42630	ARL	-	-	-	Dipeptidyl-peptidase 8-like
BS43165	SKL	-	-	-	enoyl-CoA hydratase, mitochondrial precursor
BS44545	SKL	-	-	-	acyl-CoA binding protein
BS46160	SRL	-	-	-	-
BS46210	SRA	-	-	-	-
BS46720	SHL	-	-	-	hypothetical protein

BS46725	SHL	-	-	-	hypothetical protein
BS46730	SKL	-	-	-	hypothetical protein
BS46750	SKA	-	-	-	hypothetical protein
BS46785	AKL	-	-	-	alkyl-dihydroxyacetone phosphate synthase
BS47220	SKL	-	-	-	-
BS48545	AKA	-	-	-	-
BS50560	AHL	-	-	-	inositol/phosphatidylinositol phosphatase
BS51740	AKL	-	-	-	mevalonate kinase
BS52270	SKL	-	-	-	isocitrate dehydrogenase (IDH)
BS53780	AKL	-	-	-	phosphoribosylpyrophosphate synthetase
BS54070	SKL	-	-	-	hypothetical protein
BS57195	SKL	-	-	-	isovaleryl-coA dehydrogenase
BS57535	SKL	-	-	-	pyruvate phosphate dikinase (PPDK)
BS57805	SRL	-	-	-	cystathionine beta-synthase
BS58330	AKA	-	-	-	protein kinase
BS59050	SKL	-	-	-	ATP-dependent phosphofructokinase (TbPFK)
BS60885	AHL	-	-	-	kinesin
BS61510	SKA	-	-	-	acetyltransferase/Elongator-like Protein 3b
BS62875	SHA	-	-	-	-
BS64940	AKL	-	-	-	-
BS65445	SRL	-	-	-	-
BS65465	SKL	-	-	-	glycosomal malate dehydrogenase (gMDH)
BS65575	ARL	-	-	-	-
BS66180	AKL	-	-	-	glyceraldehyde 3-phosphate dehydrogenase (GAPDH)
BS66850	SKL	-	-	-	iron/ascorbate oxidoreductase family protein
BS68360	CRA	-	-	-	-
BS70000	ARA	-	-	-	-
BS70245	SRL	-	-	-	-
BS71060	SKL	-	-	-	ubiquitin carboxyl-terminal hydrolase, cysteine peptidase, Clan CA, family C19, putative

BS72240	SHL	-	-	-	hypothetical protein
BS72250	SKL	-	-	-	hypothetical protein
BS72280	AHL	-	-	-	-
BS72570	-	MKKIVHVAKHLWKGGSLAYL	-	-	hypothetical protein
BS73415	ARL	-	-	-	-
BS73655	CRA	-	-	-	hypothetical protein
BS73660	CRA	-	-	-	hypothetical protein
BS74115	AKL	-	-	-	ATP synthase F1 subunit gamma protein
BS74555	-	MSNDLQPGQKQQQQLQQLAQ	-	-	-
BS76565	SRL	-	-	-	carnitine O-palmitoyltransferase II
BS76730	SKA	-	-	-	-
BS76915	SKL	-	-	-	NADH-dependent fumarate reductase
BS77660	SRL	-	-	-	-
BS78075	SRL	-	-	-	-
BS78645	SKL	-	-	-	-
BS79955	SRA	-	-	-	-
BS80185	SKL	-	-	-	-
BS80490	SKA	-	-	-	-
BS80955	AKL	-	-	-	agmatinase
BS82170	ARL	-	-	-	-
BS82225	AKL	-	-	-	p-nitrophenylphosphatase
BS82785	SRA	-	-	-	-
BS83930	SRL	-	-	-	-
BS84000	SHL	-	-	-	-
BS84045	SKL	-	-	-	hypothetical protein
BS84115	SKL	-	-	-	hypothetical protein
BS84370	SRL	-	-	-	-
BS86915	-	MNMKQKKLSSHLNPLKNKKE	-	-	-
BS89735	CRA	-	-	-	phospholipid:diacylglycerol acyltransferase-like protein
BS90090	SRL	-	-	-	-

BS91330	ARL	-	-	-	-
BS92480	AKA	-	-	-	CCR4 associated factor
BS93250	SKA	-	-	-	protein kinase
BS93880	CRL	-	-	-	-
BS00515	-	-	SSI	FTQGTVTTS	hypothetical protein
BS01370	-	-	SLL	SLHERENSF	paraflagellar rod component
BS02725	-	-	LKI	LGPKGSSKA	hypothetical protein
BS03320	-	-	IRL	RSVSVVVDP	proteasome regulatory non-ATPase subunit 11
BS03760	-	-	LRL	SEPTLFAVD	-
BS04475	-	-	AGI	AELLEERSQ	cytosolic malate dehydrogenase
BS05750	-	-	SSL	QGNPIVVKR	carnitine O-acetyltransferase
BS06080	-	-	SDI	ALSIADEPF	leucine-rich repeat protein (LRRP)/hypothetical protein
BS06895	-	-	AYL	RQATDADHI	-
BS07000	-	-	AMM	KKESGGCPC	surface glycoprotein/hypothetical protein
BS07130	-	-	PRL	KITKKEWEN	-
BS07140	-	-	ALL	YQDDSDLSH	hypothetical protein
BS07690	-	-	IRL	RTTAYATTT	-
BS07950	-	-	ASL	FKLRHVANV	-
BS08210	-	-	SKV	MEYYPPSTR	monooxygenase
BS08520	-	-	FKI	DLDTTSLLS	ATP-dependent RNA helicase
BS08860	-	-	FRL	DSPAGIPAF	hypothetical protein
BS09015	-	-	IRL	ARLQKEEDT	hypothetical protein
BS09030	-	-	SFL	RAPPPVFDD	-
BS09240	-	-	PKL	FGSRRSPLS	chromatin binding protein
BS09460	-	-	SRY	RLTGISPAK	DNA excision repair protein
BS09650	-	-	AQL	KRDERERME	hypothetical protein
BS09895	-	-	ASI	LPPLHSDVN	hypothetical protein
BS10265	-	-	SKV	LVPPPSLLK	hypothetical protein
BS10665	-	-	LRI	ALRKSDPAY	hypothetical protein
BS10915	-	-	CRM	ARRHQEIFR	-

BS11620	-	-	ASL	AHPLQPTSP	hypothetical protein
BS12425	-	-	CKI	VVHEAFEIG	2,4-dienoyl-coa reductase-like protein
BS12555	-	-	SFL	ELQCVADRL	GTPase activating protein
BS13235	-	-	ATL	DEKSATSAT	GTP-binding protein
BS13745	-	-	ADI	DNEREKPFL	phospholipid-transporting ATPase 1-like protein
BS13775	-	-	SLL	IAVTVVAPT	-
BS13940	-	-	IKL	TVDKLTCTS	-
BS14230c	-	-	SPL	VESASDSVT	hypothetical protein
BS14285	-	-	SEL	ELKSTVEDG	-
BS14405	-	-	APM	QLCGLEPTD	hypothetical protein
BS14455	-	-	PKL	EDLIGTRTS	ATP-dependent DEAD/H RNA helicase
BS14725	-	-	SLL	STPKPHFDA	-
BS14915	-	-	ASL	LSRSPRERG	-
BS15160	-	-	TRM	EDDDDGGSY	hypothetical protein
BS15355	-	-	SLL	EVADDDLAT	-
BS16250	-	-	SML	ICMRSSDAR	paraflagellar rod component (PFC2)
BS16400	-	-	SML	YGTDSRNSS	-
BS18210	-	-	SGL	RTYFYSKPP	Bis(5'-adenosyl)-triphosphatase
BS19180	-	-	VKI	PRSAFSKKS	hypothetical protein
BS19220	-	-	GKL	GLGEPTAPK	adenylate kinase
BS19350	-	-	SSI	EGSSDAVAP	hypothetical protein
BS19395	-	-	SLI	SRSVTLPTF	-
BS20095	-	-	KKL	TTSVREVFV	-
BS20950	-	-	FRL	ESEADGPSE	hypothetical protein
BS20955	-	-	FRL	ESEADGPSE	hypothetical protein
BS21285	-	-	LRL	THHTTLLHW	-
BS21305	-	-	SEL	TQHTAAHPN	variant surface glycoprotein (VSG)
BS21550	-	-	AFL	DQKGRRDAV	-
BS22085	-	-	SDM	TKTKTVPSH	-
BS22180	-	-	PRL	NCVEKKFGH	hypothetical protein

BS22365	-	-	SSI	LNDLTPVSG	hypothetical protein
BS22820	-	-	ASL	QQQEAAKKL	hypothetical protein
BS22965	-	-	SEL	HSSRRGGRR	hypothetical protein
BS23480	-	-	SSI	QSRRLPAPP	S. cerevisiae PSP1 homologue/hypothetical protein
BS23725	-	-	SSL	MWKKRKTHS	hypothetical protein
BS25270	-	-	KKI	FARDHRIKK	-
BS25765	-	-	SLM	RKPTSDKCC	ras-like small GTPase
BS26245	-	-	AQL	TTRSRKGSE	-
BS27220	-	-	SPI	KAAISREAD	hypothetical protein
BS27305	-	-	ATL	ATTVRLERA	-
BS27505	-	-	SEL	RAARGVSLN	-
BS27690	-	-	VKL	ENLLPTVEQ	phosphoinositide-binding protein
BS28485	-	-	AKV	SLSTTAADR	multidrug resistance protein
BS29055	-	-	SRV	QATPKVTCL	-
BS29100	-	-	FKI	RCIPVAVVY	hypothetical protein
BS29190	-	-	FKI	RCIPVAVVY	-
BS30175	-	-	VKL	SRVKPEEEW	hypothetical protein
BS30395	-	-	ADI	ASVVNRIAI	hypothetical protein
BS30520	-	-	SKF	KKHTAALQS	-
BS31035	-	-	SLM	PDSRRRNSH	-
BS31265	-	-	PKL	NGCTFPDEP	adenosine kinase
BS31500	-	-	SFL	EELVERRGS	-
BS32975	-	-	SFL	HLSVRRSVS	-
BS33105	-	-	LKL	LHEDEIYSS	-
BS33380	-	-	SDL	IDRDADDAM	hypothetical protein
BS34285	-	-	AEL	GIRRPAGTA	chaperone protein DNAj
BS34350	-	-	GRL	EEAEAEEAI	-
BS34680	-	-	QRL	ATVRNSRRP	-
BS34780	-	-	AEL	DVNVEKRAF	DNA-directed RNA polymerase II
BS34810	-	-	AEL	MKRALDDAT	-

BS34825	-	-	SNL	GIVEIDETS	-
BS35635	-	-	KKL	AEKSILMSM	pyrroline-5-carboxylate reductase
BS35640	-	-	SKV	AEKSLLMAK	pyrroline-5-carboxylate reductase
BS36495	-	-	SLL	LPVRVPVFA	-
BS37575	-	-	ATI	VFFLRPSSP	transporter
BS38020	-	-	AEL	TFELPSEDR	protein disulfide isomerase
BS38415	-	-	KRL	AADKKKAAL	ubiquitin-conjugating enzyme
BS38545	-	-	STL	KMLLDSSTT	hypothetical protein
BS39155	-	-	ASL	RIDVLRRER	-
BS39170	-	-	SYL	RLTKKKKYR	-
BS40230	-	-	SCL	ETDELQALC	-
BS40710	-	-	SFL	TAMDDDEFM	-
BS41120	-	-	SQI	LSAAENAAT	-
BS41190	-	-	GRM	RSSSTVVAV	hypothetical protein
BS41490	-	-	LRM	WVRLITSSR	-
BS41580	-	-	ASL	LESRASASS	protein tyrosine phosphatase
BS43425	-	-	AFI	IVCKGTHTK	hypothetical protein
BS43570	-	-	LKL	QSHKKAQVD	hypothetical protein
BS43615	-	-	KKI	DEEEITSAR	-
BS44440	-	-	STI	LGAAAAASA	hypothetical protein
BS44720	-	-	SLL	PAHHAITTH	Polyadenylate-binding protein 1
BS44880	-	-	FRL	EKKKKPFDG	hypothetical protein
BS44955	-	-	LRM	SGIVCPSST	-
BS45005	-	-	ARY	LRFIKKSKI	phenylalanyl-tRNA synthetase alpha chain
BS45180	-	-	QRL	TTASPSCIL	-
BS45730	-	-	AML	DDDDGVELN	-
BS46545	-	-	APL	DWSVISRER	ABC transporter
BS47715	-	-	AML	EQEADEEEA	-
BS47950	-	-	SSM	GYTAHTFSP	hypothetical protein
BS48180	-	-	ADL	TVRRAPIED	stomatin-like protein

BS48650	-	-	AMI	RAVRMDYGS	-
BS48945	-	-	GKI	DKAVPNYPG	-
BS49325	-	-	SRF	VAGTSSRRA	-
BS50685	-	-	SLM	RPTSAPNSN	-
BS50940	-	-	SSL	KAEETTAAA	Iron-sulfur cluster assembly protein
BS51330	-	-	ATL	EVQGGPDAD	folate transporter
BS51455	-	-	SSL	TSSEFICPS	-
BS52195c	-	-	SRY	VEKEGDTIG	hypothetical protein
BS52550	-	-	ARI	LAALPEATQ	-
BS52810	-	-	SDI	PARKPVHDF	hypothetical protein
BS52990	-	-	KRL	DAATPKNRL	-
BS53150	-	-	SSL	AAKPSSSLQ	hypothetical protein
BS54045	-	-	ATL	KQRKRRRKE	-
BS54580	-	-	SRF	DRDERSNNG	hypothetical protein
BS54925	-	-	GRL	CADNRRREE	hypothetical protein
BS55150	-	-	GKM	DATTAFATS	-
BS55605	-	-	AYI	CLYVHEESP	hypothetical protein
BS55760	-	-	SDL	DIFGSDSDD	-
BS55765	-	-	SDL	DIFGSDSDD	-
BS56040	-	-	SKY	LHNKSKTHT	hypothetical protein
BS56160	-	-	SQL	LFVAVAALW	hypothetical protein
BS56530	-	-	STI	DAVMDDDEP	paraflagellar rod component (PFC2)
BS56550	-	-	ASI	RAYKVRDEL	-
BS57180	-	-	SEL	RAAKALGLL	-
BS57460	-	-	SSL	VSSSLASSS	-
BS57465	-	-	AGL	LREYDDEHL	-
BS57775	-	-	LKL	LTSNSSSEP	hypothetical protein
BS58215	-	-	STM	QKFDDDDLI	hypothetical protein
BS58615	-	-	SSL	YVERVLEEC	hypothetical protein
BS60000	-	-	ALL	SRTHQTTVF	-

BS60540	-	-	SDL	DDSDDRSAD	subtilisin-like serine peptidase, serine peptidase, clan SB, family S8-like protein
BS60570	-	-	KRL	GGTPAATRQ	-
BS60735	-	-	LKM	SAPSFSFSH	hypothetical protein
BS60925	-	-	SGI	EILNSKSPQ	-
BS61010	-	-	GRL	PSSASHNGF	hypothetical protein
BS61450	-	-	SKM	TAMIVRRRQ	-
BS61680	-	-	AEM	SGGTRSSPN	-
BS61700	-	-	AKI	KWGLVAPTV	-
BS61750	-	-	SQL	CTSPYDVSK	-
BS62105	-	-	SLM	EDTPICQAC	kinesin-like protein
BS62140	-	-	AGL	RAAAVDDPL	hypothetical protein
BS62290	-	-	ATL	ERRSKALHS	membrane-bound acid phosphatase 2 (MBAP2)
BS62515	-	-	AQL	GPSTPPIPL	-
BS62550	-	-	SLL	CPCCIHTPS	-
BS63435	-	-	SNL	LQHAEPIAS	-
BS64295	-	-	GRL	VVTKRLLAI	-
BS64690	-	-	ATL	WTGVSKEVL	hypothetical protein
BS64925	-	-	QRI	PSAKLEPFE	-
BS64955	-	-	ASL	KWLRDTATG	-
BS65225	-	-	ASL	LLPKRRSLR	nucleosome assembly protein
BS65290	-	-	TRL	LIGLAERVQ	agmatinase
BS65600	-	-	ATI	IGDSCRDVP	hypothetical protein
BS66040	-	-	AGL	PPPPLNRRR	pre-mRNA splicing factor
BS66045	-	-	AGL	PPPPLNRRR	pre-mRNA splicing factor, ATP-dependent RNA helicase
BS66105c	-	-	LRM	TTSTSHPES	hypothetical protein
BS66190	-	-	SLM	IEDPHRATS	-
BS66715	-	-	SML	GRRPESIDL	-
BS67250	-	-	ALL	AGDVDASVV	-
BS67285	-	-	ALI	PPVLQAAAK	hypothetical protein

BS67825	-	-	AML	SRKSLTTYG	ABC transporter
BS67830	-	-	AML	SRKSLTTYG	-
BS68045	-	-	SKF	VFSLNRVAK	ABC transporter
BS68460	-	-	AMM	APTTSKRLA	-
BS68545	-	-	FRI	SCVHSRITA	paraflagellar rod component (PFC2)
BS68550	-	-	FRI	SCVHSRITA	paraflagellar rod component (PFC2)
BS68880c	-	-	AQL	EARKWSAMG	ATP-dependent DEAD/H RNA helicase
BS68885c	-	-	AQL	EARKWSAMG	ATP-dependent DEAD/H RNA helicase
BS68890c	-	-	AQL	EARKWSAMG	ATP-dependent DEAD/H RNA helicase
BS69005	-	-	SRF	IARLHHEAA	-
BS69190	-	-	SEL	SHEALEVAA	ubiquitin-activating enzyme e1
BS69225	-	-	STI	VRDHVNTRF	-
BS69965	-	-	AGL	LSGDRTAAK	lysophospholipase, alpha/beta hydrolase
BS70080	-	-	SCL	RCSPCARTA	paraflagellar rod component (PFC2)
BS70085	-	-	KKI	SSGSERAKP	-
BS70335	-	-	AEI	RRRREAPVL	-
BS70570	-	-	SPL	GIPLCVHAP	-
BS70715	-	-	ASL	AAATLLLTK	3,2-trans-enoyl-CoA isomerase
BS71430	-	-	SLL	GGGVQRSDV	protein tyrosine phosphatase
BS71475	-	-	SRV	EEWQPPARQ	65 kDa invariant surface glycoprotein
BS71850	-	-	AEL	LDHEAEDAH	-
BS71900	-	-	SSL	VKTEELSGT	-
BS72200	-	-	ADL	VAATIHHQD	-
BS72725	-	-	GRL	ELVVSVVSG	-
BS72810	-	-	LKM	AEATQRKKK	ubiquitin carboxyl-terminal hydrolase, cysteine peptidase, Clan CA, family C19, putative
BS72910	-	-	SLL	YTPTAPMSV	-
BS72915	-	-	AEM	IDEVFDHRD	-
BS73080	-	-	APM	ATDVEPPRE	-
BS73385	-	-	SMM	RTAESICPC	-

BS73690	-	-	SSI	DDSPLMKLR	-
BS74410	-	-	SGI	FVRDRCAST	kinesin K39
BS74440	-	-	LKL	KVVEIVKER	eukaryotic translation initation factor 4 gamma
BS75660	-	-	GRL	ARKPNYQAC	-
BS75700	-	-	SYL	SPTSFSRSE	phosphatidic acid phosphatase
BS75775	-	-	AYI	ALRHRTTFD	-
BS75990	-	-	SDL	RREAFLDSA	-
BS76145	-	-	SEL	ENMMLKRSS	-
BS76495	-	-	LRI	DGAAKVRRR	-
BS76655	-	-	SRI	VQRRSTVLR	-
BS76885	-	-	STL	KVVLTTTPT	oxidoreductase
BS77065	-	-	GRL	DTSWKRKPN	-
BS77120	-	-	TRI	LRFTRAESE	-
BS77250	-	-	ASL	DVNINDPKK	kinesin
BS77895	-	-	VRL	HNDNGSAAV	arginine N-methyltransferase
BS78310	-	-	AEL	TIDDGKEDD	hypothetical protein
BS78940	-	-	VKI	AKFPKRVFA	60S ribosomal protein L35
BS79270	-	-	AMM	EKEDTSCSC	protein kinase
BS81520	-	-	KKI	RKTVSEFRG	elongation factor
BS81705	-	-	SMM	GQPGNRSHS	hypothetical protein
BS82440	-	-	ARM	ETLANSAEV	cysteine peptidase precursor, cysteine peptidase, Clan CA, family C1, Cathepsin L-like
BS83170	-	-	SKY	SPRRVERFC	-
BS83245	-	-	SQL	CCKFDTSTQ	-
BS83555	-	-	GKI	RTESQPHVK	-
BS83970	-	-	ARF	RADRQVVKG	ATP-dependent DEAD/H RNA helicase
BS84215	-	-	LRL	DLLPPASAH	-
BS84355	-	-	SGL	FDDSDCRFF	-
BS84850c	-	-	GKI	PKAPTAGAG	-
BS84865	-	-	PRL	LAMRQAVDV	-

BS85985	-	-	SPL	EIVANTTSA	-
BS86495	-	-	TKL	EKFLESVDS	glycosylphosphatidylinositol-specific phospholipase C
BS86585	-	-	ADL	LRADASDGE	-
BS86905	-	-	TKI	SKEEQNAGS	-
BS87675	-	-	AKM	QKEAAAAKI	-
BS87865	-	-	SQL	QHKDKRPSV	hypothetical protein
BS88165	-	-	ATL	SRSKKSTKG	-
BS88820	-	-	SKI	HGGRSAHYA	paraflagellar rod component (PFC2)
BS89460	-	-	SKY	DEQDKATAF	TFIIF-stimulated CTD phosphatase
BS89575	-	-	STI	STSKSFSDS	hypothetical protein
BS89710	-	-	ASL	EDDDYESVK	hypothetical protein
BS89725	-	-	QRL	DDATRNYML	hypothetical protein
BS90430	-	-	SPI	AKKSKGGSA	protein kinase
BS90540	-	-	SGL	TKHNARSTV	-
BS90780	-	-	ASI	MTISVSHLA	-
BS90855	-	-	QRI	LDHPTQPRT	-
BS91760	-	-	ADL	VETQSASSR	-
BS93245	-	-	PRL	DLAEVDVLV	hypothetical protein
BS93575	-	-	GKI	CKTVKGWTG	-
BS93780	-	-	ASM	AKARIFNAV	hypothetical protein
BS93875	-	-	SDI	SSMRCEVSV	-
BS94135	-	-	SLL	AITDAKRCS	-
BS94570	-	-	QRI	MFMMRATTI	protein kinase

Appendix 17A: *Naegleria fowleri PEX* genes

Protein	Accession	Length	R. pHMMer (*H. sapiens*)	R. pHMMer (*S. cerevisiae*)	R. pHMMer (*N. crassa*)	R. pHMMer (*A. thaliana*)	R. pHMMer (*N. gruberi*)
Pex1	augustus_g6024	1133	2.50E-114	-	4.50E-112	2.50E-115	-
Pex2	augustus_g8328	481	3.80E-31	-	8.10E-26	4.60E-33	-
Pex3	augustus_g3435	404	2.30E-25	-	1.10E-13	9.30E-13	-
Pex4	augustus_g4750	165	-	-	-	4.10E-59	-
Pex5	augustus_g3785	376	-	3.10E-06	4.00E-10	7.50E-24	-
Pex6	augustus_g2952	718	-	3.80E-10	-	1.00E-15	-
Pex7	augustus_g9074	397	8.90E-81	4.90E-65	7.80E-57	4.10E-99	-
Pex10	augustus_g1001	293	3.60E-16	8.50E-13	1.00E-20	1.10E-25	-
Pex11	augustus_g3959	257	8.70E-03	6.50E-03	4.40E-02	1.10E-02	-
	augustus_g1404	351	-	2.10E-04	4.50E-04	1.60E-04	-
Pex12	augustus_g8470	396	9.00E-43	7.90E-22	2.60E-23	2.60E-47	-
Pex16	augustus_g11467	680	4.80E-03	-	3.60E-04	1.00E-07	-
Pex19	augustus_g440	297	1.30E-20	2.90E-09	3.00E-16	1.30E-17	-

Appendix 17B: *Naegleria fowleri* glycolysis enzymes

Protein	*N. fowleri* ID	R. pHMMer (*N. gruberi*)	PTS
fructose-bisphosphate aldolase (ALD)	augustus_g9439	1.80E-203	-
glucose-6-phosphate isomerase (GPI)	augustus_g8520	5.60E-293	-
glyceraldehyde 3-phosphate dehydrogenase (GAPDH)	augustus_g9438	1.30E-209	-
glycerol kinase (GK)	augustus_g10225	6.30E-294	-
glycerol-3-phosphate dehydrogenase [NAD+] (G3PDH)	augustus_g4032	6.70E-272	-
	augustus_g404	7.40E-182	-
hexokinase (HK)	augustus_g10647	9.80E-239	-
phosphofructokinase (PFK)	augustus_g2742	8.50E-244	-
phosphoglycerate kinase (PGK)	augustus_g10212	4.10E-237	-
triosephosphate isomerase (TPI)	augustus_g158	3.40E-138	-