Comparative genomics of peroxisome biogenesis proteins: making sense of the PEX mess

Renate L.M. Jansen[1,*], Carlos Santana Molina[2,*], Marco van den Noort[1,3], Damien P. Devos[2,4] and Ida J. van der Klei[1,4]

[1]Molecular Cell Biology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9800CC, the Netherlands

[2]Centro Andaluz de Biología del Desarrollo, CSIC, Universidad Pablo de Olavide, Carretera de Utrera, Km.1, Seville 41013, Spain

[3]Current address: Membrane Enzymology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9800CC, the Netherlands

[4]Corresponding authors: i.j.van.der.klei@rug.nl, ORCID ID 0000-0001-7165-9679, P.O. Box 11103, 9700CC, Groningen, The Netherlands and damienpdevos@gmail.com, ORCID ID 0000-0002-9453-4753, Carretera de Utrera, Km.1, Seville 41013, Spain

*These authors equally contributed.

Abstract

PEX genes encode proteins involved in peroxisome biogenesis and proliferation. Using a comparative genomics approach, we clarify the evolutionary relationships between the 37 known PEX proteins in a representative set of eukaryotes, including all common model organisms, pathogenic unicellular eukaryotes and human. A large number of previously unknown PEX orthologs were identified. We analysed all PEX proteins, their conservation and domain architecture and defined the minimum set of PEX proteins that is required to make a peroxisome. The molecular processes in peroxisome biogenesis in different organisms were put into context, showing that peroxisomes are not static organelles in eukaryotic evolution. Organisms that lack peroxisomes still contain a few PEX proteins, which probably play a role in alternative processes. Finally, the relationships between PEX proteins of two large families, the Pex11 and Pex23 families, were clarified, thereby contributing to the understanding of their complicated and sometimes incorrect nomenclature. We provide an exhaustive overview of this important eukaryotic organelle.

Keywords

Comparative genomics/peroxisome/PEX/evolution/protein domains

Introduction

Peroxisomes occur in almost all eukaryotes. Their number, size and protein composition are highly variable. In lower eukaryotes, such as yeast, peroxisome proliferation is stimulated by specific growth substrates. In higher eukaryotes, peroxisome abundance and composition vary with organism, tissue and developmental stage. Conserved peroxisomal pathways are the β-oxidation of fatty acids and hydrogen peroxide degradation. Examples of specialized pathways are the biosynthesis of bile acids and ether lipids in man, photorespiration in plants and the biosynthesis of antibiotics in certain filamentous fungi (Smith & Aitchison, 2013). The crucial role of peroxisomes for human health is illustrated by the occurrence of inborn errors that cause severe diseases and are often lethal. However, roles in non-metabolic processes such as ageing, anti-viral defence and cancer show that the significance of peroxisomes in human health goes far beyond the relatively rare inherited peroxisomal disorders (Islinger et al, 2018).

Peroxisomes are very simple organelles that consist of a protein rich matrix surrounded by a single membrane. Peroxisomal enzymes almost exclusively occur in the matrix. The membrane contains transporters, pores for solute transport and proteins involved in diverse processes such as matrix and membrane protein sorting, organelle fission and movement (*figure 1*).

In 1996, the term peroxin was coined for proteins “involved in peroxisome biogenesis (inclusive of peroxisomal matrix protein import, membrane biogenesis, peroxisome proliferation, and peroxisome inheritance)” (Distel et al, 1996). Peroxins are encoded by PEX genes and also called PEX proteins. So far, 37 PEX proteins have been described. Some are highly conserved, whereas others only occur in a limited number of species. Since 1996, tremendous progress has been made in our understanding of the molecular mechanisms involved in peroxisome biology. However, with the increasing number of PEX proteins, their nomenclature became more and more complex (Smith & Aitchison, 2013).

Here, we present an exhaustive up-to-date overview of all the PEX protein families. We analysed PEX proteins in a highly diverse set of eukaryotes, including all common model organisms, pathogenic unicellular eukaryotes and higher eukaryotes. Using this information, we combine phylogenetic reconstructions with other protein features (e.g., Pfam domain, protein disorder and transmembrane domain predictions) to understand the evolution of these proteins, clarifying certain inconsistencies in the nomenclature of PEX proteins. Important questions that we answer are (i) how are the different PEX genes conserved across eukaryotes, (ii) what is the minimum set of PEX genes to make a canonical peroxisome and (iii) what are the typical features of the PEX proteins.

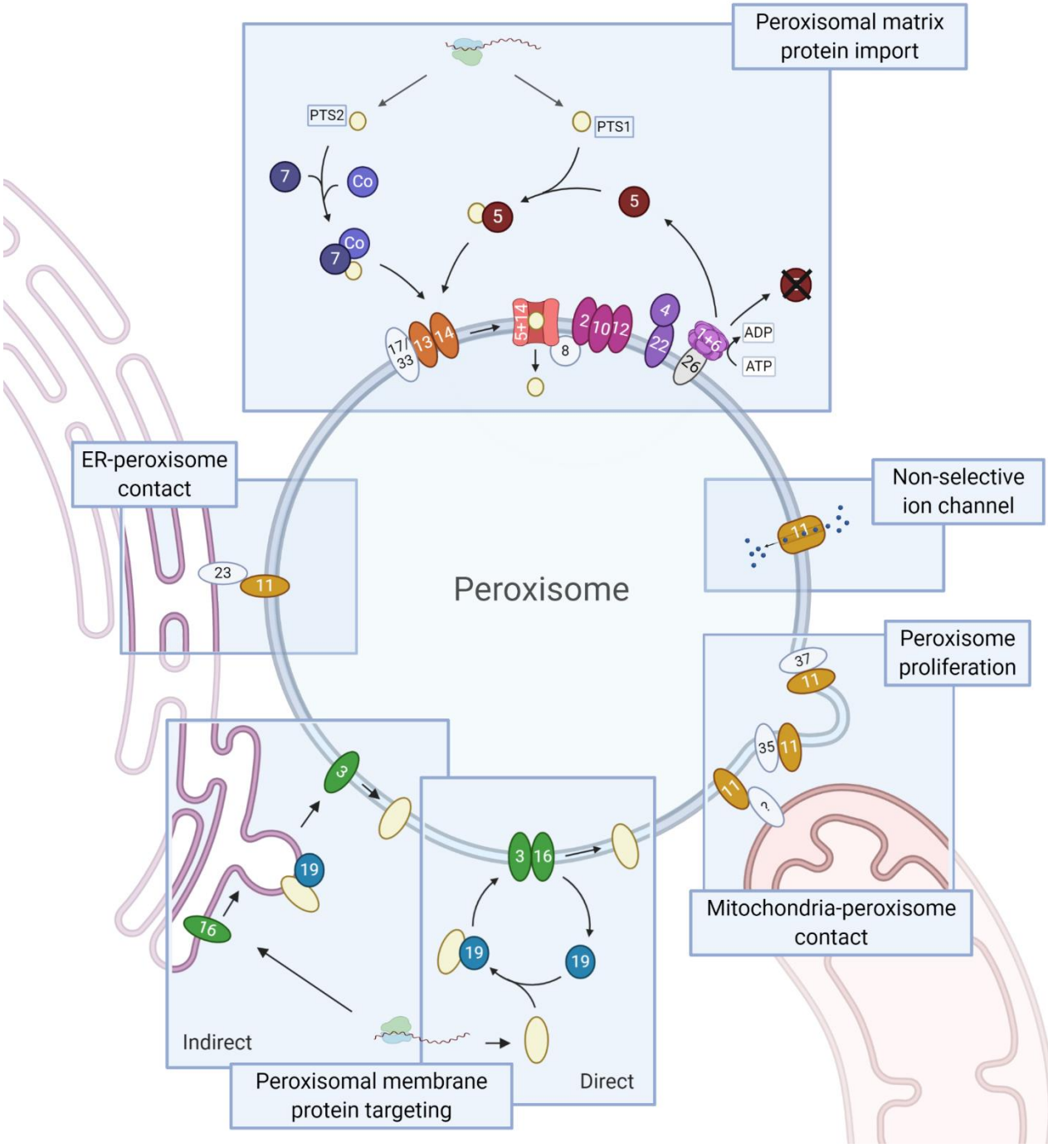

Figure 1. Schematic representation of the PEX proteins. Core conserved PEX proteins (shapes in dark colours, names in white), fungi-specific proteins (light, names in black) and the moderately conserved PEX protein Pex26/15 (grey, name in black, which is only present in metazoa and fungi) are depicted. Membrane proteins are ovals, soluble proteins round. **Matrix protein import.** Peroxisomal matrix proteins contain a peroxisomal targeting signal (PTS) that is recognized by cytosolic receptors: a C-terminal PTS1 or (less commonly) an N-terminal PTS2, recognized by PEX5 and PEX7 respectively. PTS2 import involves a co-receptor (Co): PEX5 (animals, plants and protists), PEX18/21 (*S. cerevisiae*) or PEX20 (fungi). Next, the receptor-cargo complex associates with the docking complex, consisting of PEX13/14 (and in fungi PEX17 or PEX33). Upon cargo translocation and release, the PTS (co-)receptor is ubiquitinated and recycled. Ubiquitination involves the ubiquitin conjugating enzyme (E2) PEX4 (recruited to the membrane by PEX22) and the ubiquitin ligase (E3) activities of the RING finger complex, consisting of PEX2/10/12. Receptor extraction requires the AAA+ ATPase complex PEX1/6, which is recruited to the membrane via PEX26 (PEX15 in *S. cerevisiae*). PEX8 bridges the docking and RING finger complexes, and functions in receptor-cargo dissociation. **Peroxisomal**

membrane protein (PMP) sorting involves PEX3, PEX19 and PEX16. PMPs can sort directly to peroxisomes or indirectly via the ER. In the direct pathway PEX19 acts as receptor/chaperone, while it functions at the ER in PMP sorting via the indirect pathway. The **Pex11 protein family** (all show as PEX11) and the fungal peroxins PEX35 and PEX37 have been mainly implicated in peroxisome proliferation. Pex11 family proteins are also present in mitochondria-peroxisome contact sites and PEX11 functions as non-selective ion channel. Members of the fungal **Pex23 protein family** localize to the ER and are involved in the formation of peroxisome-ER membrane contact sites. Created with BioRender.com.

Results

The proteomes of 38 eukaryotes were investigated to identify all PEX proteins known to date. Not all eukaryotes contain peroxisomes (Žárský & Tachezy, 2015) and several protist species of our initial analysis were found to lack most PEX proteins, namely *Cryptosporidium parvum, Theileria annulata, Babesia bovis, Monosiga brevicollis, Plasmodium falciparum, Blastocystis hominis* and *Entamoeba histolytica*. To facilitate comparison between species containing and (likely) lacking peroxisomes, the latter species was included in further analyses, but all others likely lacking peroxisomes were omitted. *Table 1* shows the 31 remaining species containing peroxisomes, plus *Entamoeba histolytica*. An overview of all orthologs identified can be found in *table S1*.

Table 1: Overview of proteomes investigated.

Eukaryotic supergroup	Kingdom	Other labels	Species	Uniprot	Description	Note
Opisthokonta	Metazoa (animals)	Vertebrate, Mammalian	*Homo sapiens*	HUMAN	Human	
		Vertebrate, Mammalian	*Mus musculus*	MOUSE	House mouse	
		Vertebrate	*Danio rerio*	DANRE	Zebrafish	
			Drosophila melanogaster	DROME	Fruit fly	
			Caenorhabditis elegans	CAEEL	Nematode	
	Fungi	Yeasts*	*Saccharomyces cere-visiae*	YEAST	Baker's yeast	
			Komagataella phaffii[a]	KOMPG	Yeast	Methylotropic
			Candida albicans	CANAL	Opportunistic pathogenic yeast/fungus	Causes candidiasis
			Ogataea polymorpha[b]	PICAN	Yeast	Methylotropic
			Penicillium rubens	PENRW	Filamentous fungus	Produces penicillin
			Aspergillus niger	ASPNC	Filamentous fungus	Causes black mold
			Colletotrichum higginsi-anum	COLHI	Plant pathogen	
			Gibberella fujikuroi	GIBF5	Plant pathogen	Causes bakanae dis-ease in rice
			Neurospora crassa	NEUCR	Red bread	

					mold	
			Schizosaccharomyces pombe	SCHPO	Fission yeast	Smallest known genome sequence for eukaryote
			Schizophyllum commune	SCHCM	Split gill (mushroom)	Edible
			Cryptococcus neoformans	CRYNJ	Filamentous, encapsulated, pathogenic yeast and obligate aerobe	Can live in plants/animals. Causes cryptococcosis
Amoebozoa	Amoebozoa	Protist	*Dictyostelium discoideum*	DICDI	Amoeba	Slime mold
			Entamoeba histolytica	ENTHI	Parasitic anaerobic amoeba	Causes amoebiasis, lacks peroxisomes
Archaeplastida	Viridiplantae	Plant	*Arabidopsis thaliana*	ARATH	Mouse-ear cress (plant)	Relatively small genome, diploid
			Physcomitrella patens	PHYPA	Moss	Highly efficient homologous recombination (good for creating knock-outs)
		Protist	*Ostreococcus tauri*	OSTTA	Green algae	Compact genome
	Rhodophyta		*Galdieria sulphuraria*	GALSU	Red algae	Horizontal gene transfer from archaea and bacteria (5% of genome). Extremophile
SAR[c]	Alveolata		*Toxoplasma gondii*	TOXGV	Parasite, protozoa	Causes toxoplasmosis
			Tetrahymena thermophila	TETTS	Ciliate	Nuclear dimorphism
	Stramenopiles		*Phytophthora infestans*	PHYIT	Water mold	Causes potato blight
			Thalassiosira pseudonana	THAPS	Marine diatom	Relatively small genome
Excavata	Heterolobosea		*Naegleria gruberi*	NAEGR	Amoeboflagellate	Can change from amoeba to flagellate
	Euglenozoa		*Euglena gracilis*	EGRACILIS[d]	Single-celled alga	Has chloroplasts
			Trypanosoma brucei	TRYB2	Parasitic kinetoplastid	Causes sleeping sickness in animals. kinetoplast (organelle)
			Leishmania major	LEIMA	Parasite, trypanosomatid	Causes zoonotic cutaneous leishmaniasis
			Bodo saltans	BODSA	Non-parasitic kinetoplastid protozoa	

*) These four species belong to the order of Saccharomycetales and are sometimes referred to as “true yeasts”. They are hereafter referred to in the text as “yeasts”.

a) Previously named *Pichia pastoris*

b) Previously named *Hansenula polymorpha*

c) SAR = Stramenophiles, Alveolates and Rhizaria
d) Proteome not from Uniprot, but from recently identified proteome described by Ebenezer et al. (2019)(Ebenezer et al, 2019).

1. Distribution and general description of PEX proteins across eukaryotic lineages

The results of our computational survey are summarized in *figure 2*. We detect a core of PEX proteins that are **broadly conserved** in all eukaryotic lineages, encompassing PEX3/19/16 (peroxisomal membrane protein (PMP) sorting), PEX1/6, PEX2/10/12, PEX13/14 and PEX5/7 (matrix protein import) and proteins of the Pex11 family (peroxisome proliferation and contact sites). Some detected **absences** are probably real. On the other hand, in other cases the function of missing PEX proteins may be taken over by other homologous proteins. For instance, the function of the ubiquitin conjugating enzyme (E2 enzyme) PEX4 in receptor ubiquitination is performed by proteins of the E2D family in Metazoa, which lack a PEX4 ortholog (Grou et al, 2008). Similarly, the function of PEX26 is complemented by the homologous protein APEM9 in plants (Cross et al, 2016) and PEX15 in *S. cerevisiae*. Furthermore, we observe an important bias towards fungi (yeasts and filamentous fungi) reflected in the large number of PEX proteins that are specific to fungi, such as PEX8, PEX20/18/21 and the Pex23 family (*figure 2*). This is a result of the fact that the large majority of studies investigating peroxisomes, in particular their biogenesis, have been performed in yeast models such as *S. cerevisiae, O. polymorpha, K. phaffii* and *Yarrowia lipolytica*.

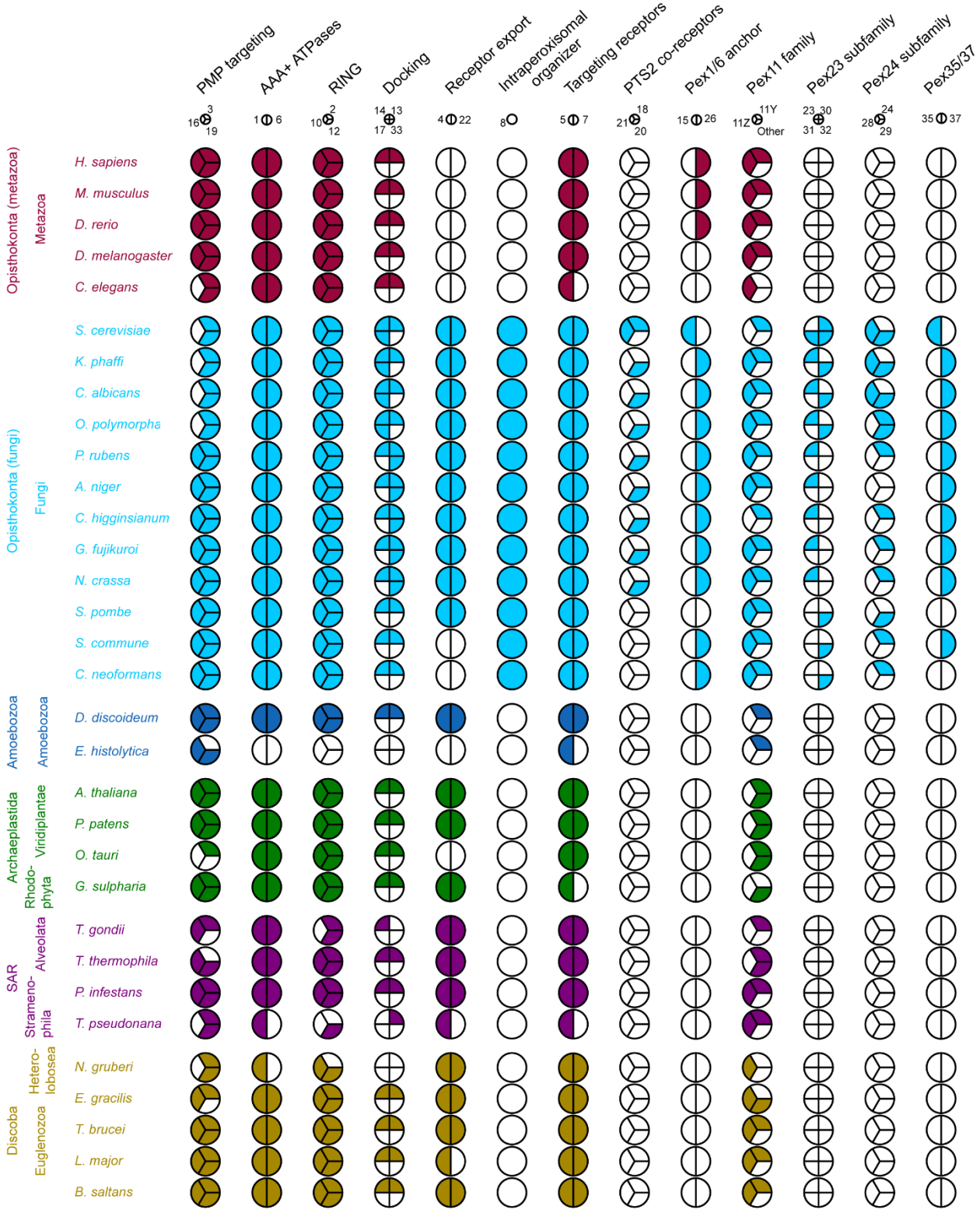

Figure 2: Coulson plot demonstrating the presence (filled) or absence (empty) of PEX protein orthologs in 32 eukaryotic proteomes. PEX proteins are divided into functional groups (columns) including homologous and non-homologous proteins, represented by a pie. Every wedge represents a PEX protein, with the exception of the Pex11 family, where each wedge represents a subfamily. The PEX11Y subfamily contains among others fungal PEX11/25/27/34/36 and mammalian PEX11α/β. The PEXZ subfamily contains fungal PEX11C and mammalian PEX11γ. Pex11 family proteins that do not belong to the PEX11Y or PEX11Z subfamilies are placed in “Other”. Organisms are grouped by eukaryotic supergroup (colour-coded for clarity) and kingdom. PEX proteins are designated by their number.

We analysed the structural features of the PEX proteins (see *table 2*). Structural protein disorder seems to be a common feature among some PEX proteins. In some of them, structural disorder is only predicted for a short fragment, but others like PEX19, PEX18/20/21, PEX14/33-13 are predicted as almost entirely disordered. Also, transmembrane helical domains are usually present in certain PEX proteins, such as PEX3, PEX14 and PEX26. Several PEX proteins have common eukaryotic structural domains, like the E2 enzyme PEX4 and the AAA+ ATPase domain present in PEX1 and PEX6. We also detect several functional domain associations, such as the RING finger (zinc finger) domain in PEX2/10/12 and SH3 domains in PEX13, both being involved in signal transduction and controlling protein-protein interactions. Other recognizable fold types in PEX proteins include α-solenoid formed by the TPR repeat domains in PEX5 and the β-propeller formed by WD40 repeats in PEX7.

Table 2: Overview of PEX proteins and their main features.

Main members	Actual groups	Naming inconsistency	Protein disorder	Trans-membrane	Pfam	Distribution
PEX1, 6	PEX1, 6		Some regions	-	AAA, AAA_lid_3, PEX-1N	Eukaryotes
PEX2, 10, 12	PEX2, 10, 12		Some regions	+	Pex2_Pex12, Zn-finger	Eukaryotes
PEX3	PEX3		Some regions	+	peroxin-3	Eukaryotes
PEX4	PEX4		-	-	UB_con	Eukaryotes
PEX5, 9	PEX5	PEX9 is actually a duplication of PEX5, specific to *S. cerevisiae*	+, N-terminal	-	TPR	Eukaryotes
PEX7	PEX7		-	-	WD40 (B-propeller)	Eukaryotes
PEX8	PEX8		-	-	unknown	Fungi
Pex11 family, including PEX25, 27, 34, 36	Pex11Y subfamily (incl. group PEX25/27/34/36), 11Z subfamily, Other	Current PEX11A/B names in different organisms are inconsis-tent. PEX25, 27, 34 and 36 belong to the same fungi-specific paralog group	-	+	PEX11	Pex11Y-Z Eukaryotes / PEX25/27/34/36 Fungi
PEX13	PEX13		+ (N-terminal)	+	Pex13, SH3	Eukaryotes
PEX14, 33	PEX14, 33		+	+	Pex14_N	PEX14-Eukaryotes / PEX33-fungi
PEX15, 26	PEX15, 26		Some regions	+	Pex26	Fungi-metazoa
PEX16	PEX16		Some regions	+	Pex16	Eukaryotes
PEX17	PEX17		-	+	unknown	Fungi
PEX19	PEX19		+	+?	Pex19	Eukaryotes
PEX18, 20, 21	PEX20	PEX18 and PEX21 of *S. cerevi-siae* are the result of the duplica-tion of the fungi specific PEX20 form	+	-	unknown	Fungi
PEX22	PEX22		-	+	peroxin_22	Fungi-Plants-Protist
PEX23, 24, 28, 29, 30, 31, 32, 23-like, TECPR1	TCPR1, Pex23 subfamily, Pex24 subfamily, 23-like (sporula-tion)	PEX28 and PEX24 are actually the same protein in different organisms (*O. polymorpha* and *S. cerevisiae*). PEX30 and PEX31 are a specific duplication in *S. cerevisiae* of PEX23 form	Some regions	+ (TCPR1 does not)	pex24p, Hyd_WA (TCPR1)	Fungi / TECPR1 metazoa
PEX35	PEX35		-	+	unknown	*Saccharomycetaceae*
PEX37	PEX37		Some, C-terminal	+	Mpv17_PMP22	Fungi

The actual groups indicate to the main groups (deep paralogs) identified in phylogenetic reconstructions. Protein disorder was predicted using IUPRED and transmembrane helices through TMHMM software. Functional protein domains

annotated using Pfam database. Question marks indicate those features that were present in a subset of protein sequences from our data set.

The functional diversification of proteins is caused by the duplication of the respective genes. This is one of the main sources of cellular complexity and development. This process is called paralogization, where paralogous proteins are those having a common origin, i.e., belonging to the same protein family. These gene duplications (paralogizations) can be ancestral (deep paralogs) or they can be asynchronous during evolution: appearing later and being restricted to specific taxonomic clades (in-paralogs). The paralogization of PEX proteins seems to have been relevant for the development of peroxisomes in Eukarya domain. Indeed, some of these paralogizations preceded the diversification of eukaryotes, like the peroxins of the AAA+ ATPase protein family PEX1/6, the RING finger proteins PEX2/10/12 and proteins of the Pex11 protein family. On the other hand, some other PEX proteins have been duplicated in specific eukaryotic taxons. These proteins have often been inconsistently named, since newly discovered proteins were sometimes given a new number. This should be kept in mind when studying such proteins. For instance, the *S. cerevisiae* PEX9 is actually a copy of PEX5 (in-paralogs, not ancestral duplication in fungi). Similarly, the fungal PTS2 co-receptors PEX18/20/21 should be considered as a single group: PEX18 and PEX21 of *S. cerevisiae* are actually the result of a duplication of the ancestral PEX20 form. The PEX23 family is specifically found in fungi and encompasses multiple copies in specific organisms, such as PEX30/31/32 and PEX28/29 in *S. cerevisiae*, resulting from the duplication of PEX23 and PEX24, respectively. In the previous examples, different proteins derived from the same ancestral protein, i.e., belonging to the same protein family, have received different numbers. On the other hand, the opposite has happened for certain other PEX proteins. Many members of the Pex11 family have the same number, but were given a different appendix instead: for instance, PEX11α/β/γ or PEX11A/B/C/D/E. We detected that these paralogs originated from independent paralogizations in different lineages, but their naming does not always reflect this. For instance, fungal PEX11C belongs to the same subfamily as human PEX11γ, but PEX11C from *A. thaliana* does not. Similarly, *A. thaliana* PEX11A is not equivalent to human PEX11α. Based on phylogenetic reconstructions, we propose that two different subfamilies can be distinguished within the Pex11 family. In addition, the Pex11 family includes an in-paralog group specific to fungi, containing PEX25/27/34/36.

Therefore, in some cases, the nomenclature ascribed to the PEX protein paralogizations could lead to confusion, because there is no uniformity in the way in which paralogous, in-paralogous or non-paralogous/unrelated proteins have been named. Furthermore, some paralogizations have led to

paralogs of PEX proteins that may no longer function in peroxisome biology. For instance, vertebrates express a PEX5 paralog called PEX5R (TRIP8b), whose only known function is the regulation of hyperpolarization-activated cyclic nucleotide-gated (HCN) channels - key modulators of neuronal activity(Han et al, 2020).

Taking into account all of the above, we review the role of these PEX proteins below, in order to gain a comprehensive understanding of their functional classification.

2. A core set of PEX proteins is broadly conserved in Eukaryotes

A core set of PEX proteins is broadly conserved across all eukaryotic lineages, encompassing proteins involved in PMP sorting (PEX3, PEX19 and PEX16), matrix protein receptors (PEX5 and PEX7), components of the receptor docking site (PEX13 and PEX14), enzymes involved in receptor ubiquitinylation (PEX2, PEX10, PEX12 and PEX4), two AAA-ATPases that play a role in receptor recycling (PEX1 and PEX6) and a protein family involved in peroxisome proliferation (Pex11 family). The function of these conserved PEX proteins is central to peroxisome biology and thus maintained. In the following section, we will review how these processes define the biology of the canonical peroxisomes as well as the mechanistic models proposed in the field. Furthermore, we describe variations in the repertoire of PEX proteins in certain eukaryotes.

Sorting of PMPs (PEX3, PEX19 and PEX16)

Only three PEX proteins (PEX3, PEX16 and PEX19) are known to be involved in targeting of PMPs. Two mechanisms of PMP sorting to the peroxisome membrane have been described (see *figure 1*; for a detailed review, see (Jansen & van der Klei, 2019)). According to the direct sorting model, PEX19 binds to newly translated PMPs in the cytosol. In this pathway PEX19 acts as a chaperone and cycling receptor (Jansen & van der Klei, 2019). The PEX19-PMP complex binds to the PMP PEX3 and is subsequently inserted in the membrane by a currently unknown mechanism. In the indirect pathway, PMPs traffic first to the ER and accumulate at a subdomain, where PMP containing vesicles bud off. PEX3 plays a role in the intra-ER sorting of PMPs (Fakieh et al, 2013), while PEX19 is important for vesicle budding (Agrawal et al, 2016, Van Der Zand et al, 2012). PEX16 plays a role in the indirect pathway (Hua & Kim, 2016). Notably, PEX3 is also involved in a host of other functions, including pexophagy, peroxisome retention during yeast budding and the formation of contacts between peroxisomes and vacuoles. In all these processes, PEX3 recruits proteins to the peroxisomal membrane (e.g. Atg30/36, Inp1) (Jansen & van der Klei, 2019).

Our computational survey shows that PEX3, PEX19 and PEX16 are conserved well, with a few exceptions, suggesting minor variations in mechanisms of PMP sorting. For instance, PEX16 is widely conserved, but is absent in all (investigated) yeast species, *C. elegans* and several protists. A characteristic motif in PEX19 orthologs of many species is a **CaaX** box at the C-terminus. Farnesylation of this motif causes conformational changes in PEX19 and increases its binding affinity for PMPs (Emmanouilidis et al, 2017, Rucktaschel et al, 2009). Previous studies in *S. cerevisiae* and humans are contradictory regarding the importance of this post-translational modification for peroxisome function (Rucktaschel et al, 2009, Schrul & Kopito, 2016, Vastiau et al, 2006). Interestingly, Schrul & Kopito (2016) found that the CaaX box of human PEX19 was important for targeting of lipid droplet protein UBXD8, but not for peroxisome biogenesis (Schrul & Kopito, 2016). We checked if the CaaX box is present in all eukaryotes. We found that while this motif is present in all animals, plants and fungi, it is absent (or difficult to align) in many protists, like euglenozoa and amoebozoa, despite these organisms expressing the enzyme required for farnesylation (see e.g. (Buckner et al, 2002)) (*figure 3*). Interestingly, putative PEX19 orthologs were also identified in *Entamoeba histolytica* and *M. brevicollis*, despite these species very likely lacking peroxisomes. This may suggest an alternative function for PEX19, unrelated to peroxisomes.

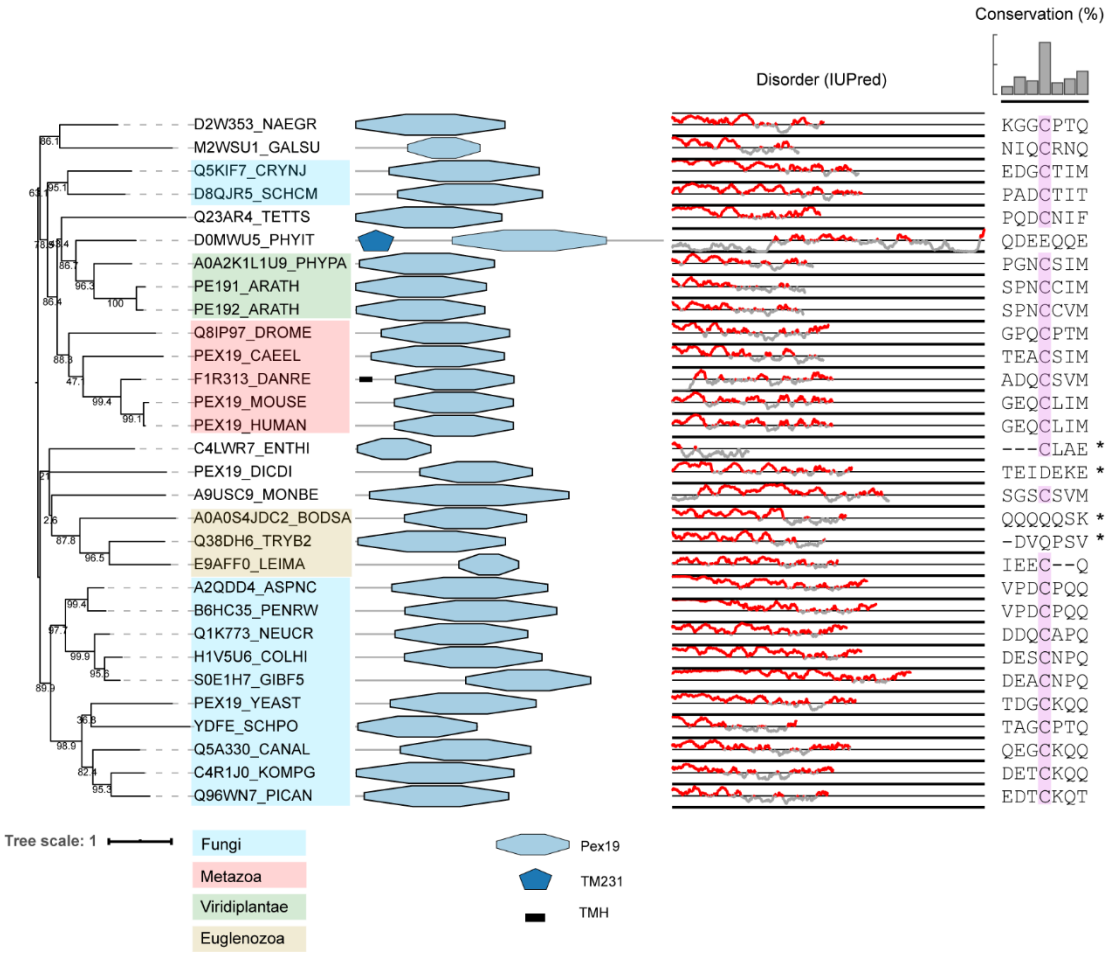

Figure 3. Phylogeny and protein features of PEX19 orthologs. The phylogeny is rooted at mid-point to ease the visualization and labels of the main taxonomic groups are coloured according to the legend. Note that the topology does not necessarily reflect the actual evolutionary trajectory of such proteins. Protein domain architecture is defined by pfam annotations and transmembrane helices (TMH) according to TMHMM software. The line-dot plot, indicates the regions predicted to be disordered (red) and not disordered (grey). The sequence alignment shows the conservation of the CaaX box in PEX19 orthologs of distant eukaryotes, with 'C' denoting Cys, 'a' an aliphatic residue and 'X' usually being a Ser, Thr, Gln, Ala or Met. Asterisk indicates forced alignments manually.

Matrix protein receptors (PEX5 and PEX7)

Newly synthesized matrix proteins are first recognized by their cytosolic **peroxisomal targeting signal (PTS) receptor**. The majority of peroxisomal matrix proteins contain a PTS1 or a PTS2, recognized by PEX5 and PEX7 respectively.

PEX7 contains WD40 repeats, which fold into a β-propellor structure that provides a platform for interaction with the PTS2 motif and PTS2 co-receptor (Pan et al, 2013). While PEX5 was identified in all eukaryotic organisms, PEX7 is absent in *C. elegans*, *T. pseudonana* and *G. sulphuraria*, which may be explained by a loss of the PTS2 targeting pathway. This was shown to be the case in *C. elegans*: proteins normally containing a PTS2 have gained a PTS1 instead (Motley et al, 2000). A similar loss of the PTS2 targeting pathway has been proposed for *T. pseudonana* and the red alga *Cyanidioschyzon merolae* (Gonzalez et al, 2011). As *G. sulphuraria* is a red alga belonging to the same family as *C. merolae* (Cyanidiaceae), it is likely that the same happened in *G. sulphuraria*. Why most species utilise multiple matrix protein targeting pathways as opposed to just one is unclear. It could be that proteins of different pathways are differentially expressed depending on growth conditions, as is the case for PEX5 and its copy PEX9 in *S. cerevisiae* (Effelsberg et al, 2016, Yifrach et al, 2016). In a similar vein, it may be a matter of targeting priority, with one pathway responsible for targeting key proteins, while the other targets proteins that are less important. Another possibility is that the location of the targeting signal at either the N- or C-terminus affects protein function, making one of the targeting signals not feasible for a particular protein.

PEX5 is conserved in all eukaryotes analysed and is characterized by a disordered region at the N-terminal and several tetratricopeptide repeats (TPR) at the C-terminal (*figure 4*). While the TPR domains are responsible for its interaction with the PTS1 motif (Gatto et al, 2000), the N-terminal region interacts with a rarer PTS, PTS3 (Rymer et al, 2018) and with docking proteins PEX13 and PEX14 (Otera et al, 2002, Saidowsky et al, 2001), with the interacting regions partially overlapping (Rymer et al, 2018). As previously recognized, the structurally disordered region at the N-terminal of some PEX5 proteins shares sequence similarities with the fungi-specific PEX20 proteins (Kiel et al, 2006). These similarities between the PEX5 N-terminal and PEX20 rely on: i) a conserved motif

at the N-terminal domain, ii) followed by one or more WxxxF/Y motifs and iii) a PEX7-binding domain (Schliebs & Kunau, 2006). The conserved N-terminal domain of PTS2 co-receptors contains a highly conserved cysteine residue (Schliebs & Kunau, 2006), which has been implicated in (co-)receptor recycling and cargo translocation (Hensel et al, 2011, Leon & Subramani, 2007, Okumoto et al, 2011). The WxxxY/F motifs are important for binding to PEX14 and PEX13 (Otera et al, 2002, Saidowsky et al, 2001). These WxxxF motifs are not only found in PTS2 co-receptors, but also in PEX5 of species where PEX5 does not act as PTS2 co-receptor but only as PTS1 receptor (Schliebs et al, 1999). As the name implies, the PEX7-binding domain allows the co-receptors to bind to PEX7. We checked the conservation of this domain by manually generating a hidden Markov model of the fungal PEX20, and found that this domain is detected in some but not all PEX5 orthologs that act as PTS2 co-receptors (see *figure 4*; Pex20* domains in PEX5).

Phylogeny shows that vertebrates and *S. cerevisiae* have duplicated their PEX5 gene independently (*figure 4*). In *S. cerevisiae*, PEX5 works as a general import receptor for all PTS1-containing peroxisomal matrix proteins, while its paralog PEX9 acts as a condition-specific receptor for a subset of PTS1 proteins (Effelsberg et al, 2016, Yifrach et al, 2016). PEX9 has lost the N-terminal disordered region that is normally present in PEX5 (see *figure 4*). Vertebrates express PEX5R, a PEX5-related protein also called TRIP8b. PEX5R is preferentially expressed in the brain and can bind PTS1-containing proteins *in vitro* (Amery et al, 2001). Nevertheless, it is unclear whether PEX5R plays any role in matrix protein targeting, although the paralogizations of PEX5 could involve different functional novelties for peroxisome protein import as in *S. cerevisiae*.

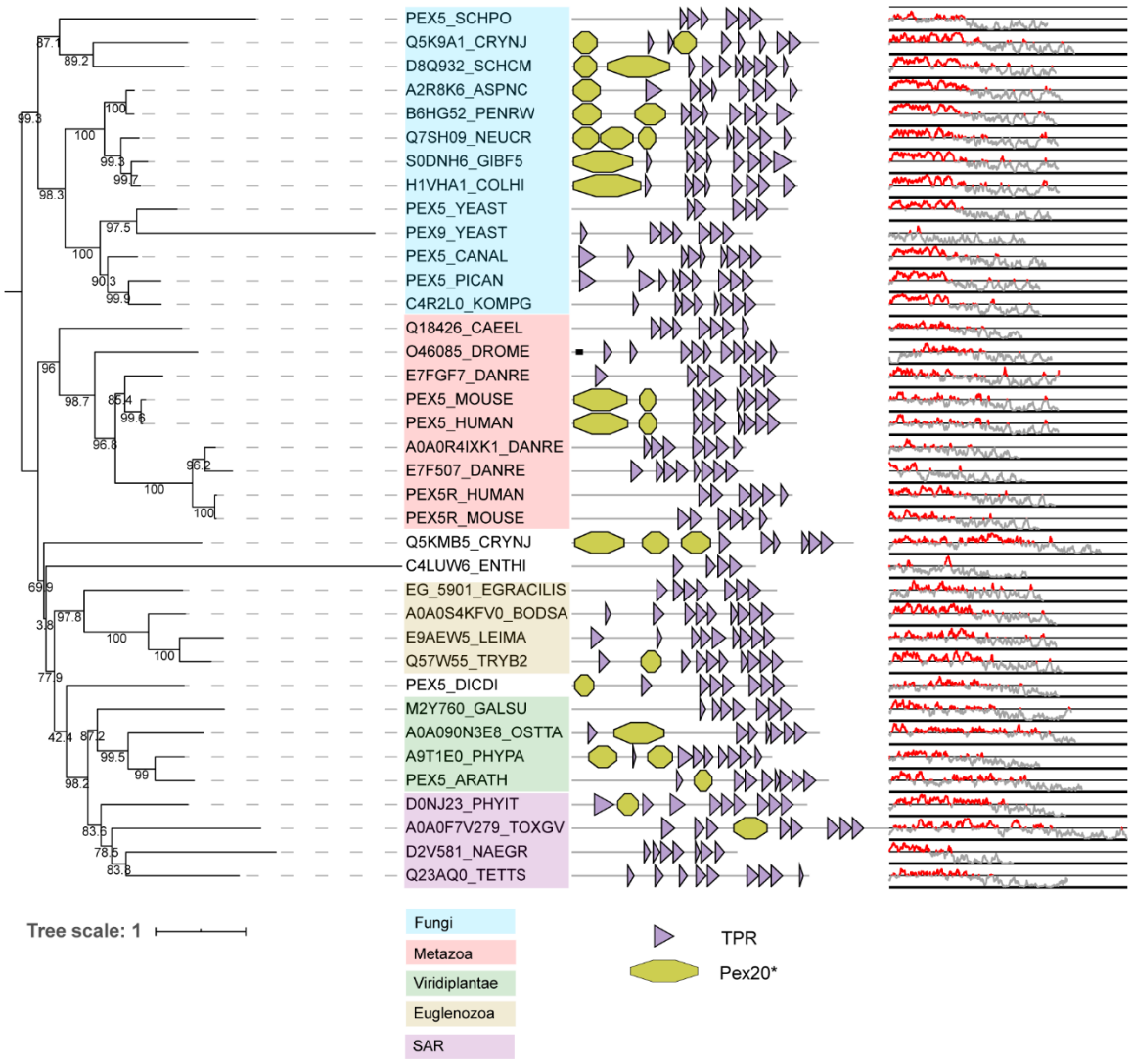

Figure 4. Phylogeny and protein features of PEX5 orthologs. The phylogeny is rooted at mid-point to ease the visualization and labels of the main taxonomic groups are coloured accordingly to the legend. Note that the topology does not necessarily reflect the actual evolutionary trajectory of such proteins. Protein domain architecture is defined by pfam annotations. The Pex20* is a manually generated hidden Markov model (CSM, this study). The line-dot plot indicates the regions predicted be disordered (red) and not disordered (grey).

The docking site (PEX13 and PEX14)

Once the peroxisomal matrix protein is bound to its receptor, the receptor-cargo complex associates to the docking complex, consisting of PEX13 and PEX14 (and in fungi PEX17 or PEX33), at the peroxisomal membrane (*figure 1*).

Transmembrane helices were predicted in some, but not all, PEX13 orthologs (*figure S1A*). In addition, only in Opisthokonta organisms (fungi and metazoa) and amoebozoa, PEX13 has a predicted SH3 domain at the C-terminal (*figure S1A*), which likely controls its interaction with other proteins. PEX14 also contains a predicted transmembrane helix, but seems to be largely structurally disordered (*figure S1B)*, although it also includes several coiled-coil domains (e.g. (Lill et al, 2020)). *In vitro* protease protection experiments using human PEX13 and PEX14 confirmed that both proteins are integral membrane proteins. Human PEX14 has an N_{in}-C_{out} topology, while

PEX13 adopts an N_{out} -C_{in} topology, thereby exposing its SH3 domain to the peroxisomal matrix (Barros-Barbosa, Ferreira et al, 2019). The architecture of the *S. cerevisiae* PEX14-PEX17 complex was recently elucidated and revealed that PEX14 forms a 3:1 heterotetrameric complex with PEX17, forming a rod-like structure of approximately 20 nm that is exposed to the cytosol (Lill et al, 2020). This structure is mainly formed by the coiled-coil domains of PEX14 and PEX17. Besides its coiled-coil domains, PEX14 has a predicted intrinsically disordered C-terminal domain, which may be involved in recruiting import receptor PEX5 (Lill et al, 2020).

After docking, the cargo is translocated into the peroxisomal matrix. For *S. cerevisiae* PTS1 protein import it was shown that PEX5 integrates into the peroxisomal membrane to form a transient translocation pore alongside PEX14 (Meinecke et al, 2010). For PTS2 import, the pore is formed by PEX14, PEX17 and PEX18 (Montilla-Martinez et al, 2015). Little is known about the matrix protein import pores in other organisms, but the involvement of PEX14 seems to be a common denominator (Barros-Barbosa, Rodrigues et al, 2019). After formation of the translocation pore, the cargo is released into the peroxisomal matrix.

Receptor ubiquitination (PEX4, PEX22, PEX2, PEX10 and PEX12)

After cargo release, the PTS (co-)receptor needs to be extracted from the peroxisomal membrane, so it can be used in subsequent rounds of peroxisomal matrix protein import (Platta et al, 2014). PEX5 is mono-ubiquitinated at a conserved cysteine, leading to its extraction and recycling (Platta et al, 2014). In most eukaryotes, this ubiquitination depends on the ubiquitin-conjugating enzyme (Ubc or E2 enzyme) PEX4, associated to the peroxisomal membrane via PEX22, and on the ubiquitin ligase activities of PEX2, PEX10 and PEX12. Notably, PEX4 and PEX22 are absent in metazoa. However, mono-ubiquitination of PEX5 occurs in a comparable manner in mammalian cells through the E2D proteins UbcH5a/b/c (Grou et al, 2008). They are the closest functional counterparts to PEX4. Their actual orthologs in fungi, the Ubc enzymes, are also involved in PEX5 mono-ubiquitination (Platta et al, 2014). This reveals that in the absence of a PEX4 ortholog, functional compensation in specific organisms is possible, showing that the ubiquitination process can be shifted between subfamilies of the whole ubiquitin conjugating enzyme family. Thus, for other organisms lacking PEX4 and PEX22, it could be expected that other E2 enzymes perform this function.

The RING finger complex proteins PEX2/10/12 have ubiquitin (E3) ligase activity (Platta et al, 2014) and are broadly conserved in eukaryotes (*figure 1*). The three paralogous proteins PEX2, PEX10 and PEX12 form a heterotrimeric complex (El Magraoui et al, 2012). Characteristic for these three proteins is a highly conserved region at the N-terminus (annotated as Pex2_Pex12 pfam)

and a zf-RING finger domain at the C-terminus. While the first domain can display a transmembrane helix (predicted in some of the species, suggesting membrane anchoring), the latter domain is responsible for the E3 ubiquitin ligase activity of the proteins (Platta et al, 2014) (*figure 5*). The strong conservation of both domains in most of the sequences could indicate that the cooperation of both domains is crucial for peroxisome biology. The phylogeny of these enzymes, which clearly establishes the three main subfamilies (PEX2, PEX10 and PEX12 that each contain organisms from almost all lineages), suggests that they are deep paralogs and that their functional speciation was important and early in eukaryotic evolution.

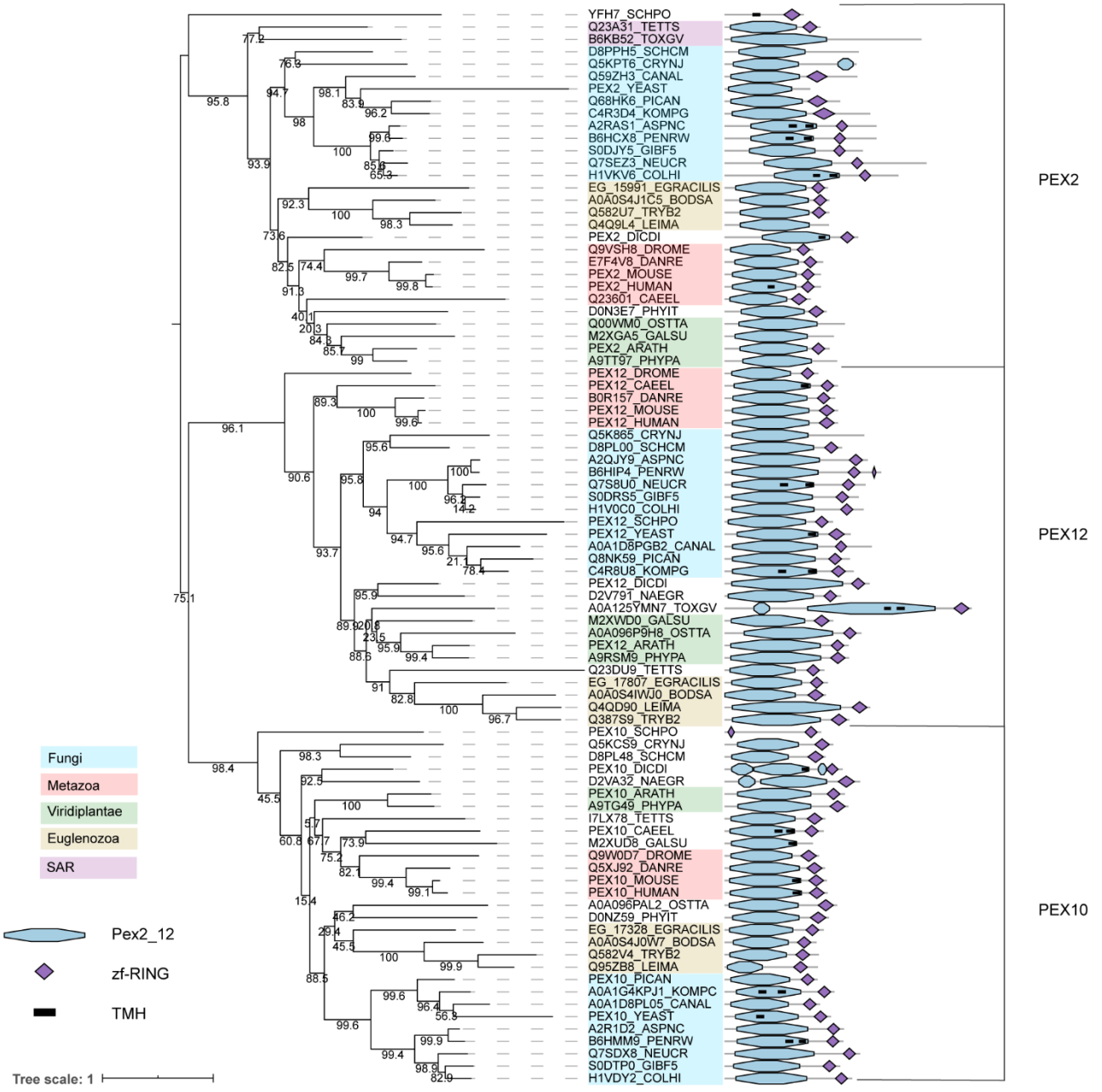

Figure 5. Phylogeny and protein features of PEX2/10/12 orthologs. The phylogeny is rooted at mid-point to ease the visualization and labels of the main taxonomic groups are coloured accordingly to the legend. Note that the topology does not necessarily reflect the actual evolutionary trajectory of such proteins. Protein domain architecture is defined by pfam annotations and transmembrane helix according to TMHMM software.

Receptor extraction (PEX1/6)

Once PEX5 is ubiquitinated, peroxisomal AAA+ ATPases PEX1 and PEX6 are responsible for PEX5 export from the peroxisomal membrane in order to recycle it back to the cytosol. PEX1 and PEX6 belong to the AAA (ATPase associated with diverse cellular activities) family (Pedrosa et al, 2018), a group of protein motors that use ATP binding and hydrolysis to mechanically unfold, disaggregate or remodel substrates (Olivares et al, 2016). Proteins of this family form ring structures with a central channel, through which they can translocate their substrates (Gates &

Martin, 2020). PEX1 and PEX6 form a hetero-hexameric complex with alternating subunits in a double-ring structure (Blok et al, 2015, Gardner et al, 2015). In *S. cerevisiae*, the complex mechanically unfolds its substrates via progressive threading in an ATP-dependent manner (Gardner et al, 2015). Pedrosa et al. (2018) demonstrated using an *in vitro* setup that the PEX1/PEX6 complex directly interacts with ubiquitinated (human) PEX5, unfolding it during extraction (Pedrosa et al, 2018). The phylogeny of PEX1 and PEX6 splits both subfamilies, while their protein domain architecture shows that the architecture is more conserved in PEX1 than in PEX6 (*figure S2*). Similar to PEX2/10/12, these facts suggest that the functional speciation of PEX1 and PEX6 was also important and early in Eukaryotes.

The Pex11 family

Pex11 family proteins coordinate peroxisome proliferation (Koch et al, 2010). The Pex11 family is a large and complex protein family, with some members containing predicted transmembrane helices. Its phylogeny shows that it has been differentially extended in specific organisms (*figure 6***)**, meaning that in different lineages, independent paralogizations have occurred over time. Notwithstanding the low sequence conservation, provoking weak support in some basal nodes in the phylogeny (bootstraps lower than 80%), we can distinguish two main groups within the Pex11 protein family, which we call Pex11Y and Pex11Z here (*figure 6*). Both groups contain organisms from most taxonomic lineages, with the exception of plants, which apparently do not have Pex11Z, although they have intermediary Pex11 sequences that fall outside our Pex11Y/Z groups (along with other Pex11 protist sequences; *figure 6*). Due to the limitations of this phylogeny, it is unclear whether these forms are actually deep paralogs or whether they represent alternative evolutionary histories. However, we will follow our proposed Pex11Y/Z nomenclature to ease the functional contextualization of these paralogues.

The phylogeny of PEX11 shows that these (hypothetic) deep paralogues Pex11Y/Z have subsequently undergone independent paralogizations in different lineages. For instance, Pex11Y was clearly duplicated independently in vertebrates and in several filamentous fungi. Notably, certain further paralogizations seem to have undergone extreme sequence divergence, probably providing artefactual clustering like the fungi-specific PEX25/27/34/36 subgroup within PEX11Y, which contains shortened proteins up to 144 amino acids. Fungal PEX11 and human PEX11α and PEX11β (all PEX11Y subfamily) contain a conserved amphipathic helix capable of tubulating negatively charged membranes *in vitro* (Opaliński et al, 2011, Yoshida et al, 2015). We mapped this amphipathic helix onto the multiple sequence alignment of PEX11 family proteins, observing that

three positively charged residues are generally conserved in these proteins. However, the second positively charged position is not conserved in the Pex11Z subfamily (*figure S3*), suggesting possible functional difference between Pex11Y and Pex11Z. Furthermore, we observed that *S. cerevisiae* PEX34 has lost this amphipathic helix, while the C-terminal region is conserved (*figure S3*).

Several members of the Pex11 protein family have been studied. So far, the majority of studies have investigated members of the PEX11Y subfamily, which includes PEX11 from fungi and PEX11α/β from mammals. In yeasts, the absence of PEX11 results in fewer and larger peroxisomes, while cells overexpressing PEX11 have increased peroxisome numbers with smaller (Erdmann & Blobel, 1995, Joshi et al, 2012, Krikken et al, 2009). Similarly, overproduction of PEX11α or PEX11β in vertebrates induces peroxisome proliferation, while reduction of protein levels resulted in lower peroxisome numbers (Li & Gould, 2002, Schrader et al, 1998). This led to the hypothesis that these proteins play a role in peroxisome fission. Peroxisome fission takes place in three steps: organelle elongation, constriction and scission (Schrader et al, 2016). PEX11 plays a role in the first step where it functions in membrane remodelling (Schrader et al, 2016). So far, no proteins have been identified that are responsible for organelle constriction. Peroxisomal fission shares several components with the mitochondrial fission machinery, such as the dynamin related protein Dnm1 (Drp1/DLP1), Fis1 and Mff (Schrader et al, 2016). Human PEX11β recruits DRP1 to the peroxisomal membrane (Koch & Brocard, 2012, Li & Gould, 2003), and both *S. cerevisiae* PEX11 and human PEX11β have been reported to function as GTPase activating protein (GAP) for Dnm1 (DRP1) (Williams et al, 2015).

Several other functions have been attributed to proteins of the Pex11Y subfamily. *O. polymorpha* PEX11 has been implicated in peroxisome segregation during cell division (Krikken et al, 2009). *S. cerevisiae* PEX11 and PEX34 are involved in peroxisome-mitochondria contact sites (Shai et al, 2018, Ušaj et al, 2015), while *O. polymorpha* PEX11 has been implicated in peroxisome-ER contact sites (Wu et al, 2020). *S. cerevisiae* PEX11 has also been proposed to act as a pore-forming protein (Mindthoff et al, 2016) and has been implicated in medium chain fatty acid oxidation as well (van Roermund et al, 2000). As only a subset of proteins from the PEX11Y subfamily have been investigated, perhaps other functions will still be discovered.

Much less is known about proteins of the PEX11Z subfamily, which includes PEX11γ from metazoa, fungal PEX11C and GIM5A/B from *T. brucei*. However, they also play a role in peroxisome proliferation (see e.g. (Koch & Brocard, 2012, Opaliński et al, 2012)). PEX11γ has

been suggested to coordinate peroxisomal growth and division via heterodimerisation with other mammalian PEX11 paralogs and interaction with Mff and Fis1 (Schrader et al, 2016). *O. polymorpha* PEX11C is downregulated upon shifting from peroxisome repressing (glucose) to peroxisome inducing (methanol) growth conditions (van Zutphen et al, 2010) suggesting that PEX11C is not required for peroxisome proliferation. In *Penicillium rubens*, deletion of PEX11C has no significant effect on peroxisome number or size, while overexpression strongly stimulates peroxisome proliferation (Opaliński et al, 2012). In *T. brucei*, the absence of both GIM5A and GIM5B is fatal, due to cellular fragility (Voncken et al, 2003). In *S. cerevisiae* proteins of the PEX11Z subfamily are absent.

The remaining proteins, which do not have a clear evolutionary relationship with each other and fall outside the Pex11Y/Z subfamilies, we call the Pex11-like proteins. The most studied proteins from this group are *A. thaliana* PEX11C/D/E. These proteins cooperate with FIS1b and DRP3A in peroxisome growth and division during the G_2 phase just prior to mitosis (Lingard et al, 2008). Interestingly, in cells where PEX11C, PEX11D and PEX11E were silenced simultaneously, peroxisomes were enlarged, but not elongated, suggesting that these proteins act in peroxisome growth, but not tubulation (Lingard et al, 2008).

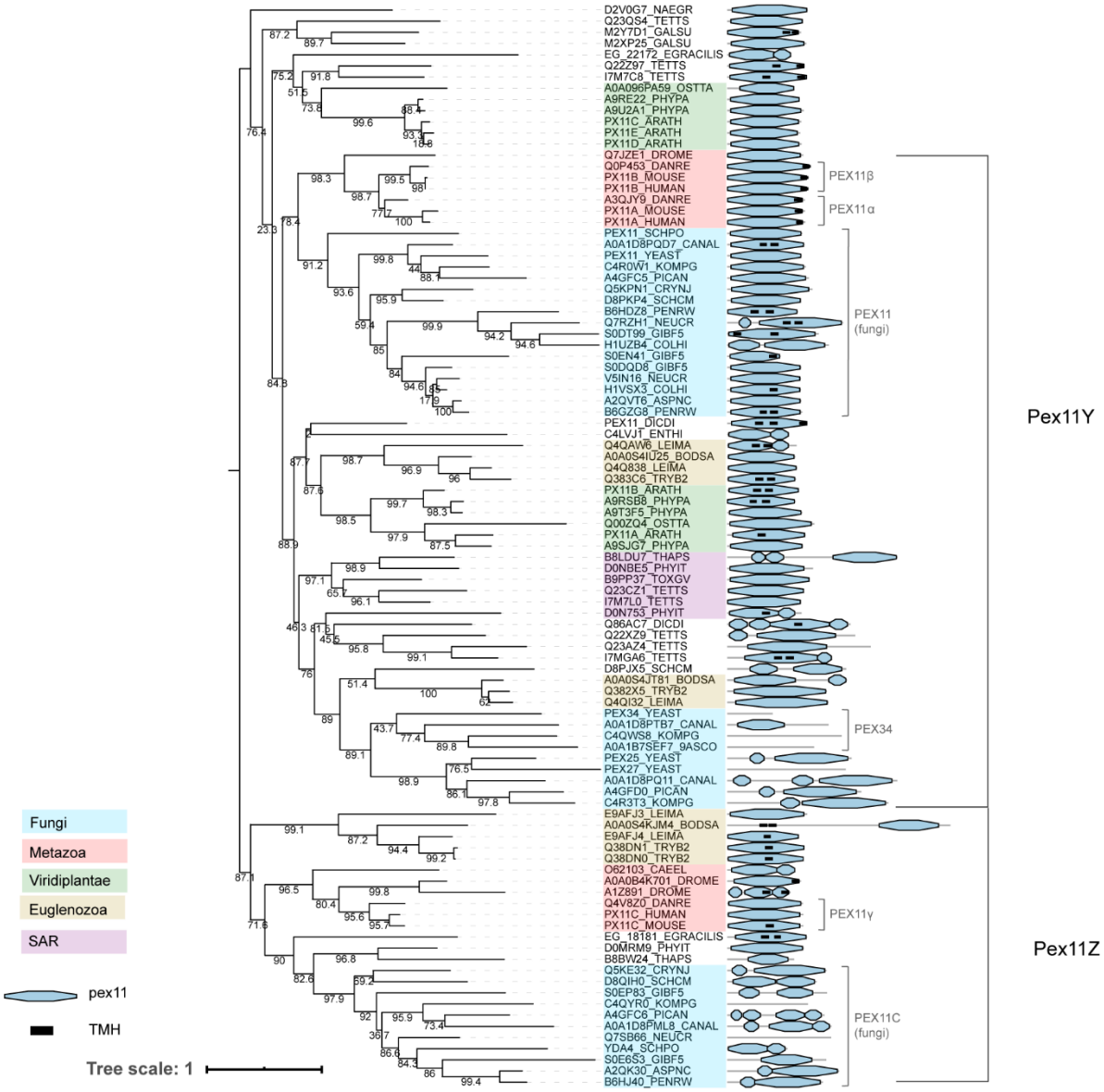

Figure 6. Phylogeny and protein features of PEX11 family proteins. The phylogeny is rooted at mid-point to ease the visualization and labels of the main taxonomic groups are coloured accordingly to the legend. Note that the topology does not necessarily reflect the actual evolutionary trajectory of such proteins. Protein domain architecture is defined by pfam annotations and transmembrane helix prediction (black box). The Pex11Y and Pex11Z subfamilies are named according to the most supported and basal bootstraps and their taxonomic compositions. Note that Viridiplantae organisms do not appear to have Pex11Z, although they have other paralogs outside of both defined subfamilies.

3. PEX proteins specific for fungi

Several PEX proteins are specific to fungi. The high number of known fungal PEX proteins is probably due to the extensive screens for yeast peroxisome-deficient mutants that have been performed in the past (Erdmann et al, 1997). Additionally, current peroxisome biogenesis research is still taking advantage of a wealth of genetic and biochemical toolboxes to analyse the molecular biology of these organelles in yeast.

The PEX7 co-receptors (PEX18, PEX20, PEX21)

In plants, animals and protists like TRYPB2, *D. discoideum* and *L. major*, (a longer splicing variant of) PEX5 acts as PEX7 co-receptor for PTS2 protein import (Schliebs & Kunau, 2006). In contrast, in many fungi the PEX7 co-receptor is a separate PEX protein, namely PEX18, PEX20 or PEX21 (see for more detailed reviews e.g. (Kunze, 2020, Schliebs & Kunau, 2006)). Duplication of the ancestral PEX20 in *S. cerevisiae* (see *figure S4*), resulted in the partially redundant paralogs PEX18 and PEX21 that perform the same function (Purdue et al, 1998). Therefore, these proteins can be considered as a single PEX20 group. As previously described, some sequence features relate PEX20 with the N-terminus of PEX5 proteins: a conserved cysteine, WxxxF motifs and PEX7 binding domain (Schliebs & Kunau, 2006). Due to the fact that that PEX5 is present in most eukaryotes and Pex20 domains can be found at the N-terminus of many such proteins, it is most likely that PEX20 is the result of a protein domain separation specific to fungi, rather than the previously proposed protein fusion of PEX5 and PEX20 (Kiel et al, 2006).

PEX17 and PEX33

In all species, PEX13 and PEX14 are components of the receptor docking site. An additional component of the docking site in yeasts is PEX17, while in filamentous fungi PEX33 is part of the docking complex. PEX17 is characterized by a single transmembrane helix at the N-terminal. As described above, *S. cerevisiae* PEX14 and PEX17 together form a rod-like structure at the peroxisomal membrane (Lill et al, 2020). PEX33 is a paralog of PEX14, whereas PEX17 is a protein partially aligning to the C-terminal of PEX14 and PEX33, suggesting PEX17 is a PEX14-like protein. The exact functions of PEX17 and PEX33 are still unclear, but PEX17 in *S. cerevisiae* is a main component of the PTS2 import pore (Montilla-Martinez et al, 2015) and seems to increase the efficiency of binding of import receptors PEX5 and PEX7 to the docking complex (Lill et al, 2020).

PEX8

In fungi, intraperoxisomal protein PEX8 bridges the docking and RING finger complexes (PEX2/10/12) (Agne et al, 2003). Little else is known about PEX8, but it has been implicated in cargo release from the PTS1 receptor PEX5 (Ma et al, 2013, Wang et al, 2003).

Pex23 family proteins

PEX23, PEX24, PEX29, PEX32 (for *O. polymorpha* for example) and PEX28, PEX29, PEX30 PEX31, PEX32 (for *S. cerevisiae*) are homologous proteins containing a highly conserved domain called the Pex24p domain (pfam). This domain contains a Dysferlin (DysF) motif at the C-terminal

region, the function of which is still unclear (Wu et al, 2020). At the N-terminal, these proteins have several transmembrane domains suggesting that these proteins are anchored to membranes. A group of proteins related to this Pex23 protein family are the Pex23-like proteins (Kiel et al, 2006), including SPO73, a protein involved in sporulation. Pex23-like proteins do not usually present the region containing the predicted transmembrane helices. The phylogeny of all these proteins can be divided into three main groups that here we call PEX23 subfamily, PEX24 subfamily and Pex23-like proteins (*figure 7A*). The sequences from the PEX23 and PEX24 subfamilies appear to differ in protein extensions at their C- and N- termini respectively, with predicted structural protein disordered regions. Due to the fact that the main PEX23 and PEX24 subfamilies contain most of the fungi analysed, it is likely that both subfamilies originated from an ancestral duplication in fungi. Later, these PEX23 and PEX24 paralogs duplicated in yeasts leading to amongst others PEX28/PEX29 and PEX30/31/32 in the ancestor of *S. cerevisiae.* In filamentous fungi on the other hand, no duplication occurred, and these fungi express only one protein of each group. Thus, these proteins have diversified differentially in Fungi.

Unlike other peroxins, proteins of the Pex23 family localise to the ER instead of peroxisomes. Although initially reported at the peroxisome (Brown et al, 2000, Tam & Rachubinski, 2002, Vizeacoumar et al, 2003, Vizeacoumar et al, 2004), later studies either reported dual localization to peroxisomes and ER (David et al, 2013, Yan et al, 2008) or exclusive localization at ER subdomains (Joshi et al, 2016, Mast et al, 2016, Wu et al, 2020). A recent study characterizing *O. polymorpha* Pex23 family members reported the involvement of PEX24 and PEX32 in peroxisome-ER contact sites (Wu et al, 2020). This could explain the previous contradictory reports on their localization, as they can be expected to be present in spots where peroxisomes and ER interact. Furthermore, *S. cerevisiae* PEX30 and PEX31 are ER membrane shaping proteins (Joshi et al, 2016). *S. cerevisiae* PEX30 plays a role in regulating budding of pre-peroxisomal vesicles and lipid droplets from specific ER subdomains (Joshi et al, 2016, Joshi et al, 2018). It has been proposed to facilitate this by collaborating with seipin to organize ER subdomains to alter the membrane lipid composition (Wang et al, 2018). In humans, no orthologs of PEX30 have been identified, but MCTP2 has been suggested to act as a functional analog (Joshi et al, 2018).

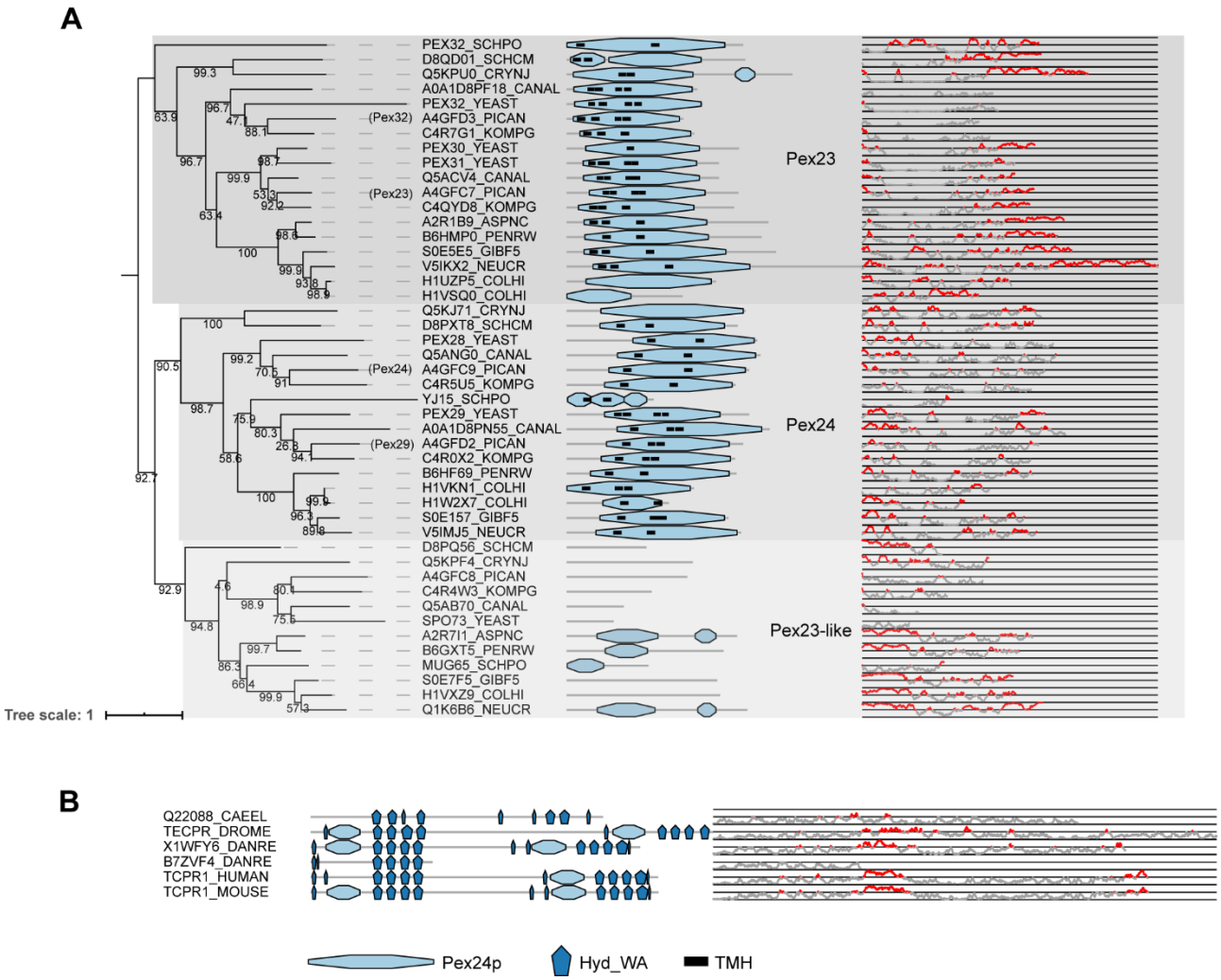

Figure 7. Phylogeny and protein features of Pex23 protein family and TECPR1 proteins.

A) Phylogeny and protein features of the fungal Pex23 protein family.

The phylogeny is rooted at mid-point to ease the visualization. The main phylogenetic groups are named and highlighted according the protein names of *O. polymorpha*, indicated between brackets. Protein domain architecture is defined by pfam annotations and transmembrane helix according to TMHMM software. The Pex24p pfam domain contains the DysF motifs. The line-dot plot, indicates the region predicted be disordered (red) and not disordered (grey).

B) Protein features of TECPR1 protein family from metazoa.

PEX23 homologs were found in metazoa, but these proteins cannot be considered orthologs of PEX23. These proteins were previously published as metazoan PEX23 orthologs (e.g. (Di Cara et al, 2017, Jeynov et al, 2006, Mast et al, 2011)) and are also annotated as such in some databases (e.g. protein Q9VWB0|TECPR_DROME annotated as PEX23 in Uniprot and FlyBase). However, their domain architecture (see *figure 7B*) is clearly different from previously established Pex23 family proteins, and they actually belong to the TECPR1 family of proteins. TECPR1 proteins are localized to lysosomes and play a role in autophagy (Chen & Zhong, 2012). While TECPR1 proteins do contain a DysF domain, like the proteins from the PEX23 family, they also contain several tectonin repeats (TECPR) and a PH domain, in addition to a beta-propellor structure

(Ogawa et al, 2011). It is therefore unlikely that they perform a function similar to PEX23 family proteins in fungi and they cannot be considered PEX23 orthologs.

Little is known about both PEX35 and PEX37, but both seem to play a role in regulating peroxisome proliferation. PEX35 is unique to *S. cerevisiae* and closely related species in the Saccharomycetaceae family, while PEX37 is found in most other yeast species and filamentous fungi. PEX35 has no known functional domains or similarity to other known PEX proteins. Only one study investigating PEX35 has been published to date, showing that PEX35 is a PMP that interacts with vesicle budding inducer Arf1 and localizes at the proximity of proteins from the Pex11 family (Yofe et al, 2017). The authors speculate that PEX35 may regulate peroxisome fission alongside proteins of the Pex11 family. *O. polymorpha* PEX37 is a peroxisomal transmembrane protein that affects peroxisome segregation and proliferation under peroxisome-repressing conditions, but not on peroxisome-inducing conditions. So far, only one study has investigated this protein (Singh et al, 2020). PEX37 belongs to the same protein family as human PXMP2, *N. crassa* Woronin body protein Wsc and *S. cerevisiae* mitochondrial inner membrane protein Sym1 and its human homolog MPV17, many of which are thought to act as channels. Human PXPM2 is able to partially rescue the phenotype present in the absence of *O. polymorpha* PEX37, suggesting that these proteins have similar functions (Singh et al, 2020).

4. Moderately conserved PEX proteins

In many species, the PEX1/PEX6 complex is recruited to the peroxisomal membrane via an anchoring protein. These membrane anchors are much less conserved than PEX1 and PEX6 themselves, with different homologous, but not orthologous, proteins acting as anchoring protein in different species. In vertebrates and most fungi, the anchoring protein is PEX26, while in *S. cerevisiae* and closely related species in the Saccharomycetaceae family it is PEX15 (Kiel et al, 2006) and in plants it is APEM9 (Cross et al, 2016). Despite sharing only weak sequence identity, PEX15, PEX26 and APEM9 do have several features in common. All three proteins are tail-anchored proteins (Cross et al, 2016, Halbach et al, 2006) and tether the PEX1/PEX6 complex to the peroxisomal membrane via PEX6 (Birschmann et al, 2003, Goto et al, 2011, Matsumoto et al, 2003) .

Discussion

We used a comparative genomics approach to provide an up-to-date overview of all PEX protein families known to date, in a range of representative organisms from all eukaryotic lineages. Our

computational survey identified a **core set of PEX proteins** that is broadly conserved across all eukaryotic lineages (PEX1/2/3/5/6/7/10/11/12/13/14/16/19). This means that ancestral versions of these PEX proteins were already present in the last eukaryotic common ancestor (LECA) and that this set of proteins defines the minimum set of PEX proteins that is required to make a peroxisome.
Besides a broadly conserved core set of PEX proteins, we found that a large number of **PEX proteins is specific to the kingdom of fungi**. Although there is increasing consensus that homology detection failure is frequent (Weisman et al, 2020), our inner controls (see methods) still suggest that these fungi-specific proteins are absent in other lineages. This exposes not only a bias in peroxisome research towards fungi, but also reveals that peroxisomes are dynamic organelles, their composition evolving under different evolutionary pressures. The loss of specific PEX proteins in some eukaryotes, such as the loss of proteins associated with the PTS2 targeting pathway in *C. elegans* and the loss of PEX16 in *S. cerevisiae* and other yeasts further supports this notion.

Intriguingly, PEX proteins in human pathogens like *T. gondii*, *T. brucei* and *L. major* were often difficult to detect. Moreover, these PEX proteins frequently had additional domains, which could indicate that they may have obtained additional functions. The low homology between PEX proteins of human and human pathogens may be advantageous for the identification of specific drug targets.

In some species lacking peroxisomes such as *E. histolytica*, we still identified some PEX proteins such as PEX5, PEX16 and PEX19 (see *figure 2, 3* and *4*). This could mean that these organisms have lost this organelle relatively recently and thus have not entirely lost all PEX proteins yet, but it could also suggest that the remaining PEX proteins retain non-peroxisomal functions. This is not completely unthinkable, as some PEX proteins have already been suggested to be involved in non-peroxisomal pathways. For instance, human PEX3 and PEX19 have been implicated in targeting of lipid droplet protein UBXD8 (Schrul & Kopito, 2016). On the other hand, the example of *E. histolytica* illustrates drastic evolutionary changes in the peroxisomal biology, a fact already observed in other amoeba species like *Mastigamoeba balamuthi* (Le et al, 2020). Understanding the reason behind these evolutionary adaptations will improve our understanding about the peroxisome biology.

The vast majority of the core PEX proteins (PEX1/2/5/6/7/10/12/13 and 14) are involved in matrix protein import, while only a few (PEX3, PEX16 and PEX19) play a role in PMP sorting. In addition to these core PEX proteins all eukaryotes contain multiple proteins of the Pex11 family, which are

involved in several peroxisome-related processes. It is unclear why so few proteins have been identified that plays a role in PMP sorting. Proteins of the common ER protein sorting machineries, such as the Sec and GET translocons, have been reported to function in the indirect pathway of PMP sorting. The absence of these proteins is lethal in yeast, explaining that such mutants have not been obtained in screens for yeast peroxisome deficient mutants. For the direct pathway of PMP sorting it is unlikely that the entire sorting/insertion machinery consists of only three, or even two for yeast (PEX3/19), proteins.

Most of the currently known PEX genes have been identified in the nineties of the previous century by very successful genetic approaches to identify peroxisome deficient (*pex*) yeast mutants. Yeast *pex* mutants are viable and have distinct growth phenotypes (e.g., deficiency to grow on oleic acid or methanol), which greatly facilitated the isolation of these mutants and cloning of the corresponding genes by functional complementation. Most likely this caused the bias towards fungal PEX genes. In addition to *S. cerevisiae*, which is the main yeast model in cell biology, a few other yeast species were used to identify PEX proteins (*Komagataella phaffiia* [formerly *Pichia pastoris*], *Ogataea polymorpha* [formerly *Hansenula polymorpha*] and *Yarrowia lipolytica*). Notably, several conserved PEX proteins that are present in the latter three yeast species are absent in *S. cerevisiae* (for instance PEX20, PEX26, PEX37 and proteins of the PEX11Z subfamily), while orthologs of the *S. cerevisiae* PEX proteins PEX9, PEX15, PEX35 are absent in all other species that we analysed. This stresses the importance of using several yeast models besides *S. cerevisiae* in cell biology research.

Fusion of human cell lines, derived from patients suffering from peroxisome biogenesis disorders, resulted in the classification of these patients in 12 genotypes/complementation groups (Fujiki, 2016). Using known yeast PEX genes, human orthologues were identified by homology searches on the human expressed sequence tag database. By functional complementation of the cell lines with these putative human PEX genes, 12 of the currently known human PEX genes were identified. Because mislocalisation of the PTS1 protein catalase was used as criterion for peroxisome deficiency, the human PTS2 receptor PEX7 was not identified by this approach (Fujiki, 2016). Together with the results of functional complementation of mutant Chinese hamster ovary (CHO) cell lines, at present 16 mammalian PEX proteins are known (compared with 29 in *S. cerevisiae*).

It is unlikely that all human/mammalian PEX proteins have been identified. Mutations in human/mammalian PEX genes could cause lethal phenotypes, explaining why they have not been isolated in mutant screens. Also, there may be functional redundancy among human PEX genes,

which prevents their identification by mutant complementation approaches. Conversely, mutations in yet unknown mammalian PEX genes could cause relatively weak phenotypes and hence were overlooked. Indeed, the approaches used so far resulted in the identification of PEX11β, but not of PEX11α and PEX11γ. Alternative approaches, like the identification of novel peroxisomal proteins using proteomics of isolated mammalian peroxisomes may result in the characterization of novel mammalian PEX proteins.

PEX proteins (peroxins) were originally defined as proteins "involved in peroxisome biogenesis (inclusive of peroxisomal matrix protein import, membrane biogenesis, peroxisome proliferation, and peroxisome inheritance)" (Distel et al, 1996). However, proteins fitting this definition are not always named as such. For instance, *T. brucei* GIM5A is a member of the Pex11 protein family, but is not named 'PEX'. Also, two proteins involved in peroxisome inheritance, Inp1 and Inp2, are not called PEX. Therefore "inheritance" could be omitted from the original definition of PEX proteins, or these proteins could be renamed. Some proteins that fulfil the PEX protein definition are also involved in other processes and obviously not called PEX. This is for instance the case for the organelle fission proteins FIS1 and DRP1, and ER proteins that play a role in the indirect sorting pathways of PMPs.

Current PEX protein nomenclature has several issues and inconsistencies that can easily lead to confusion. As PEX proteins are numbered chronologically, there is no intuitive link between their names and their function and/or conservation. Additionally, there are several naming inconsistencies relating to PEX protein families. For instance, in higher eukaryotes Pex11 protein family members are named PEX11'X' (e.g., PEX11α/β/γ, PEX11A/B/C). The nomenclature of yeast proteins does not allow the addition of the extra symbol 'X'. These genes invariably consist of a three-letter code (PEX) followed by a number, explaining why PEX11 orthologs in yeast have been designated PEX25, PEX27 and PEX34, not PEX11X.

Since most PEX proteins were initially identified in yeast species and numbered in the order in which they were described, proteins belonging to the same protein family have received different names. For instance, the two AAA ATPases are called PEX1 and PEX6, while the three RING proteins are called PEX2, PEX10 and PEX12. Lastly, there are proteins carrying the same name that are not actually orthologs (e.g., PEX23 in metazoa). In summary, current PEX protein nomenclature can easily lead to confusion as it is often far from intuitive, sometimes inconsistent and occasionally wrong. This not only leads to confusion within the peroxisome field, but the large number of PEX proteins numbering up to 37 can be quite intimidating for researchers from other fields.

We therefore suggest that it may be prudent to come up with a **new naming system**. Although it is beyond the scope of the current paper, similar new naming systems are not unprecedented. Indeed, the name PEX protein itself was devised to unify nomenclature regarding proteins involved in peroxisome biogenesis (Distel et al, 1996), thereby re-naming the 13 proteins known at the time to be involved in peroxisome biogenesis. More recently, proteins involved in mitochondrial contact site and cristae organizing system (MICOS) (Pfanner et al, 2014), autophagy-related proteins(Klionsky et al, 2003) and ribosomal proteins (Ban et al, 2014) have been re-named. In addition, we recommend setting up guidelines for naming newly discovered 'PEX proteins', taking into account phylogeny to extend to ortho- and in-paralogues. Moreover, we propose amending the definition of 'PEX proteins' as posed in 1996 (Distel et al, 1996). Proteins involved in peroxisome inheritance such as Inp1 and Inp2 have so far been named differently and should be removed from the definition.

Adopting an entirely new naming system may be very difficult. However, it would already be very helpful to only re-name the most confusing and inconsistent parts. The two largest protein families, the Pex11 family and the Pex23 family, together make up about one-third of all PEX numbers and are arguably the most confusingly named.

Acknowledgements

RLMJ is supported by a grant from the Netherlands Organization of Scientific Research (NWO), section Earth and Life Sciences (ALWOP161). DPD was supported by the Spanish Ministry of Economy and Competitiveness (Grant nos. BFU2016-78326-P). CSM is supported by the Joined Gordon and Betty Moore Foundation with the Simons foundation (Grant Agreement #9733).

Author contributions

RLMJ, CSM, MVDN, DPD and IvdK conceived the project; RLMJ, CSM, MVDN, DPD analysed the data and prepared the figures; RLMJ and IJvdK wrote the original draft. All contributed to reviewing and editing the manuscript.

Conflict of interest

The authors declare no conflict of interest.

Methods

Ortholog identification of PEX proteins

For the ortholog detection of PEX proteins, we systematically used two approaches: reciprocal searches of single protein sequences and reciprocal searches based on protein profiles (Hidden Markov models). We selected a set of eukaryotic proteomes from UniProt (Anonymous , 2017) (see *table 1*) and for both approaches, performed the reciprocal searches starting from the sequences of different organisms (see table) and made a consensus for the assignment of orthologs between the searches.

The first approach was based on phmmer searches (HMMER package (Potter et al, 2018)). As peroxisomal proteins can be multidomain proteins, when the first reciprocal hit failed, we also checked the best domain e-value hit from the target proteome. In this way, we also retrieve potential orthologs taking into account alternative domain architecture. The second approach was based on reciprocal jackhmmers followed by hmmsearches (HMMER package (Potter et al, 2018)). This method is applied in order to detect divergent orthologs undetectable by the previous approach, although it can be problematic for proteins containing common domains like PEX1/6, PEX4 (containing functional domains like WD40, ATPase, zinc-finger and ubiquitin ligases; see table). Due to the diverse nature of the PEX proteins, different e-value thresholds and iterations were applied. For example, searches involving transmembrane proteins and tandem protein repeats (TPR) were conducted with 2 iterations and a relaxed e-value, 1e-2. Alternatively, for the other common domains, we applied 2 iterations and constrained e-value, 1e-20. The reciprocal detection for these common domains were often/frequently unsatisfying showing the limitation of this method for abundant and common domains.

Once the ortholog assignment of both methods combined, for each set of orthologs we manually filtered-out possible false positive by performing a multiple sequence alignment using Mafft (einsi-mode (Katoh & Standley, 2013)) followed by visual inspection. We additionally searched for missing orthologs. We built HMM profiles through Hmmbuild using the MSA generated previously and made searches into the suspect proteome through Hmmsearch (both from the HMMER package (Potter et al, 2018)). It is important to note that if no orthologs were identified for a particular PEX protein in a specific organism, this does not necessarily mean that no ortholog exists. Possible causes of not identifying orthologs are incomplete genome information and sequence divergence of the ‘true’ ortholog. For example, the *T. pseudonana* proteome seems to be incomplete in the Uniprot database: a previous study identified a *T. pseudonana* Pex12 ortholog (Mix et al, 2018) that matches our criteria for orthology, but is absent from Uniprot.

Ortholog sequences included in the final dataset were aligned with Mafft, and trimmed the gap position with Trimal using different thresholds. Phylogenetic trees were constructed using IQ-TREE (Nguyen et al, 2015) obtaining branch supports with ultrafast bootstrap (Hoang et al, 2018) and applying the automatic model selection calculated by ModelFinder (Kalyaanamoorthy et al, 2017). Trees were visualized and annotated using iTOL (Letunic & Bork, 2019). Functional domain annotation was carried out using the Pfam database (El-Gebali et al, 2019), transmembrane domains using the TMHMM server (http://www.cbs.dtu.dk/services/TMHMM/) and structural disorder with IUPred2 (Mészáros et al, 2018).

Literature

UniProt: the universal protein knowledgebase. (2017) *Nucleic Acids Res* **45**: D158-D169

Agne B, Meindl NM, Niederhoff K, Einwächter H, Rehling P, Sickmann A, Meyer HE, Girzalsky W, & Kunau W (2003) Pex8p: an intraperoxisomal organizer of the peroxisomal import machinery. *Mol Cell* **11**: 635-646

Agrawal G, Fassas SN, Xia Z, & Subramani S (2016) Distinct requirements for intra-ER sorting and budding of peroxisomal membrane proteins from the ER. *J Cell Biol* **212**: 335-348

Amery L, SANO H, MANNAERTS GP, SNIDER J, van LOOY J, Fransen M, & Van Veldhoven PP (2001) Identification of PEX5p-related novel peroxisome-targeting signal 1 (PTS1)-binding proteins in mammals. *Biochem J* **357**: 635-646

Ban N, Beckmann R, Cate JH, Dinman JD, Dragon F, Ellis SR, Lafontaine DL, Lindahl L, Liljas A, & Lipton JM (2014) A new system for naming ribosomal proteins. *Curr Opin Struct Biol* **24**: 165-169

Barros-Barbosa A, Ferreira MJ, Rodrigues TA, Pedrosa AG, Grou CP, Pinto MP, Fransen M, Francisco T, & Azevedo JE (2019) Membrane topologies of PEX 13 and PEX 14 provide new insights on the mechanism of protein import into peroxisomes. *The FEBS journal* **286**: 205-222

Barros-Barbosa A, Rodrigues TA, Ferreira MJ, Pedrosa AG, Teixeira NR, Francisco T, & Azevedo JE (2019) The intrinsically disordered nature of the peroxisomal protein translocation machinery. *The FEBS journal* **286**: 24-38

Birschmann I, Stroobants AK, van den Berg M, Schäfer A, Rosenkranz K, Kunau W, & Tabak HF (2003) Pex15p of Saccharomyces cerevisiae provides a molecular basis for recruitment of the AAA peroxin Pex6p to peroxisomal membranes. *Mol Biol Cell* **14**: 2226-2236

Blok NB, Tan D, Wang RY, Penczek PA, Baker D, DiMaio F, Rapoport TA, & Walz T (2015) Unique double-ring structure of the peroxisomal Pex1/Pex6 ATPase complex revealed by cryo-electron microscopy. *Proceedings of the National Academy of Sciences* **112**: E4017-E4025

Brown TW, Titorenko VI, & Rachubinski RA (2000) Mutants of the Yarrowia lipolytica PEX23 gene encoding an integral peroxisomal membrane peroxin mislocalize matrix proteins and accumulate vesicles containing peroxisomal matrix and membrane proteins. *Mol Biol Cell* **11**: 141-152

Buckner FS, Eastman RT, Nepomuceno-Silva JL, Speelmon EC, Myler PJ, Van Voorhis WC, & Yokoyama K (2002) Cloning, heterologous expression, and substrate specificities of protein farnesyltransferases from Trypanosoma cruzi and Leishmania major. *Mol Biochem Parasitol* **122**: 181-188

Chen D & Zhong Q (2012) A tethering coherent protein in autophagosome maturation. *Autophagy* **8**: 985-986

Cross LL, Ebeed HT, & Baker A (2016) Peroxisome biogenesis, protein targeting mechanisms and PEX gene functions in plants. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1863**: 850-862

David C, Koch J, Oeljeklaus S, Laernsack A, Melchior S, Wiese S, Schummer A, Erdmann R, Warscheid B, & Brocard C (2013) A combined approach of quantitative interaction proteomics and live-cell imaging reveals a regulatory role for endoplasmic reticulum (ER) reticulon homology proteins in peroxisome biogenesis. *Mol Cell Proteomics* **12**: 2408-2425

Di Cara F, Sheshachalam A, Braverman NE, Rachubinski RA, & Simmonds AJ (2017) Peroxisome-mediated metabolism is required for immune response to microbial infection. *Immunity* **47**: 93-106. e7

Distel B, Erdmann R, Gould SJ, Blobel G, Crane DI, Cregg JM, Dodt G, Fujiki Y, Goodman JM, & Just WW (1996) A unified nomenclature for peroxisome biogenesis factors. *J Cell Biol* **135**: 1-3

Ebenezer TE, Zoltner M, Burrell A, Nenarokova A, Vanclová AMN, Prasad B, Soukal P, Santana-Molina C, O'Neill E, & Nankissoor NN (2019) Transcriptome, proteome and draft genome of Euglena gracilis. *BMC biology* **17**: 11

Effelsberg D, Cruz-Zaragoza LD, Schliebs W, & Erdmann R (2016) Pex9p is a new yeast peroxisomal import receptor for PTS1-containing proteins. *J Cell Sci* **129**: 4057-4066

El Magraoui F, Bäumer BE, Platta HW, Baumann JS, Girzalsky W, & Erdmann R (2012) The RING-type ubiquitin ligases Pex2p, Pex10p and Pex12p form a heteromeric complex that displays enhanced activity in an ubiquitin conjugating enzyme-selective manner. *The FEBS journal* **279**: 2060-2070

El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, & Smart A (2019) The Pfam protein families database in 2019. *Nucleic Acids Res* **47**: D427-D432

Emmanouilidis L, Schütz U, Tripsianes K, Madl T, Radke J, Rucktäschel R, Wilmanns M, Schliebs W, Erdmann R, & Sattler M (2017) Allosteric modulation of peroxisomal membrane protein recognition by farnesylation of the peroxisomal import receptor PEX19. *Nature communications* **8**: 1-13

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

Erdmann R, Veenhuis M, & Kunau W (1997) Peroxisomes: Organelles at the crossroads. *Trends Cell Biol* **7**: 400-407

Fakieh MH, Drake PJ, Lacey J, Munck JM, Motley AM, & Hettema EH (2013) Intra-ER sorting of the peroxisomal membrane protein Pex3 relies on its luminal domain. *Biology open* **2**: 829-837

Fujiki Y (2016) Peroxisome biogenesis and human peroxisome-deficiency disorders. *Proc Japan Acad , Ser B* **92**: 463-477

Gardner BM, Chowdhury S, Lander GC, & Martin A (2015) The Pex1/Pex6 complex is a heterohexameric AAA motor with alternating and highly coordinated subunits. *J Mol Biol* **427**: 1375-1388

Gates SN & Martin A (2020) Stairway to translocation: AAA motor structures reveal the mechanisms of ATP-dependent substrate translocation. *Protein Science* **29**: 407-419

Gatto GJ, Geisbrecht BV, Gould SJ, & Berg JM (2000) Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. *Nat Struct Biol* **7**: 1091-1095

Gonzalez NH, Felsner G, Schramm FD, Klingl A, Maier U, & Bolte K (2011) A single peroxisomal targeting signal mediates matrix protein import in diatoms. *PloS one* **6**: e25316

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

Grou CP, Carvalho AF, Pinto MP, Wiese S, Piechura H, Meyer HE, Warscheid B, Sa-Miranda C, & Azevedo JE (2008) Members of the E2D (UbcH5) family mediate the ubiquitination of the conserved cysteine of Pex5p, the peroxisomal import receptor. *J Biol Chem* **283**: 14190-14197

Halbach A, Landgraf C, Lorenzen S, Rosenkranz K, Volkmer-Engert R, Erdmann R, & Rottensteiner H (2006) Targeting of the tail-anchored peroxisomal membrane proteins PEX26 and PEX15 occurs through C-terminal PEX19-binding sites. *J Cell Sci* **119**: 2508-2517

Han Y, Lyman KA, Foote KM, & Chetkovich DM (2020) The structure and function of TRIP8b, an auxiliary subunit of hyperpolarization-activated cyclic-nucleotide gated channels. **14**: 110-122

Hensel A, Beck S, El Magraoui F, Platta HW, Girzalsky W, & Erdmann R (2011) Cysteine-dependent ubiquitination of Pex18p is linked to cargo translocation across the peroxisomal membrane. *J Biol Chem* **286**: 43495-43505

Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, & Vinh LS (2018) UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* **35**: 518-522

Hua R & Kim PK (2016) Multiple paths to peroxisomes: mechanism of peroxisome maintenance in mammals. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1863**: 881-891

Islinger M, Voelkl A, Fahimi HD, & Schrader M (2018) The peroxisome: an update on mysteries 2.0. *Histochem Cell Biol* **150**: 443-471

Jansen R & van der Klei I (2019) The peroxisome biogenesis factors Pex3 and Pex19: multitasking proteins with disputed functions. *FEBS Lett* **593**: 457-474

Jeynov B, Lay D, Schmidt F, Tahirovic S, & Just WW (2006) Phosphoinositide synthesis and degradation in isolated rat liver peroxisomes. *FEBS Lett* **580**: 5917-5924

Joshi AS, Huang X, Choudhary V, Levine TP, Hu J, & Prinz WA (2016) A family of membrane-shaping proteins at ER subdomains regulates pre-peroxisomal vesicle biogenesis. *J Cell Biol* **215**: 515-529

Joshi AS, Nebenfuehr B, Choudhary V, Satpute-Krishnan P, Levine TP, Golden A, & Prinz WA (2018) Lipid droplet and peroxisome biogenesis occur at the same ER subdomains. *Nature communications* **9**: 1-12

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

Kalyaanamoorthy S, Minh BQ, Wong TK, von Haeseler A, & Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods* **14**: 587-589

Katoh K & Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**: 772-780

Kiel JA, Veenhuis M, & van der Klei, Ida J (2006) PEX genes in fungal genomes: common, rare or redundant. *Traffic* **7**: 1291-1303

Klionsky DJ, Cregg JM, Dunn WA,Jr, Emr SD, Sakai Y, Sandoval IV, Sibirny A, Subramani S, Thumm M, Veenhuis M, & Ohsumi Y (2003) A unified nomenclature for yeast autophagy-related genes. *Dev Cell* **5**: 539-545

Koch J & Brocard C (2012) PEX11 proteins attract Mff and human Fis1 to coordinate peroxisomal fission. *J Cell Sci* **125**: 3813-3826

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

Krikken AM, Veenhuis M, & Van Der Klei, Ida J (2009) Hansenula polymorpha pex11 cells are affected in peroxisome retention. *The FEBS journal* **276**: 1429-1439

Kunze M (2020) The type-2 peroxisomal targeting signal. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1867**: 118609

Le T, Zarsky V, Nyvltova E, Rada P, Harant K, Vancova M, Verner Z, Hrdy I, & Tachezy J (2020) Anaerobic peroxisomes in Mastigamoeba balamuthi. *Proc Natl Acad Sci U S A* **117**: 2065-2075

Leon S & Subramani S (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

Letunic I & Bork P (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**: W256-W259

Li X & Gould SJ (2003) The dynamin-like GTPase DLP1 is essential for peroxisome division and is recruited to peroxisomes in part by PEX11. *J Biol Chem* **278**: 17012-17020

Li X & Gould SJ (2002) PEX11 promotes peroxisome division independently of peroxisome metabolism. *J Cell Biol* **156**: 643-651

Lill P, Hansen T, Wendscheck D, Klink BU, Jeziorek T, Miehling J, Bender J, Drepper F, Girzalsky W, & Erdmann R (2020) Towards the molecular architecture of the peroxisomal receptor docking complex. *bioRxiv*: 854497

Lingard MJ, Gidda SK, Bingham S, Rothstein SJ, Mullen RT, & Trelease RN (2008) Arabidopsis PEROXIN11c-e, FISSION1b, and DYNAMIN-RELATED PROTEIN3A cooperate in cell cycle-associated replication of peroxisomes. *Plant Cell* **20**: 1567-1585

Ma C, Hagstrom D, Polley SG, & Subramani S (2013) Redox-regulated cargo binding and release by the peroxisomal targeting signal receptor, Pex5. *J Biol Chem* **288**: 27220-27231

Mast FD, Jamakhandi A, Saleem RA, Dilworth DJ, Rogers RS, Rachubinski RA, & Aitchison JD (2016) Peroxins Pex30 and Pex29 Dynamically Associate with Reticulons to Regulate Peroxisome Biogenesis from the Endoplasmic Reticulum. *J Biol Chem* **291**: 15408-15427

Mast FD, Li J, Virk MK, Hughes SC, Simmonds AJ, & Rachubinski RA (2011) A Drosophila model for the Zellweger spectrum of peroxisome biogenesis disorders. *Dis Model Mech* **4**: 659-672

Matsumoto N, Tamura S, & Fujiki Y (2003) The pathogenic peroxin Pex26p recruits the Pex1p–Pex6p AAA ATPase complexes to peroxisomes. *Nat Cell Biol* **5**: 454-460

Meinecke M, Cizmowski C, Schliebs W, Krüger V, Beck S, Wagner R, & Erdmann R (2010) The peroxisomal importomer constitutes a large and highly dynamic pore. *Nat Cell Biol* **12**: 273-277

Mészáros B, Erdős G, & Dosztányi Z (2018) IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res* **46**: W329-W337

Mindthoff S, Grunau S, Steinfort LL, Girzalsky W, Hiltunen JK, Erdmann R, & Antonenkov VD (2016) Peroxisomal Pex11 is a pore-forming protein homologous to TRPM channels. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1863**: 271-283

Mix A, Cenci U, Heimerl T, Marter P, Wirkner M, & Moog D (2018) Identification and Localization of Peroxisomal Biogenesis Proteins Indicates the Presence of Peroxisomes in the Cryptophyte Guillardia theta and Other "Chromalveolates". *Genome biology and evolution* **10**: 2834-2852

Montilla-Martinez M, Beck S, Klümper J, Meinecke M, Schliebs W, Wagner R, & Erdmann R (2015) Distinct pores for peroxisomal import of PTS1 and PTS2 proteins. *Cell reports* **13**: 2126-2134

Motley AM, Hettema EH, Ketting R, Plasterk R, & Tabak HF (2000) Caenorhabditis elegans has a single pathway to target matrix proteins to peroxisomes. *EMBO Rep* **1**: 40-46

Nguyen L, Schmidt HA, Von Haeseler A, & Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**: 268-274

Ogawa M, Yoshikawa Y, Kobayashi T, Mimuro H, Fukumatsu M, Kiga K, Piao Z, Ashida H, Yoshida M, & Kakuta S (2011) A Tecpr1-dependent selective autophagy pathway targets bacterial pathogens. *Cell host & microbe* **9**: 376-389

Okumoto K, Misono S, Miyata N, Matsumoto Y, Mukai S, & Fujiki Y (2011) Cysteine ubiquitination of PTS1 receptor Pex5p regulates Pex5p recycling. *Traffic* **12**: 1067-1083

Olivares AO, Baker TA, & Sauer RT (2016) Mechanistic insights into bacterial AAA proteases and protein-remodelling machines. *Nature Reviews Microbiology* **14**: 33

Opaliński Ł, Bartoszewska M, Fekken S, Liu H, De Boer R, Van Der Klei I, Veenhuis M, & Kiel JA (2012) De novo peroxisome biogenesis in Penicillium chrysogenum is not dependent on the Pex11 family members or Pex16. *PloS one* **7**: e35490

Opaliński Ł, Kiel JA, Williams C, Veenhuis M, & Van Der Klei, Ida J (2011) Membrane curvature during peroxisome fission requires Pex11. *EMBO J* **30**: 5-16

Otera H, Setoguchi K, Hamasaki M, Kumashiro T, Shimizu N, & Fujiki Y (2002) Peroxisomal targeting signal receptor Pex5p interacts with cargoes and import machinery components in a spatiotemporally differentiated manner: conserved Pex5p WXXXF/Y motifs are critical for matrix protein import. *Mol Cell Biol* **22**: 1639-1655

Pan D, Nakatsu T, & Kato H (2013) Crystal structure of peroxisomal targeting signal-2 bound to its receptor complex Pex7p–Pex21p. *Nature structural & molecular biology* **20**: 987

Pedrosa AG, Francisco T, Bicho D, Dias AF, Barros-Barbosa A, Hagmann V, Dodt G, Rodrigues TA, & Azevedo JE (2018) Peroxisomal monoubiquitinated PEX5 interacts with the AAA ATPases PEX1 and PEX6 and is unfolded during its dislocation into the cytosol. *J Biol Chem* **293**: 11553-11563

Pfanner N, van der Laan M, Amati P, Capaldi RA, Caudy AA, Chacinska A, Darshi M, Deckers M, Hoppins S, & Icho T (2014) Uniform nomenclature for the mitochondrial contact site and cristae organizing system. *J Cell Biol* **204**: 1083-1086

Platta HW, Hagen S, Reidick C, & Erdmann R (2014) The peroxisomal receptor dislocation pathway: to the exportomer and beyond. *Biochimie* **98**: 16-28

Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, & Finn RD (2018) HMMER web server: 2018 update. *Nucleic Acids Res* **46**: W200-W204

Purdue PE, Yang X, & Lazarow PB (1998) Pex18p and Pex21p, a novel pair of related peroxins essential for peroxisomal targeting by the PTS2 pathway. *J Cell Biol* **143**: 1859-1869

Rucktaschel R, Thoms S, Sidorovitch V, Halbach A, Pechlivanis M, Volkmer R, Alexandrov K, Kuhlmann J, Rottensteiner H, & Erdmann R (2009) Farnesylation of pex19p is required for its structural integrity and function in peroxisome biogenesis. *J Biol Chem* **284**: 20885-20896

Rymer L, Kempinski B, Chelstowska A, & Skoneczny M (2018) The budding yeast Pex5p receptor directs Fox2p and Cta1p into peroxisomes via its N-terminal region near the FxxxW domain. *J Cell Sci* **131**: 10.1242/jcs.216986

Saidowsky J, Dodt G, Kirchberg K, Wegner A, Nastainczyk W, Kunau WH, & Schliebs W (2001) The di-aromatic pentapeptide repeats of the human peroxisome import receptor PEX5 are separate high affinity binding sites for the peroxisomal membrane protein PEX14. *J Biol Chem* **276**: 34524-34529

Schliebs W & Kunau W (2006) PTS2 co-receptors: diverse proteins with common features. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1763**: 1605-1612

Schliebs W, Saidowsky J, Agianian B, Dodt G, Herberg FW, & Kunau WH (1999) Recombinant human peroxisomal targeting signal receptor PEX5. Structural basis for interaction of PEX5 with PEX14. *J Biol Chem* **274**: 5666-5673

Schrader M, Costello JL, Godinho LF, Azadi AS, & Islinger M (2016) Proliferation and fission of peroxisomes—an update. *Biochimica Et Biophysica Acta (BBA)-Molecular Cell Research* **1863**: 971-983

Schrader M, Reuber BE, Morrell JC, Jimenez-Sanchez G, Obie C, Stroh TA, Valle D, Schroer TA, & Gould SJ (1998) Expression of PEX11β mediates peroxisome proliferation in the absence of extracellular stimuli. *J Biol Chem* **273**: 29607-29614

Schrul B & Kopito RR (2016) Peroxin-dependent targeting of a lipid-droplet-destined membrane protein to ER subdomains. *Nat Cell Biol* **18**: 740-751

Shai N, Yifrach E, van Roermund CW, Cohen N, Bibi C, IJlst L, Cavellini L, Meurisse J, Schuster R, & Zada L (2018) Systematic mapping of contact sites reveals tethers and a function for the peroxisome-mitochondria contact. *Nature communications* **9**: 1-13

Singh R, Manivannan S, Krikken AM, de Boer R, Bordin N, Devos DP, & van der Klei, Ida J (2020) Hansenula polymorpha Pex37 is a peroxisomal membrane protein required for organelle fission and segregation. *The FEBS Journal* **287**: 1742-1757

Smith JJ & Aitchison JD (2013) Peroxisomes take shape. *Nature reviews Molecular cell biology* **14**: 803-817

Tam YYC & Rachubinski RA (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

Ušaj MM, Brložnik M, Kaferle P, Žitnik M, Wolinski H, Leitner F, Kohlwein S, Zupan B, & Petrovič U (2015) Genome-wide localization study of yeast Pex11 identifies peroxisome–mitochondria interactions through the ERMES complex. *J Mol Biol* **427**: 2072-2087

Van Der Zand A, Gent J, Braakman I, & Tabak HF (2012) Biochemically distinct vesicles from the endoplasmic reticulum fuse to form peroxisomes. *Cell* **149**: 397-409

van Roermund CW, Tabak HF, van den Berg M, Wanders RJ, & Hettema EH (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**: 489-498

van Zutphen T, Baerends RJ, Susanna KA, De Jong A, Kuipers OP, Veenhuis M, & Van der Klei, Ida J (2010) Adaptation of Hansenula polymorpha to methanol: a transcriptome analysis. *BMC Genomics* **11**: 1

Vastiau I, Anthonio E, Brams M, Brees C, Young S, Van de Velde S, Wanders R, Mannaerts G, Baes M, & Van Veldhoven PP (2006) Farnesylation of Pex19p is not essential for peroxisome biogenesis in yeast and mammalian cells. *Cellular and Molecular Life Sciences CMLS* **63**: 1686-1699

Vizeacoumar FJ, Torres-Guzman JC, Bouard D, Aitchison JD, & Rachubinski RA (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 FJ, Torres-Guzman JC, Tam YYC, Aitchison JD, & Rachubinski RA (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, van Hellemond JJ, Pfisterer I, Maier A, Hillmer S, & Clayton C (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

Wang D, Visser NV, Veenhuis M, & van der Klei IJ (2003) Physical interactions of the peroxisomal targeting signal 1 receptor pex5p, studied by fluorescence correlation spectroscopy. *J Biol Chem* **278**: 43340-43345

Wang S, Idrissi F, Hermansson M, Grippa A, Ejsing CS, & Carvalho P (2018) Seipin and the membrane-shaping protein Pex30 cooperate in organelle budding from the endoplasmic reticulum. *Nature communications* **9**: 1-12

Weisman CM, Murray AW, & Eddy SR (2020) Many but not all lineage-specific genes can be explained by homology detection failure. *BioRxiv*

Williams C, Opalinski L, Landgraf C, Costello J, Schrader M, Krikken AM, Knoops K, Kram AM, Volkmer R, & van der Klei, Ida J (2015) The membrane remodeling protein Pex11p activates the GTPase Dnm1p during peroxisomal fission. *Proceedings of the National Academy of Sciences* **112**: 6377-6382

Wu F, de Boer R, Krikken AM, Akşit A, Bordin N, Devos DP, & van der Klei, Ida J (2020) Pex24 and Pex32 are required to tether peroxisomes to the ER for organelle biogenesis, positioning and segregation in yeast. *J Cell Sci* **133**

Yan M, Rachubinski DA, Joshi S, Rachubinski RA, & Subramani S (2008) Dysferlin domain-containing proteins, Pex30p and Pex31p, localized to two compartments, control the number and size of oleate-induced peroxisomes in Pichia pastoris. *Mol Biol Cell* **19**: 885-898

Yifrach E, Chuartzman SG, Dahan N, Maskit S, Zada L, Weill U, Yofe I, Olender T, Schuldiner M, & Zalckvar E (2016) Characterization of proteome dynamics during growth in oleate reveals a new peroxisome-targeting receptor. *J Cell Sci* **129**: 4067-4075

Yofe I, Soliman K, Chuartzman SG, Morgan B, Weill U, Yifrach E, Dick TP, Cooper SJ, Ejsing CS, Schuldiner M, Zalckvar E, & Thoms S (2017) Pex35 is a regulator of peroxisome abundance. *J Cell Sci* **130**: 791-804

Yoshida Y, Niwa H, Honsho M, Itoyama A, & Fujiki Y (2015) Pex11mediates peroxisomal proliferation by promoting deformation of the lipid membrane. *Biology Open* **4**: 710-721

Žárský V & Tachezy J (2015) Evolutionary loss of peroxisomes–not limited to parasites. *Biology direct* **10**: 1-10

Supplementary information

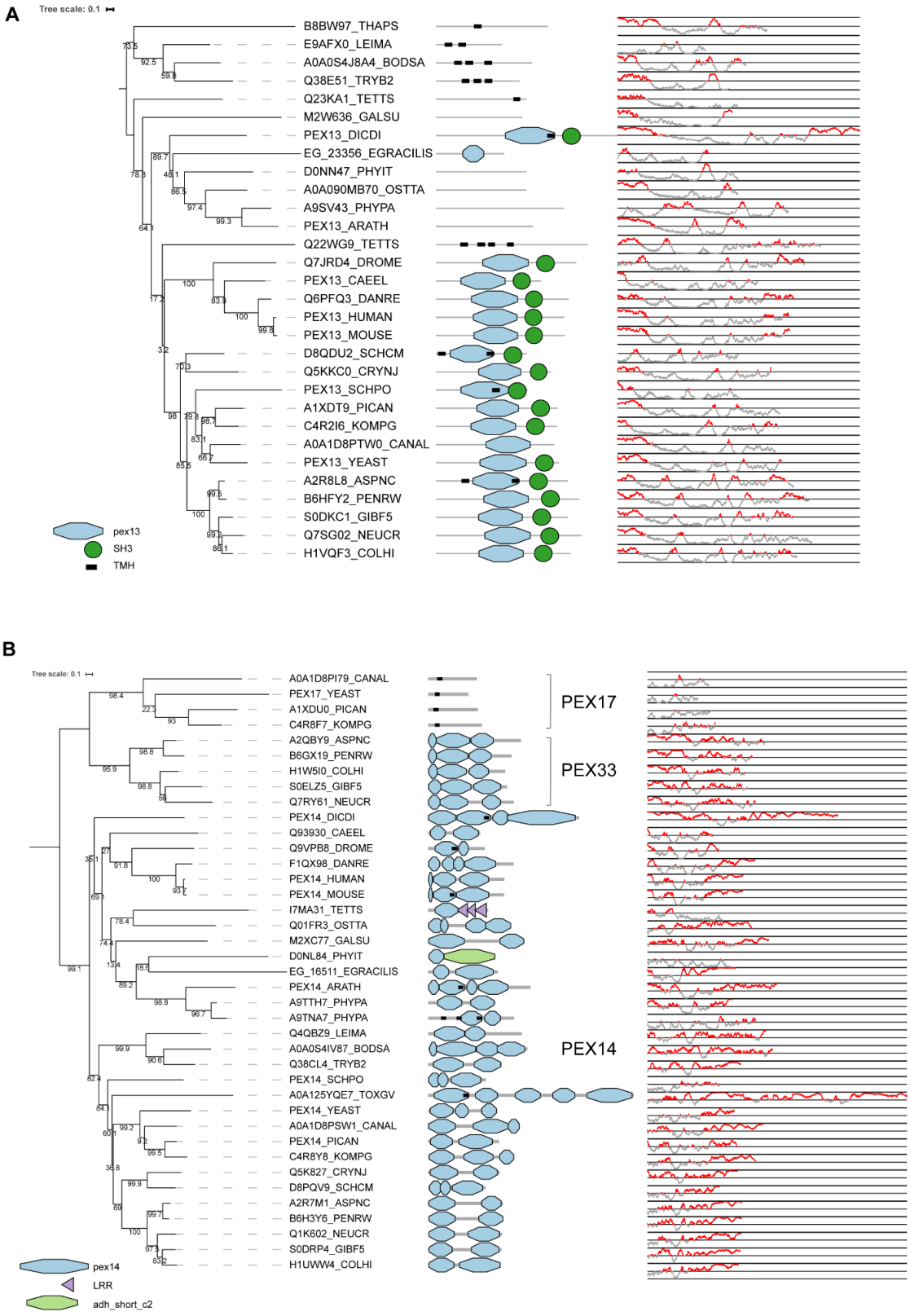

Figure S1: Phylogeny and protein features of A) PEX13 and B) PEX14/17/33 orthologs.

The phylogeny is rooted at mid-point to ease the visualization. Note that the topology does not necessarily reflect the actual evolutionary trajectory of such proteins. Protein domain architecture is defined by pfam annotations and transmembrane helix according to TMHMM software. The line-dot plot, indicates the regions predicted to be disordered (red) and not disordered (grey).

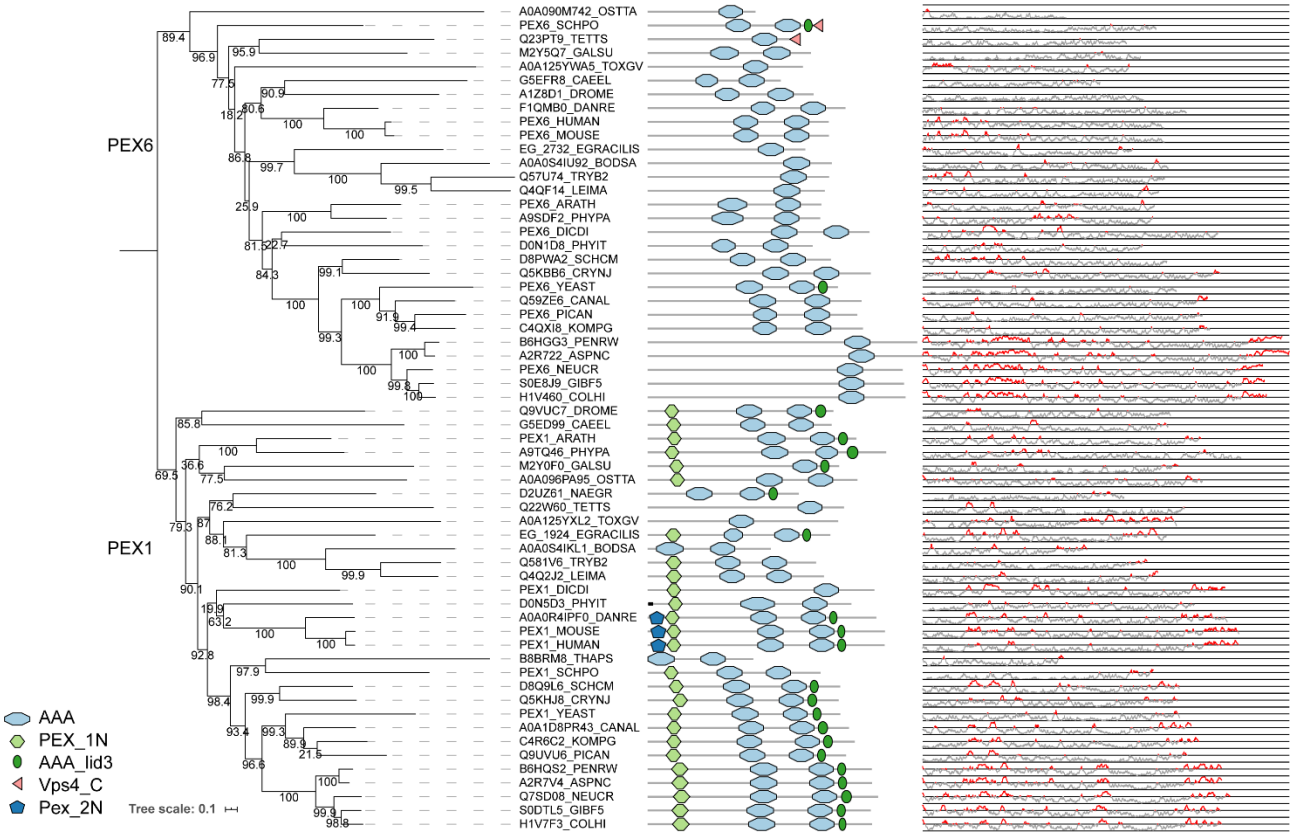

Figure S2: Phylogeny and protein features of PEX1/6 orthologs.

The phylogeny is rooted at mid-point to ease the visualization. Note that the topology does not necessarily reflect the actual evolutionary trajectory of such proteins. Protein domain architecture is defined by pfam annotations and transmembrane helix according to TMHMM software. The line-dot plot, indicates the regions predicted to be disordered (red) and not disordered (grey).

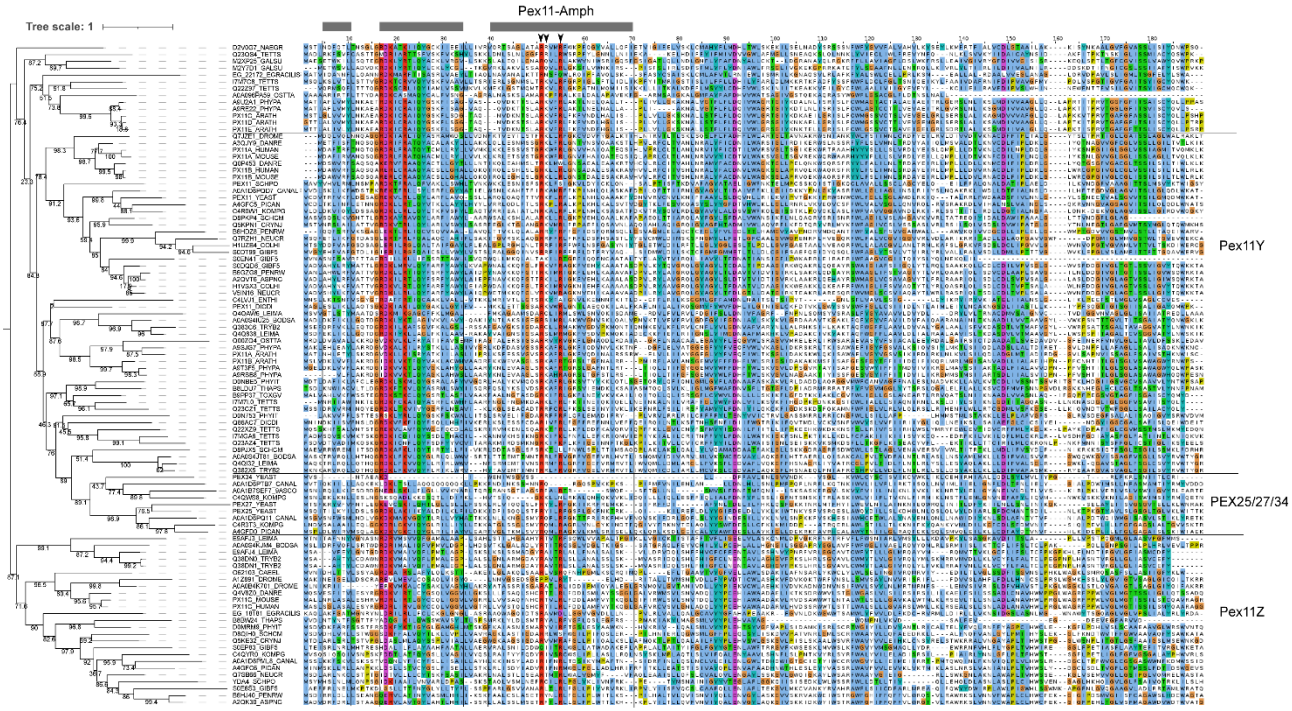

Figure S3: Multiple sequence alignment of Pex11 family proteins.

Grey bars above sequences denote predicted α-helices and the N-terminal amphipathic helix (Pex11-Amph). Residues are coloured based on physico-chemical properties according to ClustalW. The phylogeny is rooted at mid-point to ease the visualization. Note that the topology does not necessarily reflect the actual evolutionary trajectory of such proteins.

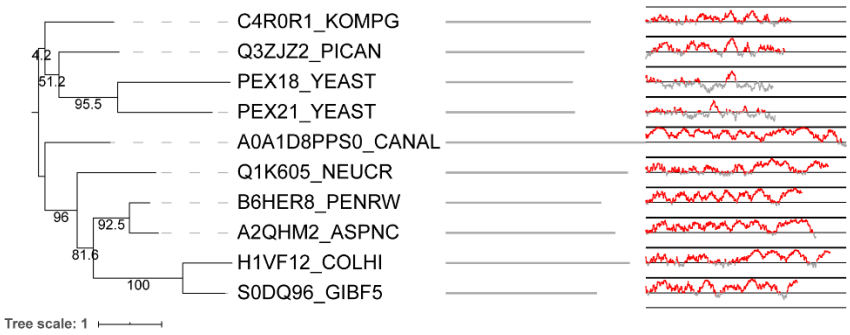

Figure S4: Phylogeny and protein features of PEX118/20/21 orthologs.

The phylogeny is rooted at mid-point to ease the visualization and labels of the main taxonomic groups are coloured accordingly to the legend. Note that the topology does not necessarily reflect the actual evolutionary trajectory of such proteins. Protein domain architecture is defined by pfam annotations and transmembrane helix according to TMHMM software. The line-dot plot, indicates the regions predicted to be disordered (red) and not disordered (grey).

Table S1: Protein codes for all PEX orthologs (Excel file)

Table 1: Protein codes for all PEX orthologs

	# PEX1	Uniprot code PEX1
EGRACILIS	1	EG_transcript_1924
TRYB2	1	Q581V6|Q581V6_TRYB2
LEIMA	1	Q4Q2J2|Q4Q2J2_LEIMA
BODSA	1	A0A0S4IKL1|A0A0S4IKL1_BODSA
PHYIT	1	D0N5D3|D0N5D3_PHYIT
TOXGV	1	A0A125YXL2|A0A125YXL2_TOXGV
TETTS	1	Q22W60|Q22W60_TETTS
NAEGR	1	D2UZ61|D2UZ61_NAEGR
THAPS	1	B8BRM8|B8BRM8_THAPS
GALSU	1	M2Y0F0|M2Y0F0_GALSU
OSTTA	1	A0A096PA95|A0A096PA95_OSTTA
PHYPA	1	A9TQ46|A9TQ46_PHYPA
ARATH	1	Q9FNP1|PEX1_ARATH
YEAST	1	P24004|PEX1_YEAST
KOMPG	1	C4R6C2|C4R6C2_KOMPG
CANAL	1	A0A1D8PR43|A0A1D8PR43_CANAL
PICAN	1	Q9UVU6|Q9UVU6_PICAN
PENRW	1	B6HQS2|B6HQS2_PENRW
ASPNC	1	A2R7V4|A2R7V4_ASPNC
COLHI	1	H1V7F3|H1V7F3_COLHI
GIBF5	2	S0DTL5|S0DTL5_GIBF5; S0E826|S0E826_GIBF5
NEUCR	1	Q7SD08|Q7SD08_NEUCR
SCHPO	1	O74941|PEX1_SCHPO
SCHCM	1	D8Q9L6|D8Q9L6_SCHCM
CRYNJ	1	Q5KHJ8|Q5KHJ8_CRYNJ
DICDI	1	Q54GX5|PEX1_DICDI
ENTHI	0	
CAEEL	1	G5ED99|G5ED99_CAEEL
DROME	1	Q9VUC7|Q9VUC7_DROME
DANRE	1	A0A0R4IPF0|A0A0R4IPF0_DANRE
MOUSE	1	Q5BL07|PEX1_MOUSE
HUMAN	1	O43933|PEX1_HUMAN

	# PEX2	Uniprot code PEX2
EGRACILIS	1	EG_transcript_15991
TRYB2	1	Q582U7|Q582U7_TRYB2
LEIMA	1	Q4Q9L4|Q4Q9L4_LEIMA
BODSA	1	A0A0S4J1C5|A0A0S4J1C5_BODSA
PHYIT	1	D0N3E7|D0N3E7_PHYIT
TOXGV	1	B6KB52|B6KB52_TOXGV
TETTS	1	Q23A31|Q23A31_TETTS
NAEGR	0	
THAPS	0	
GALSU	1	M2XGA5|M2XGA5_GALSU
OSTTA	1	Q00WM0|Q00WM0_OSTTA
PHYPA	1	A9TT97|A9TT97_PHYPA
ARATH	1	Q9CA86|PEX2_ARATH
YEAST	1	P32800|PEX2_YEAST
KOMPG	1	C4R3D4|C4R3D4_KOMPG
CANAL	1	Q59ZH3|Q59ZH3_CANAL
PICAN	1	Q68HK6_PICAN
PENRW	1	B6HCX8|B6HCX8_PENRW

ASPNC	1	A2RAS1|A2RAS1_ASPNC
COLHI	1	H1VKV6|H1VKV6_COLHI
GIBF5	1	S0DJY5|S0DJY5_GIBF5
NEUCR	1	Q7SEZ3|Q7SEZ3_NEUCR
SCHPO	1	O42845|YFH7_SCHPO
SCHCM	1	D8PPH5|D8PPH5_SCHCM
CRYNJ	1	Q5KPT6|Q5KPT6_CRYNJ
DICDI	1	Q75JQ3|PEX2_DICDI
ENTHI	0	
CAEEL	1	Q23601|Q23601_CAEEL
DROME	1	Q9VSH8|Q9VSH8_DROME
DANRE	1	E7F4V8|E7F4V8_DANRE
MOUSE	1	P55098|PEX2_MOUSE
HUMAN	1	P28328|PEX2_HUMAN

	#	Uniprot code
	PEX3	**PEX3**
EGRACILIS	1	EG_transcript_17796
TRYB2	1	Q383Q3|Q383Q3_TRYB2
LEIMA	1	Q4Q114|Q4Q114_LEIMA
BODSA	1	A0A0S4J1D1|A0A0S4J1D1_BODSA
PHYIT	1	D0MT01|D0MT01_PHYIT
TOXGV	1	B6KA01|B6KA01_TOXGV
TETTS	0	
NAEGR	1	D2VHW1|D2VHW1_NAEGR
THAPS	1	B8C9N1|B8C9N1_THAPS
GALSU	1	M2Y9B0|M2Y9B0_GALSU
OSTTA	1	Q00UU8|Q00UU8_OSTTA
PHYPA	1	A9SSS9|A9SSS9_PHYPA
ARATH	2	Q8LDG7|PEX31_ARATH; Q8S9K7|PEX32_ARATH
YEAST	1	P28795|PEX3_YEAST
KOMPG	1	C4R6F0|C4R6F0_KOMPG
CANAL	1	A0A1D8PE25|A0A1D8PE25_CANAL
PICAN	1	Q01497|PEX3_PICAN
PENRW	1	B6HHY4|B6HHY4_PENRW
ASPNC	1	E2PSX2|E2PSX2_ASPNC
COLHI	1	H1V4L9|H1V4L9_COLHI
GIBF5	1	S0DU03|S0DU03_GIBF5
NEUCR	1	Q7SBJ5|Q7SBJ5_NEUCR
SCHPO	1	O14017|YDPE_SCHPO
SCHCM	1	D8Q669|D8Q669_SCHCM
CRYNJ	1	Q5KJS9|Q5KJS9_CRYNJ
DICDI	1	Q54U86|PEX3_DICDI
ENTHI	0	
CAEEL	1	Q18028|Q18028_CAEEL
DROME	1	Q9VUL8|Q9VUL8_DROME
DANRE	1	Q5RIV3|Q5RIV3_DANRE
MOUSE	1	Q9QXY9|PEX3_MOUSE
HUMAN	1	P56589|PEX3_HUMAN

	#	Uniprot code
	PEX4	**PEX4**
EGRACILIS	1	EG_transcript_37730
TRYB2	1	Q57YD9|Q57YD9_TRYB2
LEIMA	1	Q4QIK2|Q4QIK2_LEIMA
BODSA	1	A0A0S4KIJ1|A0A0S4KIJ1_BODSA

PHYIT	1	D0N6V8|D0N6V8_PHYIT
TOXGV	1	B9QPM5|B9QPM5_TOXGV
TETTS	1	W7XEV7|W7XEV7_TETTS
NAEGR	1	D2V3A8|D2V3A8_NAEGR
THAPS	1	B8C5Q5|B8C5Q5_THAPS
GALSU	1	M2XTF8|M2XTF8_GALSU
OSTTA	0	
PHYPA	1	A9T3L8|A9T3L8_PHYPA
ARATH	1	Q8LGF7|PEX4_ARATH
YEAST	1	P29340|UBCX_YEAST
KOMPG	1	C4R826|C4R826_KOMPG
CANAL	1	A0A1D8PP96|A0A1D8PP96_CANAL
PICAN	1	O60015|UBCX_PICAN
PENRW	1	B6H6P1|B6H6P1_PENRW
ASPNC	1	A0A100IUN7_ASPNG
COLHI	1	H1UZJ1|H1UZJ1_COLHI
GIBF5	1	S0DX94|S0DX94_GIBF5
NEUCR	1	Q7SDB0|Q7SDB0_NEUCR
SCHPO	1	Q9P6I1|UBC16_SCHPO
SCHCM	0	
CRYNJ	0	
DICDI	1	Q86IZ3|UBCX_DICDI
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX5	
EGRACILIS	1	EG_transcript_5901
TRYB2	1	Q57W55|Q57W55_TRYB2
LEIMA	1	E9AEW5|E9AEW5_LEIMA
BODSA	1	A0A0S4KFV0|A0A0S4KFV0_BODSA
PHYIT	1	D0NJ23|D0NJ23_PHYIT
TOXGV	1	A0A0F7V279|A0A0F7V279_TOXGV
TETTS	1	Q23AQ0|Q23AQ0_TETTS
NAEGR	1	D2V581|D2V581_NAEGR
THAPS	1	>jgi|Thaps3|10623|fgenesh1_pg.C_chr_17000086
GALSU	1	M2Y760|M2Y760_GALSU
OSTTA	1	A0A090N3E8|A0A090N3E8_OSTTA
PHYPA	1	A9T1E0|A9T1E0_PHYPA
ARATH	1	Q9FMA3|PEX5_ARATH
YEAST	2	P35056|PEX5_YEAST; Q04364|YMP8_YEAST
KOMPG	1	C4R2L0|C4R2L0_KOMPG
CANAL	1	O74711|PEX5_CANAL
PICAN	1	Q01495|PEX5_PICAN
PENRW	1	B6HG52|B6HG52_PENRW
ASPNC	1	A2R8K6|A2R8K6_ASPNC
COLHI	1	H1VHA1|H1VHA1_COLHI
GIBF5	1	S0DNH6|S0DNH6_GIBF5
NEUCR	1	Q7SH09|Q7SH09_NEUCR
SCHPO	1	O94325|PEX5_SCHPO
SCHCM	1	D8Q932|D8Q932_SCHCM
CRYNJ	2	Q5KMB5|Q5KMB5_CRYNJ; Q5K9A1|Q5K9A1_CRYNJ

DICDI	1	Q54MD1|PEX5_DICDI
ENTHI	1	C4LUW6|C4LUW6_ENTHI
CAEEL	1	Q18426|Q18426_CAEEL
DROME	1	O46085|O46085_DROME
		E7FGF7|E7FGF7_DANRE;
		A0A0R4IXK1|A0A0R4IXK1_DANRE;
DANRE	3	E7F507|E7F507_DANRE
MOUSE	2	O09012|PEX5_MOUSE; Q8C437|PEX5R_MOUSE
HUMAN	2	P50542|PEX5_HUMAN; Q8IYB4|PEX5R_HUMAN

	#	Uniprot code
	PEX6	**PEX6**
EGRACILIS	1	EG_transcript_2732
TRYB2	1	Q57U74|Q57U74_TRYB2
LEIMA	1	Q4QF14|Q4QF14_LEIMA
BODSA	1	A0A0S4IU92|A0A0S4IU92_BODSA
PHYIT	1	D0N1D8|D0N1D8_PHYIT
TOXGV	1	A0A125YWA5|A0A125YWA5_TOXGV
TETTS	1	Q23PT9|Q23PT9_TETTS
NAEGR	0	
THAPS	0	
GALSU	1	M2Y5Q7|M2Y5Q7_GALSU
OSTTA	1	A0A090M742|A0A090M742_OSTTA
PHYPA	1	A9SDF2|A9SDF2_PHYPA
ARATH	1	Q8RY16|PEX6_ARATH
YEAST	1	P33760|PEX6_YEAST
KOMPG	1	C4QXI8|C4QXI8_KOMPG
CANAL	1	Q59ZE6|Q59ZE6_CANAL
PICAN	1	Q9UVU5|PEX6_PICAN
PENRW	1	B6HGG3|B6HGG3_PENRW
ASPNC	1	A2R722|A2R722_ASPNC
COLHI	1	H1V460|H1V460_COLHI
GIBF5	1	S0E8J9|S0E8J9_GIBF5
NEUCR	1	Q7SGP2|PEX6_NEUCR
SCHPO	1	O13764|PEX6_SCHPO
SCHCM	1	D8PWA2|D8PWA2_SCHCM
CRYNJ	1	Q5KBB6|Q5KBB6_CRYNJ
DICDI	1	Q54CS8|PEX6_DICDI
ENTHI	0	
CAEEL	1	G5EFR8|G5EFR8_CAEEL
DROME	1	A1Z8D1|A1Z8D1_DROME
DANRE	1	F1QMB0|F1QMB0_DANRE
MOUSE	1	Q99LC9|PEX6_MOUSE
HUMAN	1	Q13608|PEX6_HUMAN

	#	Uniprot code
	PEX7	**PEX7**
EGRACILIS	1	EG_transcript_18111
TRYB2	1	Q57W97|Q57W97_TRYB2
LEIMA	1	Q4VQ66|Q4VQ66_LEIMA
BODSA	1	A0A0S4JHX6|A0A0S4JHX6_BODSA
PHYIT	1	D0NPH8|D0NPH8_PHYIT
TOXGV	1	V4ZGR8|V4ZGR8_TOXGV
TETTS	1	I7M6F2|I7M6F2_TETTS
NAEGR	1	D2VFM6|D2VFM6_NAEGR

THAPS	0	
GALSU	0	
OSTTA	1	A0A090M5G9|A0A090M5G9_OSTTA
PHYPA	2	A9SVF9|A9SVF9_PHYPA; A9TZX0|A9TZX0_PHYPA
ARATH	1	Q9XF57|PEX7_ARATH
YEAST	1	P39108|PEX7_YEAST
KOMPG	1	C4QYG0|C4QYG0_KOMPG
CANAL	1	Q59WW3|Q59WW3_CANAL
PICAN	1	Q3HL98|Q3HL98_PICAN
PENRW	1	B6GWD3|B6GWD3_PENRW
ASPNC	1	A2QB31|A2QB31_ASPNC
COLHI	1	H1VIU5|H1VIU5_COLHI
GIBF5	1	S0DTB6|S0DTB6_GIBF5
NEUCR	1	Q1K8F8|Q1K8F8_NEUCR
SCHPO	1	P78798|PEX7_SCHPO
SCHCM	1	D8Q1M9|D8Q1M9_SCHCM
CRYNJ	1	Q5KNX4|Q5KNX4_CRYNJ
DICDI	1	Q54WA3|PEX7_DICDI
ENTHI	0	
CAEEL	0	
DROME	1	Q9VSN7|Q9VSN7_DROM
DANRE	1	F1QFW8|F1QFW8_DANRE
MOUSE	1	P97865|PEX7_MOUSE
HUMAN	1	O00628|PEX7_HUMAN

	#	Uniprot code
	PEX8	
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	P53248|PEX8_YEAST
KOMPG	1	C4QY68|C4QY68_KOMPG
CANAL	1	A0A1D8PJT4|A0A1D8PJT4_CANAL
PICAN	1	Q00925|PEX8_PICAN
PENRW	1	B6H292|B6H292_PENRW
ASPNC	1	A2QZ51|A2QZ51_ASPNC
COLHI	1	H1V9M3|H1V9M3_COLHI
GIBF5	1	S0DUK6|S0DUK6_GIBF5
NEUCR	1	Q7RY73|Q7RY73_NEUCR
SCHPO	1	Q9UT96|YL43_SCHPO
SCHCM	1	D8Q091|D8Q091_SCHCM
CRYNJ	1	Q5KH17|Q5KH17_CRYNJ
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	

MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX10	**PEX10**
EGRACILIS	1	EG_transcript_17328
TRYB2	1	Q582V4|Q582V4_TRYB2
LEIMA	1	Q95ZB8|Q95ZB8_LEIMA
BODSA	1	A0A0S4J0W7|A0A0S4J0W7_BODSA
PHYIT	1	D0NZ59|D0NZ59_PHYIT
TOXGV	0	
TETTS	1	I7LX78|I7LX78_TETTS
NAEGR	1	D2VA32|D2VA32_NAEGR
THAPS	0	
GALSU	1	M2XUD8|M2XUD8_GALSU
OSTTA	1	A0A096PAL2|A0A096PAL2_OSTTA
PHYPA	1	A9TG49|A9TG49_PHYPA
ARATH	1	Q9SYU4|PEX10_ARATH
YEAST	1	Q05568|PEX10_YEAST
KOMPG	1	A0A1G4KPJ1|A0A1G4KPJ1_KOMPC
CANAL	1	A0A1D8PL05|A0A1D8PL05_CANAL
PICAN	1	Q00940|PEX10_PICAN
PENRW	1	B6HMM9|B6HMM9_PENRW
ASPNC	1	A2R1D2|A2R1D2_ASPNC
COLHI	1	H1VDY2|H1VDY2_COLHI
GIBF5	1	S0DTP0|S0DTP0_GIBF5
NEUCR	1	Q7SDX8|Q7SDX8_NEUCR
SCHPO	1	Q9UUF0|PEX10_SCHPO
SCHCM	1	D8PL48|D8PL48_SCHCM
CRYNJ	1	Q5KCS9|Q5KCS9_CRYNJ
DICDI	1	Q54S31|PEX10_DICDI
ENTHI	0	
CAEEL	1	C0HKD7|PEX10_CAEEL
DROME	1	Q9W0D7|Q9W0D7_DROME
DANRE	1	Q5XJ92|Q5XJ92_DANRE
MOUSE	1	B1AUE5|PEX10_MOUSE
HUMAN	1	O60683|PEX10_HUMAN

	#	Uniprot code
	PEX11	**PEX11**
EGRACILIS	2	EG_transcript_22172; EG_transcript_18181
TRYB2	4	Q383C6|Q383C6_TRYB2; Q382X5|Q382X5_TRYB2; Q38DN0|Q38DN0_TRYB2; Q38DN1|Q38DN1_TRYB2
LEIMA	5	Q4Q838|Q4Q838_LEIMA; Q4QAW6|Q4QAW6_LEIMA; Q4QI32|Q4QI32_LEIMA; E9AFJ4|E9AFJ4_LEIMA; E9AFJ3|E9AFJ3_LEIMA
BODSA	3	A0A0S4IU25|A0A0S4IU25_BODSA; A0A0S4JT81|A0A0S4JT81_BODSA; A0A0S4KJM4|A0A0S4KJM4_BODSA
PHYIT	3	D0MRM9|D0MRM9_PHYIT; D0NBE5|D0NBE5_PHYIT; D0N753|D0N753_PHYIT
TOXGV	1	B9PP37|B9PP37_TOXGV

		Q23CZ1|Q23CZ1_TETTS; Q22Z97|Q22Z97_TETTS; Q23AZ4|Q23AZ4_TETTS; I7MGA6|I7MGA6_TETTS; I7M7L0|I7M7L0_TETTS; Q23QS4|Q23QS4_TETTS;
TETTS	8	Q22XZ9|Q22XZ9_TETTS; I7M7C8|I7M7C8_TETTS
NAEGR	1	D2V0G7|D2V0G7_NAEGR
THAPS	2	B8LDU7|B8LDU7_THAPS; B8BW24|B8BW24_THAPS
GALSU	2	M2Y7D1|M2Y7D1_GALSU; M2XP25|M2XP25_GALSU
		A0A096PA59|A0A096PA59_OSTTA;
OSTTA	2	Q00ZQ4|Q00ZQ4_OSTTA
		A9U2A1|A9U2A1_PHYPA; A9RE22|A9RE22_PHYPA; A9SJG7|A9SJG7_PHYPA; A9T3F5|A9T3F5_PHYPA;
PHYPA	5	A9RSB8|A9RSB8_PHYPA
		Q9FZF1|PX11A_ARATH; Q9STY0|PX11B_ARATH; Q9LQ73|PX11C_ARATH; O80845|PX11D_ARATH;
ARATH	5	Q84JW1|PX11E_ARATH
YEAST	1	Q12462|PEX11_YEAST
		C4QYR0|C4QYR0_KOMPG;
KOMPG	2	C4R0W1|C4R0W1_KOMPG
		A0A1D8PML8|A0A1D8PML8_CANAL;
CANAL	2	A0A1D8PQD7|A0A1D8PQD7_CANAL
PICAN	2	A4GFC5|A4GFC5_PICAN; A4GFC6|A4GFC6_PICAN
		B6GZG8|B6GZG8_PENRW;
PENRW	3	B6HDZ8|B6HDZ8_PENRW; B6HJ40|B6HJ40_PENRW
ASPNC	2	A2QVT6|A2QVT6_ASPNC; A2QK30|A2QK30_ASPNC
COLHI	2	H1UZB4|H1UZB4_COLHI; H1VSX3|H1VSX3_COLHI
		S0EP83|S0EP83_GIBF5; S0DT99|S0DT99_GIBF5; S0DQD8|S0DQD8_GIBF5; S0E6S3|S0E6S3_GIBF5;
GIBF5	5	S0EN41|S0EN41_GIBF5
		Q7SB66|Q7SB66_NEUCR; V5IN16|V5IN16_NEUCR;
NEUCR	3	Q7RZH1|Q7RZH1_NEUCR
SCHPO	2	Q10333|PEX11_SCHPO; Q10346|YDA4_SCHPO
SCHCM	2	D8QIH0|D8QIH0_SCHCM; D8PKP4|D8PKP4_SCHCM
CRYNJ	2	Q5KPN1|Q5KPN1_CRYNJ; Q5KE32|Q5KE32_CRYNJ
DICDI	2	Q54H86|PEX11_DICDI; Q86AC7|Q86AC7_DICDI
ENTHI	1	C4LVJ1|C4LVJ1_ENTHI
CAEEL	1	O62103|O62103_CAEEL
		Q7JZE1|Q7JZE1_DROME; A0A0B4K701_DROME;
DROME	3	A1Z891_DROME

DANRE	3	A3QJY9|A3QJY9_DANRE; Q0P453|Q0P453_DANRE; Q4V8Z0|Q4V8Z0_DANRE
MOUSE	3	Q9Z211|PX11A_MOUSE; Q9Z210|PX11B_MOUSE; Q6P6M5|PX11C_MOUSE
HUMAN	3	O75192|PX11A_HUMAN; O96011|PX11B_HUMAN; Q96HA9|PX11C_HUMAN

	#	Uniprot code
	PEX12	**PEX12**
EGRACILIS	1	EG_transcript_17807
TRYB2	1	Q387S9|Q387S9_TRYB2
LEIMA	1	Q4QD90|Q4QD90_LEIMA
BODSA	1	A0A0S4IWJ0|A0A0S4IWJ0_BODSA
PHYIT	0	
TOXGV	1	A0A125YMN7|A0A125YMN7_TOXGV
TETTS	1	Q23DU9|Q23DU9_TETTS
NAEGR	1	D2V791|D2V791_NAEGR
THAPS	1	>jgi|Thaps3|1084|fgenesh1_pg.C_chr_1000218
GALSU	1	M2XWD0|M2XWD0_GALSU
OSTTA	1	A0A096P9H8|A0A096P9H8_OSTTA
PHYPA	1	A9RSM9|A9RSM9_PHYPA
ARATH	1	Q9M841|PEX12_ARATH
YEAST	1	Q04370|PEX12_YEAST
KOMPG	1	C4R8U8|C4R8U8_KOMPG
CANAL	1	A0A1D8PGB2|A0A1D8PGB2_CANAL
PICAN	1	Q8NK59_PICAN
PENRW	1	B6HIP4|B6HIP4_PENRW
ASPNC	1	A2QJY9|A2QJY9_ASPNC
COLHI	1	H1V0C0|H1V0C0_COLHI
GIBF5	1	S0DRS5|S0DRS5_GIBF5
NEUCR	1	Q7S8U0|Q7S8U0_NEUCR
SCHPO	1	Q8TFH8|PEX12_SCHPO
SCHCM	1	D8PL00|D8PL00_SCHCM
CRYNJ	1	Q5K865|Q5K865_CRYNJ
DICDI	1	Q54N40|PEX12_DICDI
ENTHI	0	
CAEEL	1	Q19189|PEX12_CAEEL
DROME	1	Q9VPT5|PEX12_DROME
DANRE	1	B0R157|B0R157_DANRE
MOUSE	1	Q8VC48|PEX12_MOUSE
HUMAN	1	O00623|PEX12_HUMAN

	#	Uniprot code
	PEX13	**PEX13**
EGRACILIS	1	EG_transcript_23356
TRYB2	1	Q38E51|Q38E51_TRYB2
LEIMA	1	E9AFX0|E9AFX0_LEIMA
BODSA	1	A0A0S4J8A4|A0A0S4J8A4_BODSA
PHYIT	1	D0NN47|D0NN47_PHYIT
TOXGV	0	
TETTS	2	Q22WG9|Q22WG9_TETTS
NAEGR	0	
THAPS	1	B8BW97|B8BW97_THAPS

GALSU	1	M2W636|M2W636_GALSU
OSTTA	1	A0A090MB70|A0A090MB70_OSTTA
PHYPA	1	A9SV43|A9SV43_PHYPA
ARATH	1	Q9SRR0|PEX13_ARATH
YEAST	1	P80667|PEX13_YEAST
KOMPG	1	C4R2I6|C4R2I6_KOMPG
CANAL	1	A0A1D8PTW0|A0A1D8PTW0_CANAL
PICAN	1	A1XDT9_PICAN
PENRW	1	B6HFY2|B6HFY2_PENRW
ASPNC	1	A2R8L8|A2R8L8_ASPNC
COLHI	1	H1VQF3|H1VQF3_COLHI
GIBF5	1	S0DKC1|S0DKC1_GIBF5
NEUCR	1	Q7SG02|Q7SG02_NEUCR
SCHPO	1	O14136|PEX13_SCHPO
SCHCM	1	D8QDU2|D8QDU2_SCHCM
CRYNJ	1	Q5KKC0|Q5KKC0_CRYNJ
DICDI	1	Q54CL3|PEX13_DICDI
ENTHI	0	
CAEEL	1	Q19951|PEX13_CAEEL
DROME	1	Q7JRD4|Q7JRD4_DROME
DANRE	1	Q6PFQ3|Q6PFQ3_DANRE
MOUSE	1	Q9D0K1|PEX13_MOUSE
HUMAN	1	Q92968|PEX13_HUMAN

	#	Uniprot code
	PEX14	**PEX14**
EGRACILIS	1	EG_transcript_16511
TRYB2	1	Q38CL4|Q38CL4_TRYB2
LEIMA	1	Q4QBZ9|Q4QBZ9_LEIMA
BODSA	1	A0A0S4IV87|A0A0S4IV87_BODSA
PHYIT	1	D0NL84
TOXGV	1	A0A125YQE7|A0A125YQE7_TOXGV
TETTS	1	I7MA31
NAEGR	0	
THAPS	0	
GALSU	1	M2XC77|M2XC77_GALSU
OSTTA	1	Q01FR3|Q01FR3_OSTTA
PHYPA	2	A9TTH7|A9TTH7_PHYPA; A9TNA7|A9TNA7_PHYPA
ARATH	1	Q9FXT6|PEX14_ARAT
YEAST	1	P53112|PEX14_YEAST
KOMPG	1	C4R8Y8|C4R8Y8_KOMPG
CANAL	1	A0A1D8PSW1|A0A1D8PSW1_CANAL
PICAN	1	P78723|PEX14_PICAN
PENRW	1	B6H3Y6|B6H3Y6_PENRW
ASPNC	1	A2R7M1|A2R7M1_ASPNC
COLHI	1	H1UWW4|H1UWW4_COLHI
GIBF5	1	S0DRP4|S0DRP4_GIBF5
NEUCR	1	Q1K602|Q1K602_NEUCR
SCHPO	1	O60065|PEX14_SCHPO
SCHCM	1	D8PQV9|D8PQV9_SCHCM
CRYNJ	1	Q5K827|Q5K827_CRYNJ
DICDI	1	Q54C55|PEX14_DICDI
ENTHI	0	
CAEEL	1	Q93930|Q93930_CAEEL
DROME	1	Q9VPB8|Q9VPB8_DROME
DANRE	1	F1QX98|F1QX98_DANRE
MOUSE	1	Q9R0A0|PEX14_MOUSE

HUMAN	1	O75381|PEX14_HUMAN

	#	Uniprot code
	PEX15	**PEX15**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	Q08215|PEX15_YEAST
KOMPG	0	
CANAL	0	
PICAN	0	
PENRW	0	
ASPNC	0	
COLHI	0	
GIBF5	0	
NEUCR	0	
SCHPO	0	
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX16	**PEX16**
EGRACILIS	1	EG_transcript_9101
TRYB2	1	Q38ET6|Q38ET6_TRYB2
LEIMA	1	Q4QFA4|Q4QFA4_LEIMA
BODSA	1	A0A0S4JKY6|A0A0S4JKY6_BODSA
PHYIT	1	D0MV74|D0MV74_PHYIT
TOXGV	1	B9QI10|B9QI10_TOXGV
TETTS	1	Q22X13|Q22X13_TETTS
NAEGR	0	
THAPS	0	
GALSU	1	M2WVQ7|M2WVQ7_GALSU
OSTTA	0	
PHYPA	1	Phpat.020G026000 (gene) = Pp3c20_6130V3.1 (transcript)
ARATH	1	Q8S8S1|PEX16_ARATH
YEAST	0	
KOMPG	0	
CANAL	0	
PICAN	0	

PENRW	1	B6GYV6|B6GYV6_PENRW
ASPNC	1	A2QPQ3|A2QPQ3_ASPNC
COLHI	1	H1VQI3|H1VQI3_COLHI
GIBF5	1	S0EB43|S0EB43_GIBF5
NEUCR	1	Q7SD18|Q7SD18_NEUCR
SCHPO	1	O94516|PEX16_SCHPO
SCHCM	1	D8PNP3|D8PNP3_SCHCM
CRYNJ	1	Q5KG96|Q5KG96_CRYNJ
DICDI	1	Q550G0|PEX16_DICDI
ENTHI	1	C4M0K4|C4M0K4_ENTHI
CAEEL	0	
DROME	1	Q9VPB9|Q9VPB9_DROME
DANRE	1	Q4QRH7|PEX16_DANRE
MOUSE	1	Q91XC9|PEX16_MOUSE
HUMAN	1	Q9Y5Y5|PEX16_HUMAN

	#	Uniprot code
	PEX17	**PEX17**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	P40155|PEX17_YEAST
KOMPG	1	C4R8F7|C4R8F7_KOMPG
CANAL	1	A0A1D8PI79|A0A1D8PI79_CANAL
PICAN	1	A1XDU0|A1XDU0_PICAN
PENRW	0	
ASPNC	0	
COLHI	0	
GIBF5	0	
NEUCR	0	
SCHPO	0	
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX18	**PEX18**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	

PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	P38855|PEX18_YEAST
KOMPG	0	
CANAL	0	
PICAN	0	
PENRW	0	
ASPNC	0	
COLHI	0	
GIBF5	0	
NEUCR	0	
SCHPO	0	
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX19	**PEX19**
EGRACILIS	0	
TRYB2	1	Q38DH6|Q38DH6_TRYB2
LEIMA	1	E9AFF0|E9AFF0_LEIMA
BODSA	1	A0A0S4JDC2|A0A0S4JDC2_BODSA
PHYIT	1	D0MWU5|D0MWU5_PHYIT
TOXGV	0	
TETTS	1	Q23AR4|Q23AR4_TETTS
NAEGR	1	D2W353|D2W353_NAEGR
THAPS	1	>jgi|Thaps3|10721|fgenesh1_pg.C_chr_17000184
GALSU	1	M2WSU1|M2WSU1_GALSU
OSTTA	0	
PHYPA	1	A0A2K1L1U9|A0A2K1L1U9_PHYPA
ARATH	2	Q94EI3|PE192_ARATH; Q9SRQ3|PE191_ARATH
YEAST	1	Q07418|PEX19_YEAST
KOMPG	1	C4R1J0|C4R1J0_KOMPG
CANAL	1	Q5A330|Q5A330_CANAL
PICAN	1	Q96WN7|Q96WN7_PICAN
PENRW	1	B6HC35|B6HC35_PENRW
ASPNC	1	A2QDD4|A2QDD4_ASPNC
COLHI	1	H1V5U6|H1V5U6_COLHI
GIBF5	1	S0E1H7|S0E1H7_GIBF5
NEUCR	1	Q1K773|Q1K773_NEUCR/
SCHPO	1	Q10485|YDFE_SCHPO
SCHCM	1	D8QJR5|D8QJR5_SCHCM
CRYNJ	1	Q5KIF7|Q5KIF7_CRYNJ

DICDI	1	Q555I0|PEX19_DICDI
ENTHI	1	C4LWR7|C4LWR7_ENTHI
CAEEL	1	P34453|PEX19_CAEEL
DROME	1	Q8IP97|Q8IP97_DROME
DANRE	1	F1R313|F1R313_DANRE
MOUSE	1	Q8VCI5|PEX19_MOUSE
HUMAN	1	P40855|PEX19_HUMAN

	#	Uniprot code
	PEX20	**PEX20**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	0	
KOMPG	1	C4R0R1|C4R0R1_KOMPG
CANAL	1	A0A1D8PPS0|A0A1D8PPS0_CANAL
PICAN	1	Q3ZJZ2|Q3ZJZ2_PICAN
PENRW	1	B6HER8|B6HER8_PENRW
ASPNC	1	A2QHM2|A2QHM2_ASPNC
COLHI	1	H1VF12|H1VF12_COLHI
GIBF5	1	S0DQ96|S0DQ96_GIBF5
NEUCR	1	Q1K605|Q1K605_NEUCR
SCHPO	0	
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX21	**PEX21**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	

ARATH	0	
YEAST	1	P50091|PEX21_YEAST
KOMPG	0	
CANAL	0	
PICAN	0	
PENRW	0	
ASPNC	0	
COLHI	0	
GIBF5	0	
NEUCR	0	
SCHPO	0	
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX22	**PEX22**
EGRACILIS	1	EG_transcript_21449
TRYB2	1	Q38AL6|Q38AL6_TRYB2
LEIMA	0	
BODSA	1	A0A0S4JN26|A0A0S4JN26_BODS
PHYIT	1	D0MS36|D0MS36_PHYIT
TOXGV	1	B9QPC3|B9QPC3_TOXGV
TETTS	1	I7LWE8|I7LWE8_TETTS
NAEGR	1	D2VPE6|D2VPE6_NAEGR
THAPS	0	
GALSU	1	M2XXM3|M2XXM3_GALSU
OSTTA	0	
PHYPA	1	A9RSG2|A9RSG2_PHYPA
ARATH	1	Q9LSX7|PEX22_ARATH
YEAST	1	P39718|PEX22_YEAST
KOMPG	1	C4R500|C4R500_KOMPG
CANAL	1	Q59KM6|Q59KM6_CANAL
PICAN	1	A2T0X6|A2T0X6_PICAN
PENRW	1	B6GVZ4|B6GVZ4_PENRW
ASPNC	1	A2QA75|A2QA75_ASPNC
COLHI	1	H1VR05|H1VR05_COLHI
GIBF5	1	S0DK65|S0DK65_GIBF5
NEUCR	1	Q1K4L9|Q1K4L9_NEUCR
SCHPO	1	Q1K9B6|YFS2_SCHPO
SCHCM	0	
CRYNJ	0	
DICDI	1	Q54ZU0|Q54ZU0_DICDI
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code

	PEX23	PEX23
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	0	
KOMPG	0	
CANAL	0	
PICAN	1	A4GFC7|A4GFC7_PICAN
PENRW	1	B6HMP0|B6HMP0_PENRW
ASPNC	1	A2R1B9|A2R1B9_ASPNC
COLHI	0	
GIBF5	0	
NEUCR	1	V5IKX2|V5IKX2_NEUCR
SCHPO	0	
SCHCM	0	
CRYNJ	1	Q5KPU0|Q5KPU0_CRYNJ
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX24	
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	0	
KOMPG	0	
CANAL	0	
PICAN	1	A4GFC9|A4GFC9_PICAN
PENRW	1	B6HF69|B6HF69_PENRW
ASPNC	0	
COLHI	2	H1VKN1|H1VKN1_COLHI; H1W2X7|H1W2X7_COLHI
GIBF5	1	S0E157|S0E157_GIBF5

NEUCR	1	V5IMJ5|V5IMJ5_NEUCR
SCHPO	0	
SCHCM	1	D8PXT8|D8PXT8_SCHCM
CRYNJ	1	Q5KJ71|Q5KJ71_CRYNJ
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX25	**PEX25**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	Q02969|PEX25_YEAST
KOMPG	1	C4R3T3|C4R3T3_KOMPG
CANAL	1	A0A1D8PQ11|A0A1D8PQ11_CANAL
PICAN	1	A4GFD0|A4GFD0_PICAN
PENRW	0	
ASPNC	0	
COLHI	0	
GIBF5	0	
NEUCR	0	
SCHPO	0	
SCHCM	1	D8PJX5|D8PJX5_SCHCM
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX26	**PEX26**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	

THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	0	
KOMPG	1	C4R825|C4R825_KOMPG
CANAL	1	A0A1D8PL18|A0A1D8PL18_CANAL
PICAN	1	A4GFD1|A4GFD1_PICAN
PENRW	1	B6HJI5|B6HJI5_PENRW
ASPNC	1	A2QK20|A2QK20_ASPNC
COLHI	1	H1VSS1|H1VSS1_COLHI
GIBF5	1	S0DU76|S0DU76_GIBF5
NEUCR	1	Q7RZL8|Q7RZL8_NEUCR
SCHPO	0	
SCHCM	1	D8QIJ8|D8QIJ8_SCHCM
CRYNJ	1	Q5KQ10|Q5KQ10_CRYNJ
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	1	F1RBL0|F1RBL0_DANRE
MOUSE	1	Q8BGI5|PEX26_MOUSE
HUMAN	1	Q7Z412|PEX26_HUMAN

	#	Uniprot code
	PEX27	**PEX27**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	Q08580|PEX27_YEAST
KOMPG	0	
CANAL	0	
PICAN	0	
PENRW	0	
ASPNC	0	
COLHI	0	
GIBF5	0	
NEUCR	0	
SCHPO	0	
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	

HUMAN	0	

	#	Uniprot code
	PEX28	
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	P38848|PEX28_YEAST
KOMPG	1	C4R5U5|C4R5U5_KOMPG
CANAL	1	Q5ANG0|Q5ANG0_CANAL
PICAN	0	
PENRW	0	
ASPNC	0	
COLHI	0	
GIBF5	0	
NEUCR	0	
SCHPO	0	
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX29	
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	Q03370|PEX29_YEAST
KOMPG	1	C4R0X2|C4R0X2_KOMPG
CANAL	1	A0A1D8PN55|A0A1D8PN55_CANAL
PICAN	1	A4GFD2|A4GFD2_PICAN
PENRW	0	

ASPNC	0	
COLHI	0	
GIBF5	0	
NEUCR	0	
SCHPO	1	O13679|YJ15_SCHPO
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX30	**PEX30**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	Q06169|PEX30_YEAST
KOMPG	1	C4QYD8|C4QYD8_KOMPG
CANAL	1	Q5ACV4|Q5ACV4_CANAL
PICAN	0	
PENRW	0	
ASPNC	0	
COLHI	1	H1VSQ0|H1VSQ0_COLHI
GIBF5	0	
NEUCR	0	
SCHPO	0	
SCHCM	1	D8QD01|D8QD01_SCHCM
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX31	**PEX31**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	

TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	P53203|PEX31_YEAST
KOMPG	1	C4R7G1|C4R7G1_KOMPG
CANAL	1	A0A1D8PF18|A0A1D8PF18_CANAL
PICAN	0	
PENRW	0	
ASPNC	0	
COLHI	1	H1UZP5|H1UZP5_COLHI
GIBF5	1	S0E5E5|S0E5E5_GIBF5
NEUCR	0	
SCHPO	0	
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX32	**PEX32**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	P38292|PEX32_YEAST
KOMPG	0	
CANAL	0	
PICAN	1	A4GFD3|A4GFD3_PICAN
PENRW	0	
ASPNC	0	
COLHI	0	
GIBF5	0	
NEUCR	0	
SCHPO	1	O59807|PEX32_SCHPO
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	

DROME	0
DANRE	0
MOUSE	0
HUMAN	0

	# **PEX33**	Uniprot code **PEX33**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	0	
KOMPG	0	
CANAL	0	
PICAN	0	
PENRW	1	B6GX19|B6GX19_PENRW
ASPNC	1	A2QBY9|A2QBY9_ASPNC
COLHI	1	H1W5I0|H1W5I0_COLHI
GIBF5	1	S0ELZ5|S0ELZ5_GIBF5
NEUCR	1	Q7RY61|Q7RY61_NEUCR
SCHPO	0	
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	# **PEX34**	Uniprot code **PEX34**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	P25584|PEX34_YEAST
KOMPG	0	

CANAL	1	A0A1D8PTB7|A0A1D8PTB7_CANAL
PICAN	1	A0A1B7SEF7_9ASCO
PENRW	0	
ASPNC	0	
COLHI	0	
GIBF5	0	
NEUCR	0	
SCHPO	0	
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX35	**PEX35**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	P53293|YG3W_YEAST
KOMPG	0	
CANAL	0	
PICAN	0	
PENRW	0	
ASPNC	0	
COLHI	0	
GIBF5	0	
NEUCR	0	
SCHPO	0	
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX36	**PEX36**
EGRACILIS	0	
TRYB2	0	

LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	0	
KOMPG	1	C4QWS8|C4QWS8_KOMPG
CANAL	0	
PICAN	0	
PENRW	0	
ASPNC	0	
COLHI	0	
GIBF5	0	
NEUCR	0	
SCHPO	0	
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	PEX37	**PEX37**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	0	
KOMPG	1	C4R902|C4R902_KOMPG
CANAL	1	Q5A7L7|Q5A7L7_CANAL
PICAN	1	A0A1B7SAK9|A0A1B7SAK9_9ASCO
PENRW	1	B6HJ42|B6HJ42_PENRW
ASPNC	1	A2QK28|A2QK28_ASPNC
COLHI	1	H1V3E8|H1V3E8_COLHI
GIBF5	2	S0DRE5|S0DRE5_GIBF5; S0EML7|S0EML7_GIBF5
NEUCR	1	U9W802|U9W802_NEUCR
SCHPO	0	
SCHCM	1	D8PPM7|D8PPM7_SCHCM

CRYNJ	0
DICDI	0
ENTHI	0
CAEEL	0
DROME	0
DANRE	0
MOUSE	0
HUMAN	0

Relate non-PEX protein families

	#	Uniprot code
	PEX23-l	**PEX23-like**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	
OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	1	P40031|SPO73_YEAST
KOMPG	1	C4R4W3|C4R4W3_KOMPG
CANAL	1	Q5AB70|Q5AB70_CANAL
PICAN	1	A4GFC8|A4GFC8_PICAN
PENRW	1	B6GXT5|B6GXT5_PENRW
ASPNC	1	A2R7I1|A2R7I1_ASPNC
COLHI	1	H1VXZ9|H1VXZ9_COLHI
GIBF5	1	S0E7F5|S0E7F5_GIBF5
NEUCR	1	Q1K6B6|Q1K6B6_NEUCR
SCHPO	1	O94611|MUG65_SCHPO
SCHCM	1	D8PQ56|D8PQ56_SCHCM
CRYNJ	1	Q5KPF4|Q5KPF4_CRYNJ
DICDI	0	
ENTHI	0	
CAEEL	0	
DROME	0	
DANRE	0	
MOUSE	0	
HUMAN	0	

	#	Uniprot code
	TECPR1	**TECPR1 family**
EGRACILIS	0	
TRYB2	0	
LEIMA	0	
BODSA	0	
PHYIT	0	
TOXGV	0	
TETTS	0	
NAEGR	0	
THAPS	0	
GALSU	0	

OSTTA	0	
PHYPA	0	
ARATH	0	
YEAST	0	
KOMPG	0	
CANAL	0	
PICAN	0	
PENRW	0	
ASPNC	0	
COLHI	0	
GIBF5	0	
NEUCR	0	
SCHPO	0	
SCHCM	0	
CRYNJ	0	
DICDI	0	
ENTHI	0	
CAEEL	1	Q22088|Q22088_CAEEL
DROME	1	Q9VWB0|TECPR_DROME
DANRE	2	X1WFY6|X1WFY6_DANRE; F1R5H3|F1R5H3_DANRE
MOUSE	1	Q80VP0|TCPR1_MOUSE
HUMAN	1	Q7Z6L1|TCPR1_HUMAN

Codes marked in yellow are based on literature, where the corresponding protein codes could not be found in the Uniprot database.