

# Fat storage-inducing transmembrane (FIT or FITM) proteins are related to lipid phosphatase/phosphotransferase enzymes

Matthew J Hayes<sup>1</sup>, Vineet Choudhary<sup>2</sup>, Namrata Ojha<sup>2</sup>, John JH Shin<sup>3</sup>, Gil-Soo Han<sup>4</sup>, George M. Carman<sup>4</sup>, Christopher JR Loewen<sup>3</sup>, William A Prinz<sup>2</sup> and Timothy P Levine<sup>1,\*</sup>

<sup>1</sup> University College London Institute of Ophthalmology, 11-43 Bath Street, London, EC1V 9EL, UK.

<sup>2</sup> Laboratory of Cell and Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA.

<sup>3</sup> Department of Cellular and Physiological Sciences, Life Sciences Institute, University of British Columbia, Vancouver, British Columbia, Canada.

<sup>4</sup> Department of Food Science and Rutgers Center for Lipid Research, Rutgers University, New Brunswick, New Jersey 08901, USA.

\* Corresponding Author:

Tim Levine, University College London Institute of Ophthalmology, 11-43 Bath Street, London, EC1V 9EL, UK; Telephone: (+44/0) 20 7608 4027; E-mail: tim.levine@ucl.ac.uk

**ABSTRACT** Fat storage-inducing transmembrane (FIT or FITM) proteins have been implicated in the partitioning of triacylglycerol to lipid droplets and the budding of lipid droplets from the ER. At the molecular level, the sole relevant interaction is that FITMs directly bind to triacylglycerol and diacylglycerol, but how they function at the molecular level is not known. *Saccharomyces cerevisiae* has two FITM homologues: Scs3p and Yft2p. Scs3p was initially identified because deletion leads to inositol auxotrophy, with an unusual sensitivity to addition of choline. This strongly suggests a role for Scs3p in phospholipid biosynthesis. Looking at the FITM family as widely as possible, we found that FITMs are widespread throughout eukaryotes, indicating presence in the last eukaryotic common ancestor. Protein alignments also showed that FITM sequences contain the active site of lipid phosphatase/phosphotransferase (LPT) enzymes. This large family transfers phosphate-containing headgroups either between lipids or in exchange for water. We confirmed the prediction that FITMs are related to LPTs by showing that single amino-acid substitutions in the presumptive catalytic site prevented their ability to rescue growth of the mutants on low inositol/high choline media when over-expressed. The substitutions also prevented rescue of other phenotypes associated with loss of FITM in yeast, including mistargeting of Opi1p, defective ER morphology, and aberrant lipid droplet budding. These results suggest that Scs3p, Yft2p and FITMs in general are LPT enzymes involved in an as yet unknown critical step in phospholipid metabolism.

doi: 10.15698/mic2018.02.614

Received originally: 13.10.2017;

in revised form: 14.12.2017,

Accepted 18.12.2017,

Published 28.12.2017.

**Keywords:** endoplasmic reticulum retention motif, endoplasmic reticulum stress, lipid biosynthesis enzyme, lipid droplet, remote homology search, type 2 diabetes.

#### Abbreviations:

ER – endoplasmic reticulum,

FIT – fat-inducing transcript,

FITM – fat storage-induced transmembrane,

LPT – lipid phosphotransferase,

SAR – stramenopiles alveolates and rhizaria,

SMS – sphingomyelin synthase,

TMD – transmembrane domain.

## INTRODUCTION

All eukaryotic cells store the neutral lipids triacylglycerol and sterol ester in lipid droplets [1]. Understanding this is societally important both for feeding the growing human population [2] and conversely to address problems of the consumption of excess calories with associated metabolic syndrome, type 2 diabetes and atherosclerosis [3]. One of

the transcriptional regulators of triacylglycerol homeostasis is the peroxisome proliferator-activated receptor- $\alpha$  (PPAR $\alpha$ ) nuclear hormone receptor, activation of which by fibrate drugs may be beneficial for some atherosclerosis patients [4]. The FITM1/2 (Fat storage-Inducing Transmembrane) proteins, also called FIT1/2 (Fat-inducing transcript), are integral ER (endoplasmic reticulum) proteins

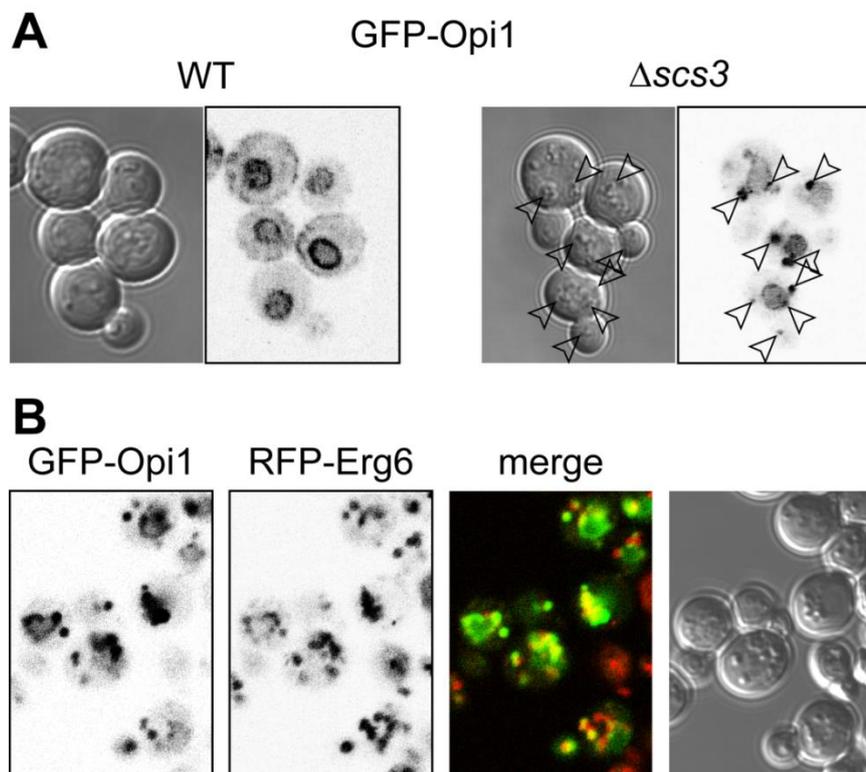
first identified transcriptional targets of fibrate [5]. FITM1 is mainly restricted to skeletal muscle, whereas FITM2 is expressed in most tissues, especially in adipose tissue. They are required for triacylglycerol storage in lipid droplets [5]. In support of their biological significance, a SNP (single nucleotide polymorphism) in the FITM2 (C20orf142) promoter is a high risk locus for type 2 diabetes in Asian populations [6-8]. Homozygous FITM2 mutations introducing a nonsense codon very near to the first residue cause deafness and dystonia [9]. This clinical picture is at odds with the effect of homozygous deletion of FITM2 in mice, which causes problems mainly through lack of lipid droplet formation [10, 11].

In the short-term, over-expression of FITM does not increase the rate of triacylglycerol synthesis, nor does it increase the expression of genes associated with triacylglycerol synthesis, nor does it reduce the rate of neutral lipid turnover [5]. This has suggested that FITMs directly act on the partitioning of triacylglycerol into the storage pathway, possibly affecting its distribution within the ER or its transport from ER to lipid droplets [11, 12]. In support of this, FITM2 binds directly to both triacylglycerol and its immediate precursor diacylglycerol [12]. Also, mutation of a conserved hydrophobic motif in one of the transmembrane domains of FITM2 increases binding to these lipids and increases lipid droplet size [12, 13], suggesting that FITM over-activation leads to increased neutral lipid storage. However, both the motif and the lipid binding site are predicted to lie within the membrane, so FITMs are unlikely to act like lipid transfer proteins that transport neutral lipid between organelles [14].

FITM homologues (more closely related to FITM2 than

FITM1) have been described so far in animals and fungi, with budding yeast *Saccharomyces cerevisiae* having two FITMs homologues called Scs3p (or FIT2b) and Yft2p (FIT2a). Budding yeast is a useful model for many aspects of neutral lipid metabolism [15-18], and it may also model pathogenic fungi, since deletion of *SCS3* in the pathogenic fungus *Candida parapsilosis* reduced its virulence [19]. Yft2p is often overlooked and its deletion has few phenotypes [19]. In contrast, Scs3p is better studied. *SCS3* is required for normal phospholipid synthesis, being one of over 200 genes deletion of which cause inositol auxotrophy, meaning that growth is reduced without the addition of extra inositol, the phospholipid headgroup. On most growth media *scs3Δ* cells have a weak inositol auxotrophy [20] that is rescued by human FITM2 [21]. However, *scs3Δ* cells have a strong inositol auxotrophy when grown in excess choline [22]. The severity of the effect of choline on inositol requirement is almost unique, indicating that Scs3p has a specific molecular function [23]. System-wide genetic analysis of yeast lacking both FITM homologues showed links to many pathways including ER stress [21]. A strong negative synthetic interaction between *scs3Δ* and deletion of the yeast atlastin (*sey1Δ*) [24] suggests that Scs3p promotes homotypic fusion of ER tubules [25]. Previously we found that FITMs are required for correct budding of nascent lipid droplets. They first form as neutral lipid lenses up to 50 nm diameter within the leaflet of the ER, and normally bud outwards into the cytoplasm. Without FITMs, lipid lenses fail to emerge from the ER lumen [26].

This evidence suggests that FITMs act in the ER to affect both that organelle and lipid droplets. However, their action at a molecular level is still not known. Here, to gain



**FIGURE 1: GFP-Opi1p re-localises to lipid droplets in cells lacking Scs3p.** (A) GFP-Opi1p was expressed in wild-type cells (left) and *scs3Δ* cells (right) that were grown without inositol. In wild-type cells, GFP-Opi1p localises most prominently to the nuclear envelope, and also to the cortical ER that forms patches adjacent to the cell periphery [30]. In *scs3Δ* cells, GFP-Opi1p is only faintly on the nuclear envelope, and mostly in puncta. Arrowheads indicate these puncta and the corresponding positions in the accompanying transmission image. (B) *scs3Δ* cells expressing GFP-Opi1p as in A (1<sup>st</sup> panel), and co-expressing RFP-Erg6p, a lipid droplet marker (2<sup>nd</sup> panel), with pseudo-coloured green and red merge (panel 3, superposition = yellow) and transmission image (panel 4).

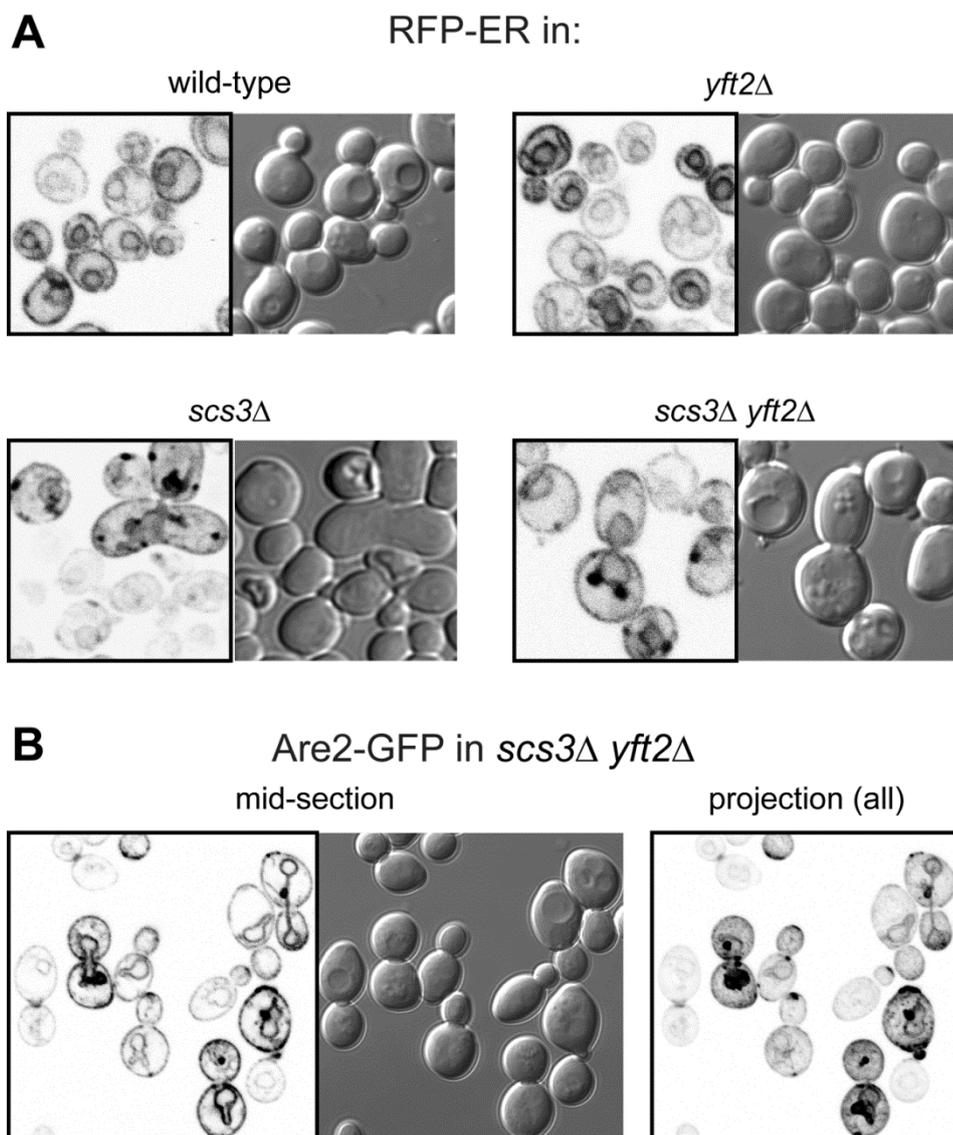
insight into FITM function, we look at their sequences, where the only published observation is that they have an unusually high proportion of tryptophan residues [27]. We identify multiple sequence motifs in FITMs, the most important of which are the catalytic triad of the large family of lipid phosphotransferase (LPT) enzymes. We then show that key residues in this triad are important for FITM function, suggesting that the sequence homology between FITMs and LPTs is physiologically relevant.

**RESULTS**

**Deletion of *SCS3* relocates Opi1p to lipid droplets**

The rate-limiting enzyme in inositol production by yeast is Ino1p, which is transcriptionally repressed by Opi1p in response to the presence of exogenous inositol, and depressed in the absence of inositol [28]. With de-repression, Opi1p localises to the ER by simultaneous detection of both the ER protein VAP (called Scs2p in yeast) [29] and phosphatidic acid [30]. Since the inositol auxo-

trophy of *scs3Δ* cells is rescued by deletion of *OPI1*, it has been suggested that Opi1p translocates to the nucleus in *scs3Δ* cells [21]. We tested this suggestion, and found that in *scs3Δ* cells grown without exogenous inositol GFP-Opi1p was not seen in the cortical (non-nuclear) ER, but instead it was found on multiple puncta (Figure 1A). By light microscopy these puncta co-localised with bodies that were highly refractile by differential interference contrast (DIC), a characteristic of lipid droplets. To study this further, we co-expressed GFP-Opi1p with the lipid droplet marker Erg6-RFP, and found strong co-localisation when this was done in a *scs3Δ* background (Figure 1B). After addition of inositol, GFP-Opi1p translocated from lipid droplets to the nucleus as in wild-type cells (data not shown). Although the GFP-Opi1p puncta in *scs3Δ* cells are bright, the overall amount of fluorescence per cell is not significantly higher than in wild-type cells (data not shown). As for the nuclear-cytoplasmic distribution of Opi1p, intra-nuclear GFP-Opi1p fluorescence in *scs3Δ* cells was double that of wild-type



**FIGURE 2: Punctate accumulations of ER form in cells lacking *SCS3*.** Fluorescent (inverted gray scale) and DIC images of cells with indicated genotype expressing either (A) RFP-ER [35] or (B) Are2-GFP [36]. For Are2-GFP only, a maximum intensity projection generated in Image-J (NIH) from a Z-stack of 17 images 380 nm apart is also shown (right-hand panel).

cells as assessed from fluorescence images (data not shown), which is consistent with the previous work showing Opi1p activation in this strain [21]. We noted that deletion of FITMs did not affect the overall number and size of lipid droplets in budding yeast, in agreement with other work [21].

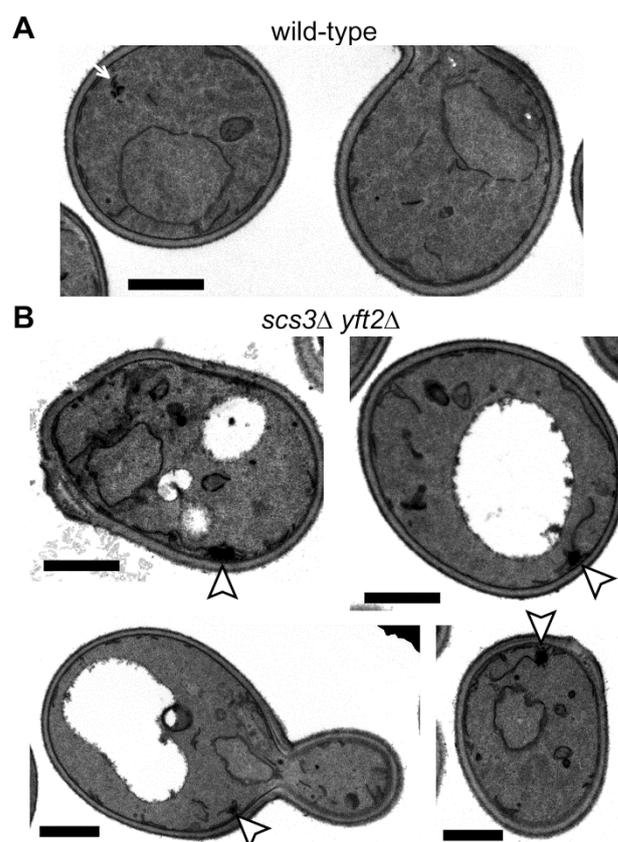
### Deletion of FITM proteins in yeast leads to ER tangles

The close proximity of Opi1-positive puncta to lipid droplets in *scs3Δ* cells is similar to the effect on Opi1p of other mutations affecting phospholipid and neutral lipid metabolism [31-33]. Since such differences in lipid metabolism can accompany changes in ER structure [15], we compared the structure of the yeast ER in wildtype, *scs3Δ*, *yft2Δ* and double delete *scs3Δ yft2Δ* cells. In *scs3Δ* and particularly in the double deletion there were punctate structures within the ER, as seen with two different markers, which did not align with lipid droplets (Figure 2A and B). In the *scs3Δ yft2Δ* double deletion there were also extended bulges of the nuclear envelope (Figure 2B). To determine the effect of lack of FITM on the ER in more detail, we examined *scs3Δ yft2Δ* cells by electron microscopy (Figure 3). This revealed tangles of membrane associated with the ER in 13% of cell profiles (27/214 from 33 consecutive random images) that were hardly ever seen in wild-type cells (1/127 cell profiles, 24 images). Both strains also contain partly stacked membranes that is likely to be the Golgi apparatus [34]. The electron micrographs do not identify the origin of the tangles with certainty, but they are likely to correlate with the additional ER-positive puncta.

### Yft2p is found in the endocytic pathway, only being partly retained in the ER

In the evolution of FITM proteins, two independent gene duplications are well documented, one producing FITM1 and FITM2 in vertebrates, the other duplication in saccharomycete fungi, the class that includes both *S. cerevisiae* and *Candida albicans*, producing Scs3p and Yft2p [5]. Despite the breadth of these duplications, and detailed study of the shared and unique genetic interactions of *SCS3* and *YFT2* [21], it is not yet known how any pair of FITM paralogues are integrated. We wondered if there is any specialisation associated with FITM duplication in yeast. One common divergence between paralogues is in their intracellular localisation. Both human FITM2 and FITM1 localise to the ER [5]. Here we expressed Scs3p and Yft2p tagged with GFP at the C-terminus, using the constitutive portion of the *PHO5* promoter, which causes moderate over-expression [37]. Scs3-GFP localised to the ER, both at high levels of expression (Figure 4A) and from its own promoter (Figure 4B), a localisation that we and others have reported before [26, 38]. For Yft2-GFP, under its own promoter no staining was detected (not shown), as it has been reported previously [38], and the lack of detectability was also true when Scs3-GFP was expressed using the *YFT2* promoter (not shown). We therefore expressed Yft2-GFP using the *PHO5* promoter. This revealed a much wider distribution than for any FITM reported previously. Yft2-GFP was present in the ER, as shown by frequent nuclear envelope pro-

files (Figure 4C). However, there were additional localisations, with diffuse fluorescence inside the lumen of the yeast vacuole (equivalent of lysosome in animal cells), and multiple puncta associated with the vacuole. Similar intravacuolar diffuse GFP was also seen in some cells expressing Yft2-GFP with the *GPD1* promoter (data not shown). The puncta are characteristic of endosomes, which are typically associated with the vacuole, which we confirmed by showing that many of the puncta were accessed rapidly by the endocytosed dye FM4-64 (Figure 4D) [39]. Since the GFP tag is predicted to be cytoplasmic, diffuse intravacuolar fluorescence implies delivery of the whole construct into the vacuolar lumen. Crossing of the vacuolar limiting membrane can result from one of two routes: either by typical secretion with vacuolar protein sorting and ESCRT-mediated (endosomal sorting complexes required for transport) inward budding of the vacuolar membrane, or by autophagy (for example ER-phagy). To differentiate between these possibilities, we expressed Yft2-GFP in strains where either the vacuolar protein sorting pathway or autophagy was inactivated. The vacuolar protein sorting pathway was required for intravacuolar targeting of Yft2-GFP, which was held up on the vacuolar limiting membrane



**FIGURE 3: Ultrastructure of double delete *scs3Δ yft2Δ* cells.** Transmission EM images of (A) wild-type cell and (B) a gallery of *scs3Δ yft2Δ* cells. In the latter there are tangles of membrane (arrowheads in B), which are absent from the former, although these occasionally have very small stacked membranes (arrow in A). Size bars are all 1 μm.

in *vps4Δ* cells, consistent with delivery of this membrane protein through the Golgi and endosomes (Figure S1). In contrast, the distribution of Yft2-GFP was unaffected by the *atg1Δ* deletion, which inhibits multiple routes of autophagy, and Scs3-GFP was unaffected by either deletion (Figure S1). This indicates that Yft2-GFP tends to leave the ER and traffic all the way through the secretory pathway to the vacuole, which it enters by ESCRT driven inward vesiculation.

**FITMs commonly have ER retention motifs, but rarely duplicate to produce homologues that lack the motifs, indicating residence outside the ER**

Based on the different localisation of Scs3-GFP and Yft2-GFP, we looked for sequence features that distinguish between them that might explain the difference. For integral

membrane proteins, ER retrieval from the Golgi in COP-I vesicles results from a combination of ill-defined luminal signals [40] and two types of well-defined cytoplasmic motifs: lysine-based signals at the extreme C-terminus (either KKxx-stop or KxKxx-stop [41]), and di-arginine motifs (RR, RxR and RxxR) that can occur anywhere (and in high copy number) in cytoplasmic extensions or loops [42, 43]. The function of di-arginine motifs may depend on precise position, for example being less active if they are very close to a transmembrane domain (TMD) [40, 44, 45]. Examination of the FITMs in *S. cerevisiae* shows that Scs3p has 6 di-arginine motifs in predicted cytoplasmic loops, while Yft2p has none. The presence across large families of polytopic proteins either of KKxx motifs and their proven variants [46] or of >3 di-arginine motifs per protein seems to reflect ER localization (Table 1A, lines 1-3). Note that where we

**Table 1. ER retention motifs in FITM proteins.**

**A**

line number	protein family	location	number of seqs	residues per protein	R motifs	...KKxx*
1	Shr3p	ER	840	269.3	2.0	59.7%
2	INSIG	ER	1277	288.1	3.1	3.2%
3	LAPTM	lysosome/ vacuole	1270	267.0	1.9	0.6%
4	FITM	all	1718	321.7	3.7	3.9%
5		FITM1 only	149	293.1	3.6	2.7%
6	Scs3p	ER	40	361.4	3.7	10.0%
7	Yft2p	vacuole	52	260.8	1.6	0.0%
8	SMS1/2	TGN & PM	840	382.7	2.7	0.4%
9	SMSr	ER	647	442.0	5.5	25.8%

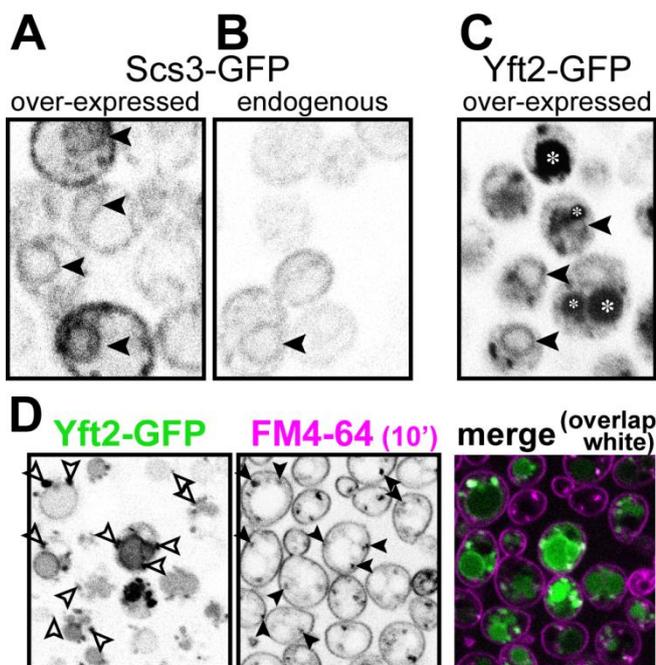
**B**

Duplicated FITM → ≥2 paralogues								Most closely related FITM (non-duplicated)					
phylum	species	# FITMs	iso-form	similarity (A cf B)	R motifs	KKxx*	predict	sim% ←→	species	related FITM	R motifs	KKxx*	predict
chordate	<i>Branchiostoma floridae</i>	2	A	56%	4	/	ER	41%	<i>Xenopus tropicalis</i>	FITM2	4	/	ER
			B		1	/	unsure	44%					
arthropod	<i>Limulus polyphemus</i>	3	A1	78% (A1 cf A2)	0	/	not ER	51%	<i>Daphnia pulex</i>	FITM2	1	/	unsure
			A2		0	/	not ER	48%					
			B	4	/	ER	45%						
arthropod	<i>Sarcoptes scabiei</i>	2	A	31%	0	/	not ER	34%	<i>Tetranychus urticae</i>	FITM2	2	/	ER
			B		1	/	unsure	33%					

**C**

phylum	species	# FITMs	isoform	similarity (A cf B)	R motifs	...KKxx*	predict
arthropod	<i>Hyalella azteca</i>	2	A	27%	1	/	unsure
			B		3	partial (Kxxxx*)	ER

(A) ER retention motifs (both di-arginine and C-terminal di-lysine) were compared across multiple large protein families (see Methods). Lines 1-3: examples of ER and lysosomal/vacuolar proteins; 4-7: FITM proteins – all and different subgroups; 8/9: Sphingomyelin synthase subgroups. (B) ER retention motifs were used to predict location in three recently duplicated FITM paralogues. Left-hand side: paralogues, including % similarity to each other; right-hand side: the FITM sequence in the most closely related organism, with central shaded column indicating % similarity between sequences on left- and right- hand sides. Prediction of location based on motifs: C-terminal di-lysine or ≥2 di-arginines = ER, 1 di-arginine = unsure, no di-arginines = not ER. (C) As B, showing one duplicated FITM pair, where the duplication appears to pre-date many speciation events, making it impossible to identify a single FITM that resembles the presumed antecedent.



**FIGURE 4: Yft2-GFP mainly escapes the ER to reach the vacuole.** (A) Scs3-GFP expressed from the *PHO5* promoter. (B) The same construct expressed from the *SCS3* promoter. Scs3-GFP localises to the ER, with diagnostic nuclear profiles (arrowheads), and only a faint signal with the endogenous promoter. (C) Yft2-GFP expressed from the *PHO5* promoter. In addition to some nuclear profiles, there are large spherical regions of diffuse staining (intravacuole – asterisks) and additional puncta. (D) Yft2-GFP expressing cells were incubated with FM4-64 on ice to label the plasma membrane, then warmed to 30°C for 10 minutes to allow dye to enter the endocytic pathway, but not to reach vacuoles. Puncta positive for Yft2-GFP (left hand panel) that were also positive for FM4-64 (middle panel) are indicated by arrowheads, and the degree of overlap is also shown in a merged coloured image (right hand panel, GFP = green, FM4-64 = magenta, overlap = white).

counted total motifs in large families all di-arginines were included, not just those in cytoplasmic loops or distant from TMDs. Applying this analysis to the fungal FITM families showed that orthologues of Scs3p consistently have more di-arginine motifs (mean=3.7, n=40) than orthologues of Yft2p (mean=1.6, n=52) (Table 1A, lines 6 & 7). Also, KKxx motifs are more common in Scs3p orthologues (4/40) than Yft2p orthologues (0/52). Therefore, simple analysis of the primary structure of fungal FITMs reveals that the split between Scs3p and Yft2p is associated with high and low propensity to be retained in the ER respectively.

Since groups of Yft2p and Scs3p sequences have a dichotomy in primary structure that explains their different localisations, we wondered whether similar dichotomies have arisen in other FITM duplications. For vertebrate FITMs, there is no such difference, with FITM1s having rates of ER retention motifs that are indistinguishable from the whole FITM family (Table 1A, lines 4 & 5). We next looked at other gene duplications of FITM/Scs3p. The literature to date has reported only the two major duplications in vertebrates and saccharomycetes [5]. We found three species outside the vertebrate taxon that have more than one FITM2, while their close relatives have just one: two FITM2s in *Branchiostoma floridae*, the most basal extant cephalochordate; three FITM2s in *Limulus polyphemus* (horseshoe crab); and two FITM2s in the mite *Sarcoptes scabiei*. In each of these species, the duplication events occurred more recently than the FITM1/2 duplication (Figure S2). A comparison of each of the duplicated proteins with the FITM2 that is most closely related, i.e. best shows properties of the common antecedent, showed that the predicted location has changed for one or both duplicates (Table 1B). We also found duplicated FITM2s and no FITM1 in the amphipod crustacean *Hyalomma*, but here the duplication event is placed prior to many other speciation events,

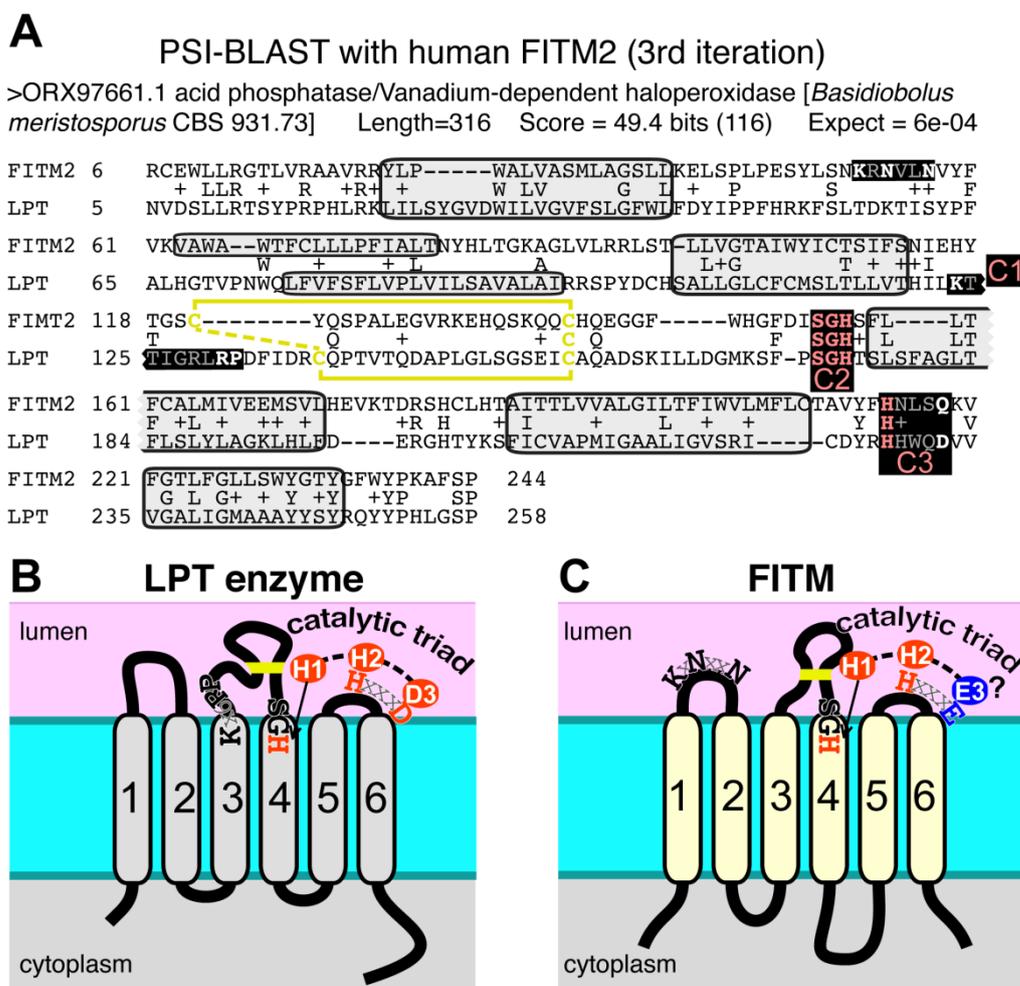
so a single related sequence cannot be identified. Again the pair of FITM2s diverge in predicted ER retention (Table 1C). We conclude that the five newly described FITM2 gene duplications in four non-vertebrate species have all led to divergent localisations for FITMs. This indicates that the divergence between Scs3p and Yft2p is common rather than unique, and that there is wide-spread pressure for having FITMs both inside and outside the ER.

#### Database mining identifies FITM proteins across the whole of eukaryote evolution

Our mining of the NCBI protein database for species with multiple FITM/Scs3p proteins revealed that several protists (for example, *Toxoplasma*) in the stramenopiles alveolates, and rhizaria (SAR) supergroup have been annotated in an automated fashion as having FITM/Scs3p proteins through the presence of domains in Pfam10261 (Table S1). We generated our own alignments using HHsearch [47] executed as HHpred by the “Tuebingen Toolkit” [48]. These searches identified FITM homologues in the protist *Plasmodium falciparum* and in the green alga *Chlamydomonas reinhardtii* (Figure S3). BLAST searches showed that groups of related proteins exist both in protists of the SAR supergroup and in several green algae (Supplementary Figures 2 and 3, and see Figure 6). The conservation is distributed along the whole length of the domain (Figure S3). Therefore, data mining and sequence-based searches with HHsearch and BLAST suggest that FITM was present in the last eukaryotic common ancestor, with selective losses since then, for example in higher plants, which is a wider distribution for FITM than previously appreciated [49].

#### FITM proteins have the key sequence LPT enzymes

Since database searches were informative on FITM evolution, we wondered if they would also provide information



**FIGURE 5: FITM proteins are related to lipid phosphatase/phosphotransferase (LPT) enzymes. (A)** The final significant hit (number 402) from PSI-BLAST with FITM2 after three iterations using the “nr50” database is a fungal LPT annotated as an acid phosphatase. The conserved residues of LPTs occur in three conserved motifs (C1 C2 C3) that are highlighted and colour coded: pink where precisely shared with FITM2 (C2: **SGH**, C3: **HxxxD/E**), white where not shared (C1: **K-x6-RP** in LPT only; **KxNxxN** in FITM2 only; C3: **HxxxD/E**). Variable positions within motifs are shown in gray. Predicted transmembrane domains (TMDs, shaded in light gray) align well in five cases. Paired cysteines likely to form a disulphide bridge in the loop between TMD3 and TMD4 are shown in yellow. **(B)** Topology of a typical LPT enzyme, Dpp1p in yeast, with the three conserved motifs, within which the catalytic triad of the most invariant residues are shown in red. **(C)** Topology of human FITM2, with the two motifs shared with LPTs and the KxNxxN motif. Residues that align with the catalytic triad are coloured red where identical with LPTs, and blue where non-identical.

on the function of FITMs. We therefore looked for proteins with functional annotations that align with FITM2 in sensitive alignments with PSI-BLAST [50]. The alignment to human FITM2 using the full protein database was dominated by >1000 highly related sequences. Therefore, we repeated the search using a non-redundant-50 (nr50) protein database, where groups of sequences >50% identical are reduced to single entries [51]. From the third iteration onwards, this produced significant hits to LPT enzymes (Table S2A). These are integral membrane enzymes in the type 2 phosphatidic acid phosphatase (PAP2, Pfam01569) family, which also includes soluble vanadium-dependent chloroperoxidases in prokaryotes [52]. With further iterations, LPTs became increasingly common in the alignments (Table S2B). Searches into the complete database (i.e. nothing excluded because of high levels of sequence identity) iden-

tified a similar range of LPTs, but the statistical dominance by genomes related to key model organisms (mammals, saccharomycetes) caused alignments with LPTs to be non-significant (data not shown). The FITM-LPT alignment was strongly supported by HHpred, the closest hit being the LPT enzyme Dpp1p, with probability of shared structure = 97% across 207 residues (data not shown). A key aspect of the alignment between FITMs and LPTs is that the residues conserved across the alignment include the LPT active site (Figure 5A) [52]. Although this does not distinguish between convergent and divergent evolution, it indicates that FITMs are functionally related to LPTs.

Which structural features of LPT enzymes are also found in FITM proteins? The LPT family consists of polytopic proteins, typically with 6 TMDs (range 4-8) (Figure 5B) [52]. The catalytic centre of LPTs is always situated in the reduc-

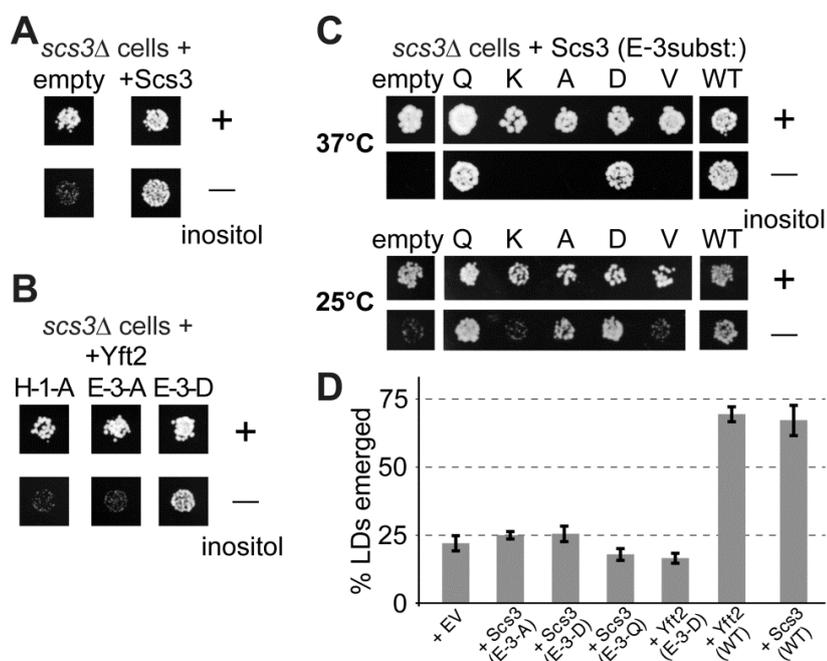


ing environment of the lumen of the secretory pathway or the external milieu, and the active site consists of three motifs called C1, C2 and C3 [53]. C1 is at the N-terminus of the long intra-luminal loop 3/4, which is structured by a disulphide bond, C2 is at the C-terminus of the same loop, and C3 is in the adjacent short loop 5/6 (Figure 5B) [53]. The catalytic triad for LPTs consists of two histidines and one aspartate (here called H-1, H-2, D-3), where the C2 motif contains H-1 typically as “SGH”, and the C3 motif contains H-2 and D-3 precisely spaced HxxxD [52]. Most of these features are found in FITM proteins. They typically have 6 TMDs (Figure 5C) [13], although human FITM1 is predicted to have an extra TMD at its N-terminus (Figure 6). FITMs also have many of the critical elements of the LPT enzymatic centre, including SGH in C2 and a disulphide bond in loop 3/4. However, C3 varies from H-x-x-x-D, as the fifth residue in that sequence is typically glutamate (making H-x-x-x-E). Sometimes this position is glutamine (Q) or H, and rarely D (Table S2C), so that the overall range of residues is similar to LPTs though in different ratios (Table S2D). Finally, C1 which contains K-x6-RP in LPTs is missing from FITMs, but the first main conserved block of FITM sequence (Figure S4) shows that the luminal loop 1/2 contains a conserved “[K]xNØØN” motif, where [K] indicates the predominance of K at that position and Ø is a hydrophobic residue (FILV) (Figure 5C).

We next looked for other primary structural features that specifically relate to possible homology between FITMs and LPT-like enzymes. Firstly, we looked to see if the divergence of ER retention motifs also exists in any group of LPTs that has undergone duplication and divergence. This applies to sphingomyelin synthases (SMS), which have two divergent groups: the active enzymes (SMS1/2) local-

ised in the late secretory pathway [54], and the SMS-related proteins (SMSr) that localise to the ER and act as lipid sensors [55]. SMSr proteins have more of both di-arginine motifs and KKxx motifs than SMS1/2 (Table 1A, lines 8/9), showing a difference in easily discernible motifs that is similar to the difference between Scs3p and Yft2p, although there are more di-arginine motifs in SMS proteins than FITMs. Secondly, we examined if there was any link between predicted cellular location and the residues in the predicted catalytic active site. Of the seven duplications (2 major and 5 minor), five are associated with divergence of a catalytic residue, including FITM1 vs. FITM2 (Figure 6). For *Limulus* where there have been two duplications, the split between FITM2A and FITM2B led to divergence of localisation but not in catalytic residues. By comparison the split between FITM2A1 and FITM2A2 has created in the latter a unique variant (H-1-D) that is likely to inactivate any LPT-like activity. The only duplication with completely conserved catalytic residues is the major one across saccharomycete fungi creating Yft2p and Scs3p. However, here there is variation in the conserved [K]xNØØN motif in loop 1/2, which is present in Yft2s (48/53), but is largely absent from homologues of Scs3p (13/40, Figure S5). Thus, the duplications within this family affect not only localisation, but also the likely architecture of the site related to the LPT catalytic centre.

Overall, many features are shared by FITMs and LPTs, indicating that the FITM family has evolved from the LPTs, and should be included in the overarching LPT family. However, the substitution at D-3 for E-3 calls into question whether FITMs are catalytically active [55]. In addition, the alignment provides no information on possible substrates.



**FIGURE 7: Growth of yeast lacking inositol requires FITMs with an intact catalytic triad. (A-C)** Equal numbers of *scs3Δ* yeast (~50) transformed with the plasmids indicated were spotted onto plates either with added inositol (100 μM) or without, and grown at 30°C or the temperature indicated for 48-72 hours. **(A)** Cells either transformed with empty plasmid (expressing GFP alone) or expressing Scs3-GFP. **(B)** Cells grown in parallel to A, but expressing Yft2-GFP, either mutated H-1-A, E-3-A or E-3-D. **(C)** Cells were expressed with mutants of Scs3-GFP, substituting E-3 to five other residues. Here cells were grown at high stringency (37°C) or low stringency (25°C), the latter allowing rescue by the partially active construct E-3-A. **(D)** Emerged lipid droplets were counted in random sections of *scs3Δ yft2Δ* yeast expressing the indicated constructs (EV = empty vector), and identified as emerged from the ER by the complete absence of wrapping membrane (≥50 cells each experiment, n=3, SD shown by error bars). Emergence, as opposed to remaining wrapped in ER, was scored as previously [26]. This means that the droplets not counted here (*i.e.* not emerged) include both the clearly wrapped, and those with an ambiguous morphology (~7% under all conditions).

### The conserved residues aligning with the catalytic centre of LPTs are functionally important

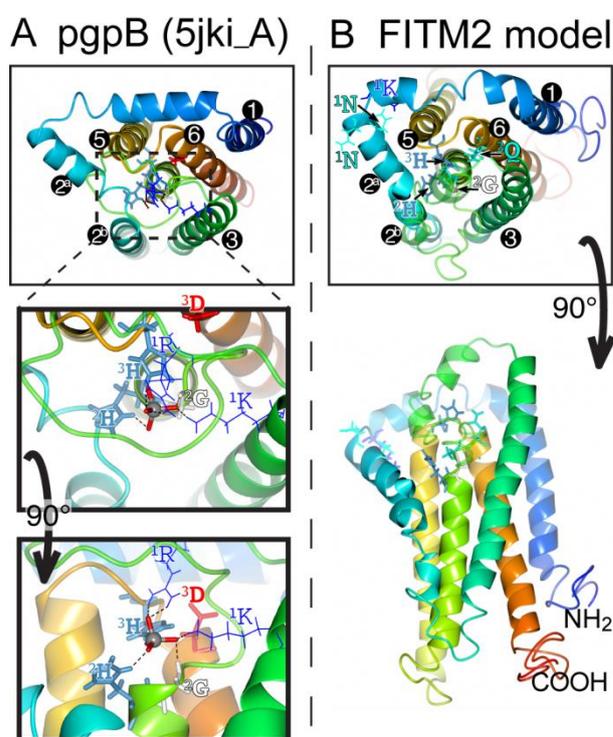
The *SCS3* gene was originally identified through a screen for deletants showing inositol auxotrophy, particularly in the presence of choline [22]. In *scs3Δ* cells inositol supplementation is required more at raised temperature (37°C compared to the usual 30°C) [23]. *yft2Δ* has no significant inositol phenotype, and the double delete *scs3Δ yft2Δ* has the same inositol auxotrophy as *scs3Δ* alone [21]. Re-expression of Scs3p on over-expressing plasmids where the open-reading frame is followed by GFP completely rescued the ability of a *scs3Δ* mutant to grow on inositol-deficient medium (Figure 7A). To test the functional relevance of the alignment of FITM with LPTs, we generated constructs mutated in key enzymatic residues. The role of H-1 was tested in Yft2p, which when over-expressed from the *PHO5* promoter rescued the inositol auxotrophy of *scs3Δ* cells. Substitution of H-1 with alanine (A) abolished the ability of over-expressed Yft2p to rescue inositol auxotrophy (Figure 7B, left-hand panel). We also tested the role of E-3 in some detail. In Yft2p, the substitution E-3-D had no effect, while E-3-A almost completely inhibited growth (Figure 7B). For Scs3p we made a wide range of substitutions of E3. The conservative changes E-3-D and E-3-Q had little effect on activity, while substitution with highly variant residues, either lysine (E-3-K) or valine (E-3-V), abolished activity completely, so that growth was as weak as with an empty plasmid (Figure 7C). By comparison, the neutral substitution E-3-A inhibited growth at 37°C, a temperature for which yeast have a high stringency for Scs3p function [23], but under less stringent conditions (25°C) the E-3-A mutant supported growth (Figure 7C). This indicates that Scs3p with the E-3-A substitution is partially active for supporting growth without inositol.

Alongside the inositol auxotrophy assays, we also determined the expression and localisation of the Scs3p mutant constructs, and found that they are all expressed at similar levels in the ER (Figure S6A/B). This also showed which variants of E3 rescue the formation of ER tangles: E-3-Q and E-3-D rescued like wild-type, but E-3A/V/K did not (Figure S6B). In addition, we studied the ability of plasmid borne Scs3p to rescue GFP-Opi1p relocalization to lipid droplets. Here we found a spectrum of activity where E-3-Q (almost as active as wildtype) was clearly different from E-3-D (only partly active), and E-3-A was inactive (Figure S6C).

Another strong phenotype of yeast lacking both FITMs (*scs3Δ yft2Δ*) is that lipid droplets fail to bud into the cytoplasm, remaining wrapped in ER membrane, a phenotype that is rescued by both Scs3p and Yft2p [26]. Here we found that mutations in E-3 of either protein, even conservative substitutions with Q or D, completely inhibited the rescue of lipid droplet budding (Figure 7D). Overall, these results indicate the necessity of the conserved triad H-1 H-2 E-3 for FITM proteins to function. For each of the four phenotypes we studied, Scs3 with the most conservative substitutions at E-3 tended to retain more activity (activity order Q>D>A>K/V). However, the phenotypes varied considerably in the extent to which any mutation was tolerated: inositol auxotrophy and ER tangles were rescued

fully by two mutants (Q and D), GFP-Opi1 mistargeting was largely rescued by one mutant (Q), and lipid droplet budding was not rescued at all, even by the Q mutant.

We also obtained two sets of negative results. We tested whether FITMs are phosphatidic acid phosphatases, an activity commonly associated with LPT enzymes and already known for two of the seven previously identified yeast LPTs (Lpp1p and Dpp1p). We expressed either Scs3p or Yft2p at high levels (*GAL1/10* promoter, cells using galactose as carbon source) in cells lacking Lpp1, Dpp1p and also the other two known phosphatidic acid phosphatases Pah1p and App1p [56]. Extracts of these cells showed no significant phosphatidic acid phosphatase activity above background (Scs3p: 0.005 U/mg, StdDev = 0.009, n=6;



**FIGURE 8: Structural model of FITM2.** (A) Luminal aspect of bacterial PgpB enzyme (accession 5jki, bottom panel), which is the highest scoring threading template for FITM2 identified by I-TASSER, numbering its 6 TMDs, except the central TMD 4, and with TMD2 split in 2 parts. Zoom in (middle panel) and rotation to side aspect (top panel) show the highlighted elements of the active site, including the tungsten ion (gray sphere) crystallised in place of the phosphate group in 5jki, and the side chains of co-ordinating conserved residues (coloured by residue type) in C1 (1K, 1R = thin bonds), C2 and C3 (2G, 2H, 3H, 3D = fat bonds) [60]. (B) Superimposed view of the modelled structure of human FITM2 with side-chains of the conserved residues in C2 and C3 as in A (bottom panel = luminal aspect, top panel = side aspect), plus the hydrophilic residues (KxNxxN) conserved in loop 1/2 (thin bonds). Models were obtained from the I-TASSER server for 3D structure prediction with standard settings [61]. Several models with seven TMDs were rejected on the basis of topology studies of Yft2p and Scs3p showing six TMDs [13]. In all panels, colouring of the chain graduates from blue (N-terminus) to red (C-terminus).

Yft2p: 0.007 U/mg, SD = 0.016 n=6), i.e. signal 1/1000<sup>th</sup> of wild-type cells (5.2 U/mg, SD = 0.12, n=3). This indicates that FITMs do not have significant phosphatidic acid phosphatase activity. Also, because of the link between FITM2 mutation and mitochondrial dysfunction in humans [9], we examined mitochondrial morphology by EM in *scs3Δ yft2Δ* double-delete cells grown on fermentable sugar to enforce the Krebs cycle, but there was no difference from wild-type (data not shown).

## DISCUSSION

FITM proteins are required to maintain the normal storage of triacylglycerol in many different animal cells. In yeast, this quantity is not affected by FITM deletion [21] instead, a more subtle effect is that FITM proteins are required for correct budding of nascent lipid droplets, a phenotype that is conserved in human cells [26]. This altered budding appears to provide a novel mechanism by which FITMs affect neutral lipid metabolism, but it in turn calls for a molecular explanation for the events that lead to LD budding. Work in yeast has identified another phenotype, a choline-sensitive inositol auxotrophy, which indicates a specific role in phospholipid metabolism. The similarity between FITMs and LPTs we have identified here suggests that FITMs interact with specific phospholipids, in addition to the previous identification of direct binding to non-phospho-lipids (di- and tri-acylglycerol) [12].

The main question raised by our study is whether FITMs are active LPT enzymes. The strongest evidence is in the sequences of FITMs dispersed across evolution. The residues FITMs share with LPTs were already identified as among the most conserved in FITMs [13], and form the core of two of the three main blocks of conserved sequence in FITMs. The FITM block that matches LPTs best includes the C2 motif and “SGH” (Figure 5A). Previously these residues have been part of a wider “FITM signature sequence” that includes most of TMD 4. Within this motif, the glycine (G in “SGH”) is not universal in FITMs, being replaced by S in FITM1, and D in all SARs and algae. These substitutions are not necessarily inactivating for an LPT enzyme, as sphingosine phosphate phosphatases have threonine at this position [52], dolicholpyrophosphatases and plant LPTs mainly have serine [57], and the antibiotic hydrolase RfiA from *Arthrobacter* has leucine [58]. Our Yft2p H-1-A mutant showed the importance of this residue for inositol auxotrophy. Unlike the C2 motif, FITMs lack the C1 motif of LPTs that helps co-ordinate the phosphate (Figure 8A). However, FITMs have a [K]xNØØN motif in the adjacent luminal loop 1/2. A structural model of FITM2 based on the structure of PgpB, a bacterial polytopic LPT, shows that this loop is wrongly placed to co-ordinate the phosphate (Figure 8B). Nevertheless, other active LPTs such as SMSs also lack the C1 motif [52], instead having a family-specific motif in loop 1/2 [54]. We note that a domain of rotaviral VP3 capping enzymes that is also distantly related to LPTs, having SGH and HxxxE/D motifs, also has a NxxN motif [59], so this may have a specific function in phosphotransferase reactions.

This leaves the C3 motif, where the third residue in the

catalytic triad tends to be glutamate (E-3) in FITMs rather than aspartate (D-3) in LPTs. Despite the D-3-E substitution abrogating the activity of some LPTs [55], D-3 is not invariant in LPTs (Table S2). For example, in dolichol pyrophosphate phosphatases this residue is always glutamine. We found that variation of E-3 reduced the activity of yeast FITMs in all the phenotypes we tested. Where the catalytic triad has been mutagenized more extensively, for example in PgpB, H-2 is essential for all activity, but variation in either H-1 or D-3 affects substrates variably [62]. The natural substrate of PgpB is phosphatidylglycerol phosphate, but it can hydrolyse several other lipid phosphate substrates, which are more sensitive to H-1-A and D-3-A mutations than is the natural ligand. Therefore, the variation in effect we detected for the same mutation in different Scs3p-dependent functions (D and Q permissive for inositol auxotrophy and ER morphology, only Q permissive for Opi1p targeting, neither substitution rescuing lipid droplet budding) may indicate that a different range of substrates can fulfill each function. Increasing the severity of mutation (E→Q<D<A<K/V) reveals that some functions of Scs3p have greater substrate specificity than others. The only reaction we have directly tested is the canonical phosphatidic acid phosphatase reaction, where extracts of yeast over-expressing either FITM failed to hydrolyse phosphatidic acid to diacylglycerol and phosphate. Overall, the conserved motifs and specifically positioned residues in FITMs are consistent with them being LPTs, though we have no direct evidence of such activity.

There are groups of pseudo-enzymes within the LPT family, so next we consider if this is the situation for FITMs. The most well-known LPT pseudo-enzymes are the lipid phosphatase-related (LPR) proteins (also called plasticity-related genes, PRGs). These have roles in the handling of extracellular phospholipids such as lysophosphatidic acid, but rather than catalyse a reaction, they seem to sense the lipid through altered ability to bind signaling partners [63, 64]. Consistent with this, H-2 in LPR/PRGs is always substituted (N or S) and some other key residues are variably missing, so that phosphatase activity is likely to be abolished [52]. In FITMs, such major changes are only found in a few FITM1s, particularly in fish. This suggests that the vast majority of FITMs are active LPTs.

The predicted lack of function in the subset of FITM1s with altered key residues is interesting in relation to the dimorphism we found throughout the FITM family. In addition to the FITM2/FITM1 and Scs3p/Yft2p duplications, we mined data on 5 further duplications. All seven events show divergence in either their catalytic motifs or their ER retention motifs (Figure 6). This indicates that there is evolutionary pressure to provide a second intracellular location and/or spectrum of substrates for FITM. In relation to this, residues adjacent to the core catalytic sites can determine specific substrates [65]. For FITMs, the wider definition of C3 is TAØYFHTØØEKØØG, with the F<sup>-1</sup> before HxxxE and K<sup>+5</sup> after HxxxE being highly conserved. This suggests that the minority of FITMs without these residues will have different specificity (e.g. K<sup>+5</sup> missing in SAR/algal sequences). FITM models based on crystal structures of PgpB

(Figure 8) might be used in future to determine what substrates fit into FITMs.

Regarding the di-arginine motifs we have found in FITMs, these have been identified widely, from humans to plants [40-43], but they have not been studied in a system-wide fashion and they are not included in any available intracellular localisation algorithm, such as PSORT [66]. Application of a crude high-throughput method that did not account for predicted topology (hence counting motifs in the lumen) still discerned between ER residents and others, and it showed a divergence between Scs3s and Yft2s. Such short motifs as a pair of arginines separated by 0, 1 or 2 residues might evolve easily, allowing the numbers to vary considerably between closely related proteins; e.g. horseshoe crab (*Limulus*) FITM2A has zero vs. FITM2B has 4 (Figure 6). A further example of apparent diversification of arginine motifs is found in *Candida albicans*. This saccharomycete has a Scs3p/Yft2p pair closely related to the *C. tropicalis* pair (both >50% identical). However, the di-arginine motifs in Scs3p in *C. tropicalis* and most fungi, are swapped in *C. albicans* into Yft2p (Figure 6, dashed arrow). This suggests that divergence between the intracellular localisation of FITM paralogues could be selected independently of which paralogue goes where.

Our results do not provide the explanation for the link between phospholipid and neutral lipid metabolism that must eventually be found, but they do provide some tools for future work. We have studied the role of Scs3 in multiple phenotypes, all linked to lipid metabolism. Is it possible to use these to identify a likely substrate/product of FITMs? Increased Opi1p targeting at an intracellular site might reflect accumulation of phosphatidic acid [30], but matters are more complex. As well as the lipid headgroup, Opi1p shows a marked preference for phosphatidic acid with short acyl chains [67], and even binds to lipid droplets because of other the features altogether related to lipid packing [31]. Thus, none of the phenotypes we studied identify a specific lesion in lipid metabolism.

Given the failure of high-throughput approaches in yeast to report a major lipid imbalance in cells lacking FITMs, we favor the possibility that their function is hard to detect in one of three ways. Firstly, it is possible that FITMs are lipid sensors, similar to SMSr enzymes that use their synthetic activity to make very small amounts of ceramide-phosphoethanolamine in the ER, this small pool having large downstream effects by regulating the much larger ER pool of the ceramide substrate [55]. This mysterious mode of action might only be assayed using purified protein, which could be the subject of future work. The second way that the action of FITMs might be hard to detect is if they carry out a phospho-exchange reaction, swapping a phosphate-containing headgroup between two highly similar lipids. Another LPT that does this is phosphatidylcholine:diacylglycerol cholinephosphotransferase (PDCT) in plant seeds, which swaps phosphocholine between two different diacylglycerols [68]. The biochemical effect of PDCT might be hard to detect because the reaction is pseudo-symmetrical, with substrates and products being identical in terms of headgroup, hence indistinguishable by

thin-layer chromatography, which identifies main lipid classes only. Finally, FITM activity might be hard to detect if the enzyme functions in a confined sub-domain of the ER where lipids cannot freely diffuse, analogous to sub-domains enriched in newly synthesized PC [69]. The strong genetic link in yeast to *SEY1*, the atlastin homologue, might be explained by FITM acting specifically where fusions are about to occur within the ER network [25]. Highly restricted activity may also occur at sites of lipid droplet budding. Future investigations might determine if rapid alterations in FITM activity has immediate effects on the levels of unique lipid species, as identified by mass spectroscopy, which splits the lipidome into thousands of lipid species [70]. A system-wide search for such subtle effects of FITMs may be the key to unlocking their function.

## MATERIALS AND METHODS

### Plasmids

For fluorescence microscopy and inositol auxotrophy assays, Scs3p and Yft2p were expressed from the constitutive portion of the *PHO5* promoter [37] and followed by the linker GTGPVEK then GFP in the plasmid pRS416 (CEN/URA). Constructs were checked by sequencing. Single point mutants were made as follows (and checked by sequencing): Scs3p E-3 (E354) to D, Q, A, K and V; Yft2p H-1 (H178) to A, and E-3 (E243) to D and A. For EM, the same mutations were introduced into plasmids pNO18 and pNO19, which contain the *YFT2* and *SCS3* open-reading frames respectively, with their own promoters in the plasmid YCplac33 (CEN/URA).

### Strains

Wild-type yeast were BY4741, and single deletion strains were obtained from freezer stocks of the yeast deletion collection (BY4741, Mat a, and  $\Delta$ *open-reading-frame::KanMX*). Double delete yeast (*scs3Δ yft2Δ*) were made from the *yft2Δ* single delete strain by PCR knock-out using the *S. pombe* HIS5 gene, as previously [37]. Cells with Erg6 tagged with RFP in the genome were obtained from a collection of strains with marked organelles [71].

### Light microscopy

Yeast growing in log phase were examined with a confocal microscopy system (AOBS SP2; Leica) at room temperature (63× NA 1.4 objective) using LCS software (Leica) for acquisition.

### Electron microscopy

Yeast in mid-log phase were fixed and embedded in Spurr's resin for transmission EM. For ER structure, fixation was in potassium permanganate 1.5% [35], and images were taken at random from low-power fields with high cell density. For lipid droplet visualisation, fixation was in 1% glutaraldehyde and 1.25% paraformaldehyde [26].

### Bioinformatics

#### Basic motifs

Di-arginine motifs (RR, RxR and RxxR) for individual proteins in Figure 6 etc. were mapped onto the topology predicted by TOPCONS [72], and only cytoplasmic motifs were counted. For the high throughput analysis of motifs in large families in Table 1A, the link between motif and predicted position in the

cells was not applied, therefore non-functional motifs (in luminal loops – typically there is one or more in loop 3/4) are included. For C-terminal di-lysine motifs, all instances of KxKxx-stop and KxKxx-stop were included, as well as proven variants RxKxx, KxRxx, HxHxx, KxHxx and RKxx [46].

#### Tree

57 FITM sequences from the organisms in Table 1B and Figure 6 and from representative eukaryotes were aligned by CLUSTAL Omega at EBI [73]. The full alignment of the 57 sequences is available as Supplementary File 1. Then, a tree was generated by submitting the alignment to PHYML 3.0 at [www.trex.uqam.ca](http://www.trex.uqam.ca) [74], with standard settings and branch support by 100 replicates for bootstrapping.

#### Data mining and alignments

The NCBI protein database was searched with the term “scs3p AND (alveolata OR stramenopiles)”. To expand the families of protist and algal FITM homologues, entries in both groups were submitted to standard BLAST at NCBI. To identify distant relatives of FITM proteins, PSI-BLAST was seeded with human FITM2 and used to interrogate the nr50 protein database in the Tuebingen Toolkit for 5 iterations [51]. Alignments were viewed in JALVIEW with Clustal X colours. HHpred at the Tuebingen Toolkit [51] was seeded with FITM sequences and used to search for homologues in PFAM and model eukaryotes. For Figure S3, the hit in *Chlamydomonas reinhardtii* was itself used to seed a reverse search at HHpred, with the width of alignment shown expanded (without changing the statistics of significance) by setting the MAC realignment threshold to 0.1 (default = 0.3).

#### Phosphatidic acid phosphatase assay

Cell extracts were prepared from the *pah1Δ app1Δ dpp1Δ lpp1Δ* mutant overexpressing Scs3p or Yft2p from the *GAL1/10* promoter. Activity was measured by following the release of water-soluble [32P]Pi from chloroform-soluble

[32P]phosphatidic acid as described previously [75]. Units are nmol/min and are expressed per mg cell protein.

#### ACKNOWLEDGEMENTS

We would like to thank Robin Howes and Peter Munro for preparation of EM samples. MH and TL were supported by the Wellcome Trust (grant #082119). VC, NO, and WP were supported by the Intramural Research Program of the National Institute of Diabetes, Digestive, and Kidney Diseases (USA). G-SH and GC were supported by National Institutes of Health (USA) (grant GM028140).

#### SUPPLEMENTAL MATERIAL

All supplemental data for this article are available online at [www.microbialcell.com](http://www.microbialcell.com).

#### CONFLICT OF INTEREST

The authors report that there was no conflict of interest in conducting this study.

#### COPYRIGHT

© 2017 Hayes *et al.* This is an open-access article released under the terms of the Creative Commons Attribution (CC BY) license, which allows the unrestricted use, distribution, and reproduction in any medium, provided the original author and source are acknowledged.

Please cite this article as: Matthew J Hayes, Vineet Choudhary, Namrata Ojha, John JH Shin, Gil-Soo Han, George M. Carman, Christopher JR Loewen, William A Prinz and Timothy P Levine (2017). Fat storage-inducing transmembrane (FIT or FITM) proteins are related to lipid phosphatase/phosphotransferase enzymes. *Microbial Cell* 5(2): 88-103. doi: 10.15698/mic2018.02.614

#### REFERENCES

- Welte MA, Gould AP (2017). Lipid droplet functions beyond energy storage. *Biochim Biophys Acta* 1862(10 Pt B): 1260-1272. doi: 10.1016/j.bbaliip.2017.07.006
- Bates PD, Stymne S, Ohlrogge J (2013). Biochemical pathways in seed oil synthesis. *Curr Opin Plant Biol* 16(3): 358-364. doi: 10.1016/j.pbi.2013.02.015
- Gluchowski NL, Becuwe M, Walther TC, Farese RV, Jr. (2017). Lipid droplets and liver disease: from basic biology to clinical implications. *Nat Rev Gastroenterol Hepatol* 14(6): 343-355. doi: 10.1038/nrgastro.2017.32
- Pires da Rosa G, Libanio D, Filipe Azevedo L (2017). Analysis of the Cochrane Review: Fibrates for secondary prevention of cardiovascular disease and stroke. *Rev Port Cardiol* 36(1): 55-58. doi: 10.1016/j.repc.2016.05.008
- Kadereit B, Kumar P, Wang WJ, Miranda D, Snapp EL, Severina N, Torregroza I, Evans T, Silver DL (2008). Evolutionarily conserved gene family important for fat storage. *Proc Natl Acad Sci U S A* 105(1): 94-99. doi: 10.1073/pnas.0708579105
- Cho YS, Chen CH, Hu C, Long J, Ong RT, Sim X, Takeuchi F, Wu Y, Go MJ, Yamauchi T, Chang YC, Kwak SH, Ma RC, Yamamoto K, Adair LS, Aung T, Cai Q, Chang LC, Chen YT, Gao Y, Hu FB, Kim HL, Kim S, Kim YJ, Lee JJ, Lee NR, Li Y, Liu JJ, Lu W, Nakamura J, et al. (2011). Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. *Nat Genet* 44(1): 67-72. doi: 10.1038/ng.1019
- Kooner JS, Saleheen D, Sim X, Sehmi J, Zhang W, Frossard P, Been LF, Chia KS, Dimas AS, Hassanali N, Jafar T, Jowett JB, Li X, Radha V, Rees SD, Takeuchi F, Young R, Aung T, Basit A, Chidambaram M, Das D, Grundberg E, Hedman AK, Hydrie ZI, Islam M, Khor CC, Kowlessur S, Kristensen MM, Liju S, Lim WY, et al. (2011). Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nat Genet* 43(10): 984-989. doi: 10.1038/ng.921
- Sakai K, Imamura M, Tanaka Y, Iwata M, Hirose H, Kaku K, Maegawa H, Watada H, Tobe K, Kashiwagi A, Kawamori R, Maeda S (2013). Replication study for the association of 9 East Asian GWAS-derived loci with susceptibility to type 2 diabetes in a Japanese population. *PLoS One* 8(9): e76317. doi: 10.1371/journal.pone.0076317
- Zazo Seco C, Castells-Nobau A, Joo SH, Schraders M, Foo JN, van der Voet M, Velan SS, Nijhof B, Oostrik J, de Vrieze E, Katana R, Mansoor A, Huynen M, Szklarczyk R, Oti M, Tranebjaerg L, van Wijk E, Scheffer-de Gooyert JM, Siddique S, Baets J, de Jonghe P, Kazmi SA, Sadananthan SA, van de Warrenburg BP, Khor CC, Gopfert MC, Qamar R, Schenck A,

- Kremer H, Siddiqi S (2017). A homozygous FITM2 mutation causes a deafness-dystonia syndrome with motor regression and signs of ichthyosis and sensory neuropathy. **Dis Model Mech** 10(2): 105-118. doi: 10.1242/dmm.026476
10. Goh VJ, Tan JS, Tan BC, Seow C, Ong WY, Lim YC, Sun L, Ghosh S, Silver DL (2015). Postnatal Deletion of Fat Storage-inducing Transmembrane Protein 2 (FIT2/FITM2) Causes Lethal Enteropathy. **J Biol Chem** 290(42): 25686-25699. doi: 10.1074/jbc.M115.676700.
11. Miranda DA, Kim JH, Nguyen LN, Cheng W, Tan BC, Goh VJ, Tan JS, Yaligar J, Kn BP, Velan SS, Wang H, Silver DL (2014). Fat storage-inducing transmembrane protein 2 is required for normal fat storage in adipose tissue. **J Biol Chem** 289(14): 9560-9572. doi: 10.1074/jbc.M114.547687
12. Gross DA, Zhan C, Silver DL (2011). Direct binding of triglyceride to fat storage-inducing transmembrane proteins 1 and 2 is important for lipid droplet formation. **Proc Natl Acad Sci U S A** 108(49): 19581-19586. doi: 10.1073/pnas.1110817108
13. Gross DA, Snapp EL, Silver DL (2010). Structural insights into triglyceride storage mediated by fat storage-inducing transmembrane (FIT) protein 2. **PLoS One** 5(5): e10796. doi: 10.1371/journal.pone.0010796
14. Wong LH, Levine TP (2017). Tubular lipid binding proteins (TULIPs) growing everywhere. **BBA Molecular Cell Research** 1864(9): 1439-1449. doi: 10.1016/j.bbamcr.2017.05.019
15. Santos-Rosa H, Leung J, Grimsey N, Peak-Chew S, Siniouoglou S (2005). The yeast lipin Smp2 couples phospholipid biosynthesis to nuclear membrane growth. **EMBO J** 24(11): 1931-1941. doi: 10.1038/sj.emboj.7600672
16. Tinkelenberg AH, Liu Y, Alcantara F, Khan S, Guo Z, Bard M, Sturley SL (2000). Mutations in yeast ARV1 alter intracellular sterol distribution and are complemented by human ARV1. **J Biol Chem** 275(52): 40667-40670. doi: 10.1074/jbc.C000710200
17. Kohlwein SD (2010). Obese and anorexic yeasts: experimental models to understand the metabolic syndrome and lipotoxicity. **Biochim Biophys Acta** 1801(3): 222-229. doi: 10.1016/j.bbaliip.2009.12.016
18. Nielsen J (2009). Systems biology of lipid metabolism: from yeast to human. **FEBS Lett** 583(24): 3905-3913. doi: 10.1016/j.febslet.2009.10.054
19. Nguyen LN, Hamari Z, Kadereit B, Trofa D, Agovino M, Martinez LR, Gacser A, Silver DL, Nosanchuk JD (2011). Candida parapsilosis fat storage-inducing transmembrane (FIT) protein 2 regulates lipid droplet formation and impacts virulence. **Microbes Infect** 13(7): 663-672. doi: 10.1016/j.micinf.2011.02.009
20. Young BP, Shin JJ, Orij R, Chao JT, Li SC, Guan XL, Khong A, Jan E, Wenk MR, Prinz WA, Smits GJ, Loewen CJ (2010). Phosphatidic acid is a pH biosensor that links membrane biogenesis to metabolism. **Science** 329(5995): 1085-1088. doi: 10.1126/science.1191026
21. Moir RD, Gross DA, Silver DL, Willis IM (2012). SCS3 and YFT2 link transcription of phospholipid biosynthetic genes to ER stress and the UPR. **PLoS genetics** 8(8): e1002890. doi: 10.1371/journal.pgen.1002890
22. Hosaka K, Nikawa J, Kodaki T, Ishizu H, Yamashita S (1994). Cloning and sequence of the SCS3 gene which is required for inositol prototrophy in *Saccharomyces cerevisiae*. **J Biochem** 116(6): 1317-1321. doi: 10.1093/oxfordjournals.jbchem.a124681
23. Villa-Garcia MJ, Choi MS, Hinz FI, Gaspar ML, Jesch SA, Henry SA (2011). Genome-wide screen for inositol auxotrophy in *Saccharomyces cerevisiae* implicates lipid metabolism in stress response signaling. **Mol Genet Genomics** 285(2): 125-149. doi: 10.1007/s00438-010-0592-x
24. Schuldiner M, Collins SR, Thompson NJ, Denic V, Bhamidipati A, Punna T, Ihmels J, Andrews B, Boone C, Greenblatt JF, Weissman JS, Krogan NJ (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. **Cell** 123(3): 507-519. doi:10.1016/j.cell.2005.08.031
25. Rogers JV, McMahon C, Baryshnikova A, Hughson FM, Rose MD (2014). ER-associated retrograde SNAREs and the Dsl1 complex mediate an alternative, Sey1p-independent homotypic ER fusion pathway. **Mol Biol Cell** 25(21): 3401-3412. doi: 10.1091/mbc.E14-07-1220
26. Choudhary V, Ojha N, Golden A, Prinz WA (2015). A conserved family of proteins facilitates nascent lipid droplet budding from the ER. **J Cell Biol** 211(2): 261-271. doi: 10.1083/jcb.201505067
27. Huston E, Houslay TM, Baillie GS, Houslay MD (2006). cAMP phosphodiesterase-4A1 (PDE4A1) has provided the paradigm for the intracellular targeting of phosphodiesterases, a process that underpins compartmentalized cAMP signalling. **Biochem Soc Trans** 34(Pt 4): 504-509. doi: 10.1042/BST0340504
28. Henry SA, Kohlwein SD, Carman GM (2012). Metabolism and regulation of glycerolipids in the yeast *Saccharomyces cerevisiae*. **Genetics** 190(2): 317-349. doi: 10.1534/genetics.111.130286
29. Loewen CJ, Roy A, Levine TP (2003). A conserved ER targeting motif in three families of lipid binding proteins and in Opi1p binds VAP. **EMBO J** 22(9): 2025-2035. doi: 10.1093/emboj/cdg201
30. Loewen CJ, Gaspar ML, Jesch SA, Delon C, Ktistakis NT, Henry SA, Levine TP (2004). Phospholipid metabolism regulated by a transcription factor sensing phosphatidic acid. **Science** 304(5677): 1644-1647. doi: 10.1126/science.1096083
31. Grippa A, Buxo L, Mora G, Funaya C, Idrissi FZ, Mancuso F, Gomez R, Muntanya J, Sabido E, Carvalho P (2015). The seipin complex Fld1/Ldb16 stabilizes ER-lipid droplet contact sites. **J Cell Biol** 211(4): 829-844. doi: 10.1083/jcb.201502070
32. Han S, Binns DD, Chang YF, Goodman JM (2015). Dissecting seipin function: the localized accumulation of phosphatidic acid at ER/LD junctions in the absence of seipin is suppressed by Sei1p(DeltaNterm) only in combination with Ldb16p. **BMC Cell Biol** 16:29. doi: 10.1186/s12860-015-0075-3
33. Wolinski H, Hofbauer HF, Hellauer K, Cristobal-Sarramian A, Kolb D, Radulovic M, Knittelfelder OL, Rechberger GN, Kohlwein SD (2015). Seipin is involved in the regulation of phosphatidic acid metabolism at a subdomain of the nuclear envelope in yeast. **Biochim Biophys Acta** 1851(11): 1450-1464. doi: 10.1016/j.bbaliip.2015.08.003
34. Jackson CL (2009). Mechanisms of transport through the Golgi complex. **J Cell Sci** 122(Pt 4): 443-452. doi: 10.1242/jcs.032581
35. Loewen CJ, Young BP, Tavassoli S, Levine TP (2007). Inheritance of cortical ER in yeast is required for normal septin organization. **J Cell Biol** 179(3): 467-483. doi: 10.1083/jcb.20070820
36. Zweytick D, Leitner E, Kohlwein SD, Yu C, Rothblatt J, Daum G (2000). Contribution of Are1p and Are2p to steryl ester synthesis in the yeast *Saccharomyces cerevisiae*. **Eur J Biochem** 267(4): 1075-1082. doi: 10.1046/j.1432-1327.2000.011103.x
37. Gatta AT, Wong LH, Sere YY, Calderon-Norena DM, Cockcroft S, Menon AK, Levine TP (2015). A new family of StART domain proteins at membrane contact sites has a role in ER-PM sterol transport. **eLife** 4:e07253. doi: 10.7554/eLife.07253
38. Tkach JM, Yimit A, Lee AY, Riffle M, Costanzo M, Jaschob D, Hendry JA, Ou J, Moffat J, Boone C, Davis TN, Nislow C, Brown GW (2012). Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. **Nat Cell Biol** 14(9): 966-976. doi: 10.1038/ncb2549
39. Vida TA, Emr SD (1995). A new vital stain for visualizing vacuolar

- membrane dynamics and endocytosis in yeast. *J Cell Biol* 128(5): 779-792. doi:10.1083/jcb.128.5.779
40. Boulaflous A, Saint-Jore-Dupas C, Herranz-Gordo MC, Pagny-Salehabadi S, Plasson C, Garidou F, Kiefer-Meyer MC, Ritzenhaller C, Faye L, Gomord V (2009). Cytosolic N-terminal arginine-based signals together with a luminal signal target a type II membrane protein to the plant ER. *BMC Plant Biol* 9:144. doi: 10.1186/1471-2229-9-144
41. Townsley FM, Frigerio G, Pelham HR (1994). Retrieval of HDEL proteins is required for growth of yeast cells. *J Cell Biol* 127(1): 21-28. PMID: 7929564
42. Michelsen K, Yuan H, Schwappach B (2005). Hide and run. Arginine-based endoplasmic-reticulum-sorting motifs in the assembly of heteromultimeric membrane proteins. *EMBO Rep* 6(8): 717-722. doi: 10.1038/sj.embor.7400480
43. Banfield DK (2011). Mechanisms of protein retention in the Golgi. *Cold Spring Harbor perspectives in biology* 3(8): a005264. doi: 10.1101/cshperspect.a005264
44. Gassmann M, Haller C, Stoll Y, Abdel Aziz S, Biermann B, Mosbacher J, Kaupmann K, Bettler B (2005). The RXR-type endoplasmic reticulum-retention/retrieval signal of GABAB1 requires distant spacing from the membrane to function. *Mol Pharmacol* 68(1): 137-144. doi: 10.1124/mol.104.010256
45. Shikano S, Li M (2003). Membrane receptor trafficking: evidence of proximal and distal zones conferred by two independent endoplasmic reticulum localization signals. *Proc Natl Acad Sci U S A* 100(10): 5783-5788. doi: 10.1073/pnas.1031748100
46. Ma W, Goldberg J (2013). Rules for the recognition of dilysine retrieval motifs by coatomer. *EMBO J* 32(7): 926-937. doi: 10.1038/emboj.2013.41
47. Soding J (2005). Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21(7): 951-960. doi: 10.1093/bioinformatics/bti125
48. Alva V, Nam SZ, Soding J, Lupas AN (2016). The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. *Nucleic Acids Res* 44(W1): W410-415. doi: 10.1093/nar/gkw348
49. Cai Y, McClinchie E, Price A, Nguyen TN, Gidda SK, Watt SC, Yurchenko O, Park S, Sturtevant D, Mullen RT, Dyer JM, Chapman KD (2017). Mouse fat storage-inducing transmembrane protein 2 (FIT2) promotes lipid droplet accumulation in plants. *Plant Biotechnol J* 15(7): 824-836. doi: 10.1111/pbi.12678
50. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25(17): 3389-3402. PMID: 9254694
51. Alva V, Lupas AN (2016). The TULIP superfamily of eukaryotic lipid-binding proteins as a mediator of lipid sensing and transport. *Biochim Biophys Acta* 1861(8 Pt B): 913-923. doi: 10.1016/j.bbali.2016.01.016
52. Sigal YJ, McDermott MI, Morris AJ (2005). Integral membrane lipid phosphatases/phosphotransferases: common structure and diverse functions. *Biochem J* 387(Pt 2): 281-293. doi: 10.1042/BJ20041771
53. Stukej J, Carman GM (1997). Identification of a novel phosphatase sequence motif. *Protein Sci* 6(2): 469-472. doi: 10.1002/pro.5560060226
54. Huitema K, Van Den Dikkenberg J, Brouwers JF, Holthuis JC (2004). Identification of a family of animal sphingomyelin synthases. *Embo J* 23(1): 33-44. doi: 10.1038/sj.emboj.7600034
55. Vacaru AM, Tafesse FG, Ternes P, Kondylis V, Hermansson M, Brouwers JF, Somerharju P, Rabouille C, Holthuis JC (2009). Sphingomyelin synthase-related protein SMSr controls ceramide homeostasis in the ER. *J Cell Biol* 185(6): 1013-1027. doi: 10.1083/jcb.200903152
56. Chae M, Han GS, Carman GM (2012). The *Saccharomyces cerevisiae* actin patch protein App1p is a phosphatidate phosphatase enzyme. *J Biol Chem* 287(48): 40186-40196. doi: 10.1074/jbc.M112.421776
57. Wang P, Chen Z, Kasimu R, Chen Y, Zhang X, Gai J (2016). Genome-wide analysis suggests divergent evolution of lipid phosphatases/phosphotransferase genes in plants. *Genome* 59(8): 589-601. doi: 10.1139/gen-2016-0061
58. Chen X, Zhang B, Xiao J, Ju F, Li S, Ren C, An L, Chen T, Liu G, Facey P, Mullins JG, Dyson P (2015). RfiA, a novel PAP2 domain-containing polytopic membrane protein that confers resistance to the FtsZ inhibitor PC190723. *Future Microbiol* 10(3): 325-335. doi: 10.2217/fmb.14.131
59. Ogden KM, Snyder MJ, Dennis AF, Patton JT (2014). Predicted structure and domain organization of rotavirus capping enzyme and innate immune antagonist VP3. *J Virol* 88(16): 9072-9085. doi: 10.1128/JVI.00923-14
60. Ghachi ME, Howe N, Auger R, Lambion A, Guiseppi A, Delbrassine F, Manat G, Roure S, Peslier S, Sauvage E, Vogeley L, Rengifo-Gonzalez JC, Charlier P, Mengin-Lecreulx D, Fogliano M, Touze T, Caffrey M, Kerff F (2017). Crystal structure and biochemical characterization of the transmembrane PAP2 type phosphatidylglycerol phosphate phosphatase from *Bacillus subtilis*. *Cell Mol Life Sci* 74(12): 2319-2332. doi: 10.1007/s00018-017-2464-6
61. Roy A, Kucukural A, Zhang Y (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature protocols* 5(4): 725-738. doi: 10.1038/nprot.2010.5
62. Tong S, Lin Y, Lu S, Wang M, Bogdanov M, Zheng L (2016). Structural Insight into Substrate Selection and Catalysis of Lipid Phosphate Phosphatase PgpB in the Cell Membrane. *J Biol Chem* 291(35): 18342-18352. doi: 10.1074/jbc.M116.737874
63. Liu X, Huai J, Endle H, Schluter L, Fan W, Li Y, Richers S, Yurugi H, Rajalingam K, Ji H, Cheng H, Rister B, Horta G, Baumgart J, Berger H, Laube G, Schmitt U, Schmeisser MJ, Boeckers TM, Tenzer S, Vlachos A, Deller T, Nitsch R, Vogt J (2016). PRG-1 Regulates Synaptic Plasticity via Intracellular PP2A/beta1-Integrin Signaling. *Dev Cell* 38(3): 275-290. doi: 10.1016/j.devcel.2016.06.019
64. Vogt J, Yang JW, Mobascher A, Cheng J, Li Y, Liu X, Baumgart J, Thalman C, Kirischuk S, Unichenko P, Horta G, Radyushkin K, Stroh A, Richers S, Sahragard N, Distler U, Tenzer S, Qiao L, Lieb K, Tuscher O, Binder H, Ferreiros N, Tegeder I, Morris AJ, Gropa S, Nurnberg P, Toliat MR, Winterer G, Luhmann HJ, Huai J, et al. (2016). Molecular cause and functional impact of altered synaptic lipid signaling due to a prg-1 gene SNP. *EMBO Mol Med* 8(1): 25-38. doi: 10.15252/emmm.201505677
65. Kol M, Panatala R, Nordmann M, Swart L, van Suijlekom L, Cabukusta B, Hilderink A, Grabietz T, Mina JGM, Somerharju P, Korneev S, Tafesse FG, Holthuis JCM (2017). Switching head group selectivity in mammalian sphingolipid biosynthesis by active-site-engineering of sphingomyelin synthases. *J Lipid Res* 58(5): 962-973. doi: 10.1194/jlr.M076133
66. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FS (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26(13): 1608-1615. doi: 10.1093/bioinformatics/btq249
67. Hofbauer HF, Schopf FH, Schleifer H, Knittelfelder OL, Pieber B, Rechberger GN, Wolinski H, Gaspar ML, Kappe CO, Stadlmann J, Mechtler K, Zenz A, Lohner K, Tehlivets O, Henry SA, Kohlwein SD (2014). Regulation of gene expression through a transcriptional repressor that senses acyl-chain length in membrane phospholipids. *Dev*

**Cell** 29(6): 729-739. doi: 10.1016/j.devcel.2014.04.025

68. Lu C, Xin Z, Ren Z, Miquel M, Browse J (2009). An enzyme regulating triacylglycerol composition is encoded by the ROD1 gene of *Arabidopsis*. **Proc Natl Acad Sci U S A** 106(44): 18837-18842. doi: 10.1073/pnas.0908848106

69. Drozd MM, Jiang H, Pytowski L, Grovenor C, Vaux DJ (2017). Formation of a nucleoplasmic reticulum requires de novo assembly of nascent phospholipids and shows preferential incorporation of nascent lamins. **Scientific reports** 7(1): 7454. doi: 10.1038/s41598-017-07614-w

70. van Meer G, Voelker DR, Feigenson GW (2008). Membrane lipids: where they are and how they behave. **Nat Rev Mol Cell Biol** 9(2): 112-124. doi: 10.1038/nrm2330

71. Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK (2003). Global analysis of protein localization in budding yeast. **Nature** 425(6959): 686-691. doi: 10.1038/nature02026

72. Tsirigos KD, Peters C, Shu N, Kall L, Elofsson A (2015). The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. **Nucleic Acids Res** 43(W1): W401-407. doi: 10.1093/nar/gkv485

73. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular systems biology** 7: 539. doi: 10.1038/msb.2011.75

74. Boc A, Diallo AB, Makarenkov V (2012). T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. **Nucleic Acids Res** 40(Web Server issue): W573-579. doi: 10.1093/nar/gks485

75. Carman GM, Lin YP (1991). Phosphatidate phosphatase from yeast. **Methods Enzymol** 197: 548-553. doi: 10.1016/0076-6879(91)97182-x

76. Lawrence CE, Altschul SF, Boguski MS, Liu JS, Neuwald AF, Wootton JC (1993). Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. **Science** 262(5131): 208-214. doi: 10.1126/science.8211139