# CHAPTER I:

# INTRODUCTION


## Abstract

Many secretory proteins are thought to rely upon transmembrane cargo receptors for efficient endoplasmic reticulum (ER)-to-Golgi transport. These receptors recognize specific cargo-encoded sorting signals. Only a few such cargo receptors have been characterized in detail, most of them in yeast. The only well-defined cargo receptor from mammalian cells, the LMAN1–MCFD2 complex, is required for the efficient secretion of coagulation factors V and VIII. Studies of this complex, coupled with recent advances in elucidating the basic machinery that mediates ER-to-Golgi transport, have provided a more detailed picture of the mechanisms underlying receptor-mediated transport in the early secretory pathway. In addition to yeast studies, insights have also come from investigations into several inherited disorders that have recently been attributed to defects in the secretory pathway.


## Overview of the early secretory pathway

A substantial proportion of all eukaryotic proteins traverse the secretory pathway en route to their final intra- or extra-cellular destinations. Collectively referred to as 'cargo proteins', this class of proteins is not only limited to soluble secreted proteins but also includes all proteins that will be sorted along the secretory pathway and eventually reside in either the Golgi complex, the lysosomal or endosomal compartments, or the

plasma membrane. Properly folded and assembled proteins exit the endoplasmic reticulum (ER) in coat protein complex-II (COPII) vesicles. The COPII coat is composed of three main subunits (Table 1-1): the Sar1p GTP-binding protein and two heteromeric protein complexes, Sec23p–Sec24p and Sec13p–Sec31p [1] and [2]. Two additional proteins, Sec12p [3] and Sec16p [4], are required for vesicle formation *in vivo*. The cellular distribution of Sec12p, presumably, restricts COPII assembly to the surface of the ER [5], and Sec16p might further limit vesicle formation to specific sites of the ER membrane termed ER-exit sites or transitional ER [6] and [7]. Upon budding, the contents of COPII vesicles are delivered to (or fuse to form) the ER–Golgi intermediate compartment (ERGIC) [8], a structure located between the ER and Golgi that is unique to higher eukaryotic cells. In the ERGIC, the COPII cargo undergoes another round of sorting and quality control. Accessory transport factors and proteins that display ER-retrieval signals are returned to the ER via COPI vesicles [9], where they can participate in subsequent rounds of vesicle formation. Some secretory proteins undergo further concentration in the ERGIC by virtue of being excluded from COPI vesicles [10]. Anterograde cargo proteins leave the ERGIC and are transported to the *cis*-Golgi along microtubules [11] and [12]. Downstream tethering and fusion of vesicles are governed by a series of large protein complexes [13].

**Recruitment of cargo proteins to COPII vesicles**

Although it is widely believed that export by bulk flow might be sufficient for the transport of some highly abundant cargo proteins [10] and [14], evidence is growing in support of a selective model of export in which cargo proteins display specific sorting signals that promote their incorporation into COPII vesicles. Based on current

knowledge, cargo proteins that are selectively recruited to budding vesicles can be divided into two broad categories: (i) those that directly bind to components of the COPII coat and (ii) those that require specific adaptors or receptors to link them to COPII vesicles. The first group is limited to proteins that span the ER membrane because they must be accessible to the cytosolic COPII-coat proteins. Transmembrane cargo proteins that bind COPII can, in turn, serve as cargo receptors for the second class of cargo, which includes both soluble and transmembrane proteins.

Studies in yeast have shown that Sec24p is the major component of COPII that interacts with transmembrane cargo proteins [15], and it contains at least three distinct cargo-binding sites [16] and [17]. *Saccharomyces cerevisiae* also contains two non-essential Sec24p paralogs, Lst1p and Iss1p, which probably increase the diversity of cargos that can be directly recruited to COPII vesicles [16]. There is also evidence that Sar1p participates in the early stages of transmembrane-cargo selection, and interactions between cargo proteins and Sar1p could further restrict vesicle biogenesis to regions of the ER membrane in which cargo proteins are concentrated [18]. COPII proteins are evolutionarily conserved; however, the mammalian genome encodes multiple isoforms of most COPII components (Table 1-1). An even greater diversity in mammalian COPII proteins could be generated by alternative splicing [19]. As a result, the diversity of COPII vesicle types in mammalian cells is likely to exceed that of yeast, with different classes of COPII-coated vesicles functioning in a tissue-specific and cell-type-specific manner to direct selective transport of a wide range of cargo proteins.

**Interactions between COPII and transmembrane cargo**

Several classes of ER export motifs that can be recognized by the COPII coat are currently known [20]. Efficient export of lectin mannose-binding protein 1 (LMAN1; also referred to as ERGIC-53) from the ER requires a Phe-Phe (FF) exit motif at the most C-terminal end of its short cytoplasmic tail. The FF motif seems to function in the context of a larger export signal because it is not sufficient to drive export of a reporter protein. Indeed, efficient export of LMAN1 is dependent upon its oligomerization, which, in turn, relies upon amino acid residues located in its transmembrane and luminal domains [21]. The FF motif is part of a larger class of di-hydrophobic export sequences that are recognized by the Sec23p–Sec24p subunit of COPII and are present in the cytoplasmic tails of a variety of transmembrane cargo proteins including members of the highly conserved p24 family of proteins, which cycle in the early secretory pathway [22], and the ER-vesicle proteins Erv41p and Erv46p [23]. In a recent study, the cytoplasmic tail of LMAN1 exhibited some binding to all four mammalian SEC24 isoforms, and transport of a modified LMAN1 reporter protein was only impaired upon simultaneous knockdown of two or more isoforms by RNAi [24]. By contrast, ER exit of GABA transporter 1 (GAT1) is specifically impaired by RNAi knockdown of SEC24D, and mutational analysis indicates that an Arg-Leu amino acid sequence motif located near the C terminus of GAT1 and two other members of the SLC6 family of neurotransmitter transporters might represent an isoform-specific export signal for SEC24D [25].

A second class of transmembrane cargo proteins interacts with the Sec23p–Sec24p subunit of COPII via a di-acidic Asp/Glu-Xaa-Asp/Glu (where Xaa is any amino acid) motif. The di-acidic ER export motif was first identified by mutational analysis of

the cytoplasmic tail of the vesicular stomatitis virus glycoprotein (VSV-G) [26], and has since been shown to direct ER export of several proteins, including the cystic fibrosis transmembrane conductance regulator (CFTR) [27], mammalian lysosomal acid phosphatase [26], the Kir1.1 and Kir2.1 potassium channels [28], and the yeast proteins Sys1p [29] and Gap1p [30]. In addition, certain transmembrane proteins that contain a C-terminal valine residue might constitute another class of cargo proteins that interacts with the COPII coat [31], [32] and [33].

Although Sec23p–Sec24p is the major COPII subunit involved in direct cargo binding, a recent study uncovered an interaction between Sar1 and a specific class of Golgi glycosyltransferases that is dependent upon a cargo-encoded di-basic Arg/Lys-Xaa-Arg/Lys export motif [34]. It seems likely that other classes of export signals that are recognized by components of the Sar1–Sec23p–Sec24p pre-budding complex also await discovery. In addition, some transmembrane cargo proteins, such as Sec22p, are efficiently packaged into COPII vesicles yet no simple export signals can be identified [35].

**Receptor-mediated transport of soluble and transmembrane cargo**

Because soluble cargo proteins are not directly accessible to the cytosolic COPII-coat proteins, selective recruitment of soluble cargo proteins to COPII vesicles is thought to rely upon interactions between cargo-encoded sorting signals and specific transmembrane adaptors or transport receptors. Some transmembrane cargo proteins also require specific transport receptors. Table 1-2 lists currently known and potential transport receptors and their corresponding cargo proteins.

**Cargo receptors in yeast**

The most well defined example of a specific transport receptor for soluble cargo proteins in the yeast *S. cerevisiae* is the ER-vesicle protein Erv29p. Deletion of *ERV29* leads to a defect in ER-to-Golgi transport of the soluble protein glycosylated pro-α-factor (gpαf) [36]. Recent studies of the interaction between Erv29p and gpαf led to the identification of the first known export signal on a soluble cargo protein that is recognized by its cognate receptor [37]. The signal consists of an Ile-Leu-Val (ILV) sequence motif located in the pro-region of gpαf, which mediates binding to Erv29p and recruitment to COPII vesicles. Furthermore, the ILV signal confers export when fused to an ER-resident protein, Kar2p. Another potential cargo, carboxypeptidase Y, also contains an ILV sequence that might be recognized by Erv29p [37].

Other known cargo receptors in yeast include Erv26p, Erv14p and members of the p24 family. Mutant *erv26Δ* strains exhibit ER accumulation of a precursor form of vacuolar alkaline phosphatase (pro-ALP), whereas other secretory proteins are unaffected. An Erv26p–pro-ALP complex can be detected in cross-linking studies, and an Erv26p-blocking antibody inhibits the incorporation of pro-ALP into COPII vesicles, indicating that Erv26p specifically directs pro-ALP into budding vesicles [38]. Haploid *erv14Δ* cells are defective in axial growth-site selection as a result of inefficient transport of the transmembrane protein Axl2p. Erv14p normally binds to and recruits Axl2p to budding vesicles via a direct interaction with the COPII coat [39]. Mutations in the *Drosophila melanogaster* homolog of *ERV14*, *cornichon* (*cni*), show a selective defect in ER export of a transforming growth factor-α-like growth factor, Gurken [40]. Interestingly, *cni* mutations lead to a cell-polarity defect in developing oocytes that bears

a striking resemblance to the axial-bud defect in *erv14Δ* cells, indicating evolutionary conservation of the receptor–cargo pairing. Diploid *erv14Δ* strains display an additional defect in sporulation that is due to inefficient transport of the transmembrane cargo protein Sma2p [41]. Simultaneous deletion of *ERV14* and a paralogous gene, *ERV15*, exacerbates this phenotype and results in accumulation of Sma2p and other cargo proteins in the ER, indicating that Erv14p and Erv15p might have more general roles in cargo transport during sporulation [41]. The yeast p24 proteins comprise a family of evolutionarily conserved transmembrane proteins that form heteromeric complexes and localize to the early secretory pathway. The strongest evidence in support of a role for p24 proteins as cargo receptors stems from the observation that two members of this family, Emp24p and Erv25p, form a complex that is required for the efficient incorporation of a glycosylphosphatidylinositol-anchored protein, Gas1p, and a soluble cargo, Suc2p (invertase), into COPII vesicles [42]. Furthermore, Gas1p can be crosslinked to Emp24p–Erv25p in COPII vesicles [42].

Emp46p and Emp47p are homologous proteins in yeast that contain a lectin-like carbohydrate- recognition domain (CRD) that shares a similar structure to the CRD of LMAN1, although with different metal ion-binding properties [43]. Emp47p and Emp46p form a heteromeric complex in the ER, and complex formation is required for efficient transport of Emp46p [44]. Therefore, Emp47p is considered to be a cargo receptor for Emp46p, a transmembrane cargo protein. The Emp46p–Emp47p complex might, in turn, function as a transport receptor for soluble cargo proteins because mutant strains lacking Emp47p, Emp46p, or both, display a defect in glycoprotein secretion [44].

Erv41p–Erv46p is another transmembrane-protein complex that is potentially a transport receptor, although no specific cargo proteins have been identified [23] and [45].

**Cargo receptors in mammals**

The best-known mammalian cargo receptor is the LMAN1–MCFD2 (multiple coagulation factor deficiency protein 2) protein complex. A potential role for LMAN1 in selective cargo export from the ER was first indicated by the discovery that mutations in *LMAN1* underlie an inherited bleeding disorder, combined deficiency of factors V and VIII (F5F8D) [46], in which plasma levels of factor V (FV) and factor VIII (FVIII) are reduced to 5–30% that of normal levels [47]. Cell biology studies also identified two lysosomal proteins, cathepsin C (catC) and cathepsin Z (catZ), as cargo proteins for LMAN1 [48] and [49]. It was later discovered that mutations in a novel gene encoding *MCFD2* also give rise to F5F8D [50] and [51]. Because mutations in either gene lead to impaired secretion of FV and FVIII, LMAN1 and MCFD2 have been proposed to function in concert to facilitate transport of FV and FVIII in the early secretory pathway.

**Structure of LMAN1 and LMAN1-MCFD2 interaction**

As described, the cytoplasmic tail of LMAN1 contains a FF ER-exit motif that interacts with the COPII coat (Figure 1-1). Although the majority of LMAN1 is localized to the ERGIC at steady state, it cycles back to the ER via a Lys-Lys (KK) ER-retrieval signal that interacts with the COPI coat. The luminal part of LMAN1 contains a series of four α helices that are predicted to form a coiled-coil domain, and a CRD that binds mannose. The crystal structure of the CRD is similar to Ca2+-dependent leguminous lectins and the ER chaperone calnexin, with two Ca2+-binding sites [52] and [53]. Structural determinants in both the luminal and transmembrane domains lead to

oligomerization of LMAN1 [21], and recent observations under native conditions suggest that LMAN1 exists exclusively in two homo-hexameric forms within the cell, the first composed of six disulfide-linked monomers and the second composed of three disulfide-linked dimers that might associate non-covalently through juxtaposition of the coiled-coil domains [54].

The interaction between LMAN1 and MCFD2 has been characterized by biochemical means, and additional insight has come from studies of mutations occurring in patients with F5F8D. MCFD2 is a 16-kDa soluble protein with two calmodulin-like EF-hand domains. Proper localization of MCFD2 to the ERGIC is LMAN1-dependent [50]. In the absence of LMAN1, the majority of MCFD2 is secreted in an O-glycosylated form and only trace amounts of intracellular MCFD2 can be detected. LMAN1 and MCFD2 form a stable, $Ca^{2+}$-dependent complex with a 1:1 stoichiometry (Figure 1-1), and nearly all endogenous LMAN1 and MCFD2 are in complex with each other in cells [55]. It seems necessary for MCFD2 to bind $Ca^{2+}$ to interact with LMAN1 because missense mutations located within the second EF-hand domain of MCFD2 disrupt the LMAN1–MCFD2 complex [50] and [51].

**Cargo recognition by the LMAN1-MCFD2 complex**

Four cargo proteins are known to require the LMAN1–MCFD2 pathway for efficient ER-to-Golgi transport, namely FV, FVIII, catC and catZ (Figure 1-2). Given the evolutionary conservation and the ubiquitous expression pattern of LMAN1 and MCFD2, additional cargo proteins that are targeted by this complex could also exist. For example, Nicastrin, a component of the γ-secretase complex, and a mutant immunoglobulin heavy chain have also been reported to interact with LMAN1 [56] and [57]. CatC was

identified as a cargo for LMAN1 using cells overexpressing a dominant-negative mutant of LMAN1 that is retained in the ER [48], and catZ was identified as a protein that interacts with LMAN1 in a chemical cross-linking study [49]. Interactions between catC or catZ and LMAN1 were also observed in live cells using a fluorescence complementation assay [58]. Upon treatment with a chemical cross-linker, an estimated 5–20% of FVIII is detected in a tertiary complex with LMAN1 and MCFD2, which provides direct evidence for the receptor activity of the LMAN1–MCFD2 complex. The interaction between LMAN1–MCFD2 and FVIII is also dependent on the intracellular $Ca^{2+}$ concentration (Figure 1-1). An MCFD2 mutant that fails to co-immunoprecipitate with LMAN1 can still bind FVIII (Figure 1-1), indicating that the interaction between MCFD2 and FVIII is, at least in part, independent of LMAN1 [55]. It is not clear whether LMAN1 can bind to FVIII independently of MCFD2. By contrast, MCFD2 seems to be dispensable for the interaction between LMAN1 and catC or catZ [59], indicating that MCFD2 is specifically required for recruitment of FV and FVIII. To date, FV and FVIII are the only cargo proteins known to be affected in patients with F5F8D. However, the levels of catC, catZ and other potential cargo proteins have not been tested. Aside from mild-to-moderate bleeding symptoms, patients with F5F8D do not exhibit any overt clinical phenotypes that could be attributed to deficiencies of other cargo proteins.

A recent study demonstrated the importance of a carbohydrate-based structural motif in the interaction of pro-cathepsin Z (pro-catZ) with LMAN1. Pro-catZ contains two potential sites for N-linked glycosylation. The first high-mannose glycan resides in the context of a β-hairpin loop, presumably forming a conformation-based motif that

accounts for LMAN1-assisted ER export of pro-catZ [60]. Structural modeling identified a similar glycan-β-hairpin-loop motif in catC [60]. The requirement for a conformational LMAN1-binding motif probably provides an additional quality-control mechanism for ER exit of properly folded proteins. At this time, the specific signal(s) that mediates the interactions between the LMAN1–MCFD2 transport-receptor complex and its cargo remain unknown. FV and FVIII are homologous proteins that share a common A1-A2-B-A3-C1-C2 domain architecture. Although the B domains of FV and FVIII share little sequence identity, both are heavily glycosylated [47]. Given the lectin activity of LMAN1, an interaction between LMAN1 and the B domains of FV and FVIII seems a likely possibility. Consistent with this notion, B-domain-deleted FVIII is not efficiently cross-linked to the LMAN1–MCFD2 complex (Figure 1-1). However, substantial cross-linking is still detected when glycosylation of FVIII is blocked [55]. This observation, combined with the essential role of MCFD2 in the secretion of FV and FVIII, indicates that direct protein–protein interactions are also important for LMAN1–MCFD2 cargo-receptor function.

Although cargo proteins are presumably released from their transport receptors upon reaching the ERGIC (Figure 1-2), it is unclear how dissociation occurs. One hypothesis is that the combined effect of lower pH and decreased $Ca^{2+}$ concentration in the ERGIC could trigger cargo release, based on the observation that neutralization of organellar pH impairs the dissociation of pro-catZ from LMAN1 [61]. Although the precise pH of the ERGIC is not known, the Golgi compartment is considerably more acidic than the ER [62].

**Other potential mammalian cargo receptors**

Apart from LMAN1–MCFD2, there are no other currently known specific cargo receptors in mammalian cells. Some of the well-defined cargo receptors in yeast are highly conserved, and the mammalian homologs of these proteins might also function as cargo receptors [63]. However, functions of the mammalian homologs of p24 proteins seem to have diverged considerably. One mammalian p24 family member, transmembrane-trafficking protein 21, was recently discovered to be a component of the presenilin complex that regulates secretase activity and has a role in the trafficking of amyloid precursor protein [64]. In addition, LMAN1 is a member of a larger family of related proteins, the animal L-type lectins, which includes ERGIC-53 (LMAN1), a LMAN1-like protein (LMAN1L), LMAN2 (also referred to as VIP36) and a LMAN2-like protein (LMAN2L; also referred to as VIPL) [65]. LMAN2 and LMAN2L have also been proposed to function in cargo transport [66] and [67]. Another potential cargo receptor is the B-cell receptor-associated protein BAP31, which has been implicated in the selective transport of human class I major histocompatibility complex (MHC) molecules [68]. However, although BAP31 overexpression leads to increased export and cell-surface expression of class I MHC molecules, depletion of BAP31 does not lead to a reciprocal decrease in the number of class I molecules at the cell surface [69]. Therefore, it seems that the function of BAP31 in class I MHC transport is either partially redundant, or perhaps more directed at quality control of subunit assembly and antigen loading. In support of a role in cargo selection, simultaneous depletion of mouse Bap31 and a related protein, Bap29, abrogates the association of mouse class I molecules with a marker of COPII [70].

**Insights from human disorders associated with defects in ER-to-Golgi transport**

Two inherited diseases are attributable to defects in COPII components. Mutations in *SAR1B*, one of two human orthologs of yeast Sar1p (Table 1-1), lead to three related lipid mal-absorption disorders [71]. A common feature of these diseases is the failure of enterocytes to transport dietary lipids, in the form of lipoprotein particles termed chylomicrons, from the luminal surface of the intestine to the basolateral surface, resulting in the accumulation of chylomicron-like particles in the intestinal epithelium. Most mutations in *SAR1B* are missense mutations that affect GDP and GTP binding. However, a splice-site mutation that is expected to disrupt the SAR1B protein structure was found to associate with patients exhibiting neuromuscular symptoms, indicating additional functions of SAR1B *in vivo*. The size of a chylomicron far exceeds that of a typical COPII vesicle [71] and [72]; therefore, SAR1B might increase the flexibility of COPII vesicles to accommodate large cargo. Alternatively, SAR1B could simply recruit the necessary machinery to the ER-exit site of pre-chylomicron particles, in analogy to a recent observation of procollagen-containing vesicles [73], for vesicle budding or later fusion events [72]. An earlier *in vitro* study of pre-chylomicron transport vesicles concluded that COPII is dispensable for ER budding but, instead, is required for Golgi fusion [74]. A missense mutation in a highly conserved amino acid residue of *SEC23A* is the cause of cranio-lenticulo-sutural dysplasia (CLSD) [75]. Children with CLSD exhibit abnormal craniofacial development, skeletal defects and distinctive Y-shaped cataracts. At the cellular level, in fibroblasts from patients homozygous for the *SEC23A* mutation, SEC31 is mislocalized and the ER is grossly dilated, presumably due to an accumulation of secretory proteins that are unable to exit. A build up of procollagen COL1A1 in the

ER is noted in patient-derived skin-fibroblast cultures; however, no difference in COL1A1 secretion is observed [75].

Mutations in the transport protein particle complex subunit (*TRAPPC2*) gene underlie the human disorder X-linked spondyloepiphyseal dysplasia tarda (SEDT) [76], a disease characterized by spine and hip deformations and early-onset osteoarthritis. Although the function of TRAPPC2 (also referred to as SEDL) remains unclear, its yeast homolog, Trs20p, is an essential component of the transport protein particle (TRAPP) tethering complex, which is required for vesicle–vesicle and vesicle–Golgi fusion events [77]. The function of Trs20p seems to be evolutionarily conserved because human TRAPPC2 can complement yeast *trs20Δ* mutants [78]. A similar disease, spondyloepiphyseal dysplasia congenita (SEDC), results from mutations in the type II procollagen gene (*COL2A1*), leading to the speculation that mutations in *TRAPPC2* might specifically disrupt protein–protein interactions that are required for a crucial step in collagen biosynthesis or secretion [79]. Interestingly, the TRAPP complex binds to Sec23p, which could provide a mechanism for selective fusion based on the vesicle coat and contents [80].

**Conclusions**

The enormous diversity and sheer number of cargo proteins that are transported through the secretory pathway raises several interesting questions. Given that so many different proteins are efficiently transported and sorted along this pathway, what proportion exit the ER by default, or bulk flow, and what proportion require specific transport receptors to guide them to their proper destinations? Furthermore, how are specificity and efficiency achieved in a developmental stage- and tissue-specific manner

at the level of the whole organism? Although the fundamental structural and biochemical features of COPII are thought to be evolutionarily conserved, mammalian cells contain multiple isoforms of most COPII components, potentially generating a much more diverse array of COPII vesicles. Surprisingly, mutations in proteins thought to have general functions cause very specific defects in humans. Functional redundancy could explain this phenomenon; however, there is clearly much to learn about how these proteins function *in vivo*.

Several classes of ER-export signals found in membrane proteins that promote COPII–cargo interactions have been delineated. By contrast, the ILV sequence motif displayed by gpαf is the only 'transplantable' sorting signal that has been identified among all known soluble cargo proteins. It is important to note that many of the export motifs that have been studied are not sufficient to yield transport of reporter proteins at wild-type rates, and are instead thought to function in the context of larger cargo-specific sequences. In general, cargo recognition might rely on multivalent interactions and 3D signals that are revealed only upon proper folding and assembly of cargo proteins.

In mammals, the only known example of a specific receptor for soluble cargo proteins is the LMAN1–MCFD2 complex. The requirement of both a transmembrane component and a soluble cofactor is a unique feature, which could represent a paradigm for the organization of other mammalian cargo receptors. In addition, the details of how LMAN1–MCFD2 recognizes, transports and releases its cargo are poorly understood, and represent important aspects for future study. The sorting and continued transport of cargo proteins from the ERGIC to the *cis*-Golgi is also poorly understood, and it will be interesting to learn whether specific cargo receptors exist for anterograde transport from

the ERGIC to the Golgi. To better understand receptor-mediated transport and determine its full range of impact on human disease, it will also be important to identify additional receptor–cargo pairs and elucidate the biochemical and molecular features that underlie their interactions in the early secretory pathway.

**Table 1-1: Evolutionary conservation of COPII.**

| ***Saccharomyces cerevisiae*** | ***Caenorhabditis elegans*** | ***Drosophila melanogaster*** | **Mammals** |
|---|---|---|---|
| *SAR1* | *ZK180.4* | *sar1* | *Sar1a* and *Sar1b* |
| *SEC23* | *sec-23* | *sec23* | *Sec23a* and *Sec23b* |
| *SEC24*, *LST1* and *ISS1* | *sec-24.1* and *sec-24.2* | *sec24* and *CG10882* | *Sec24a*, *Sec24b*, *Sec24c* and *Sec24d* |
| *SEC13* | *npp-20* | *sec13* | *Sec13* |
| *SEC31* | Unknown | *sec31* | *Sec31a* and *Sec31b* |

Designated COPII orthologs in *C. elegans* and *D. melanogaster* are based upon HomoloGene (www.ncbi.nlm.nih.gov) annotation. Abbreviations: *CG10882*, a computed gene with homology to *sec24*; *ISS1*, interacting with *sec16*; *LST1*, lethal with *sec13*; *npp-20*, nuclear pore complex protein; *ZK180.4*, a predicted protein-coding gene with homology to *SAR1*.

**Table 1-2: Confirmed and putative cargo receptors.**

| Receptor | Cargo | Species | References |
|---|---|---|---|
| Erv29p | gpαf and CPY | *S. cerevisiae* | [36] and [37] |
| Erv26p | p-ALP | *S. cerevisiae* | [38] |
| Erv14p | Axl2p and Sma2p | *S. cerevisiae* | [39] and [41] |
| Cornichon | Gurken | *D. melanogaster* | [40] |
| Emp24p–Erv25p | Gas1p and Suc2p | *S. cerevisiae* | [42] |
| Emp46p–Emp47p[a] | Glycoproteins of unknown identity | *S. cerevisiae* | [44] |
| Erv41p–Erv46p[a] | None identified | *S. cerevisiae* and *M. musculus* | [23] and [45] |
| LMAN1–MCFD2[b] | FV and FVIII; catC and catZ | *M. musculus* and *H. sapiens* | [46], [48], [49], [50], [55], [58] and [59] |
| BAP31[a] | MHC class I | *M. musculus* and *H. sapiens* | |

Confirmed cargo receptors meet the following three criteria: (i) contains a transmembrane component and cycles in the early secretory pathway, (ii) evidence that deficiency selectively impairs cargo transport, and (iii) evidence for a specific receptor–cargo interaction. Putative cargo receptors meet two of the three criteria. Abbreviations: Axl2p, axial budding pattern protein; Gas1p, glycophospholipid-anchored surface protein; Sma2p, spore membrane assembly protein; Suc2p, sucrose fermentation protein (also referred to as invertase). [a] Putative cargo receptors. [b] MCFD2 seems to be dispensable for transport of catC and catZ.

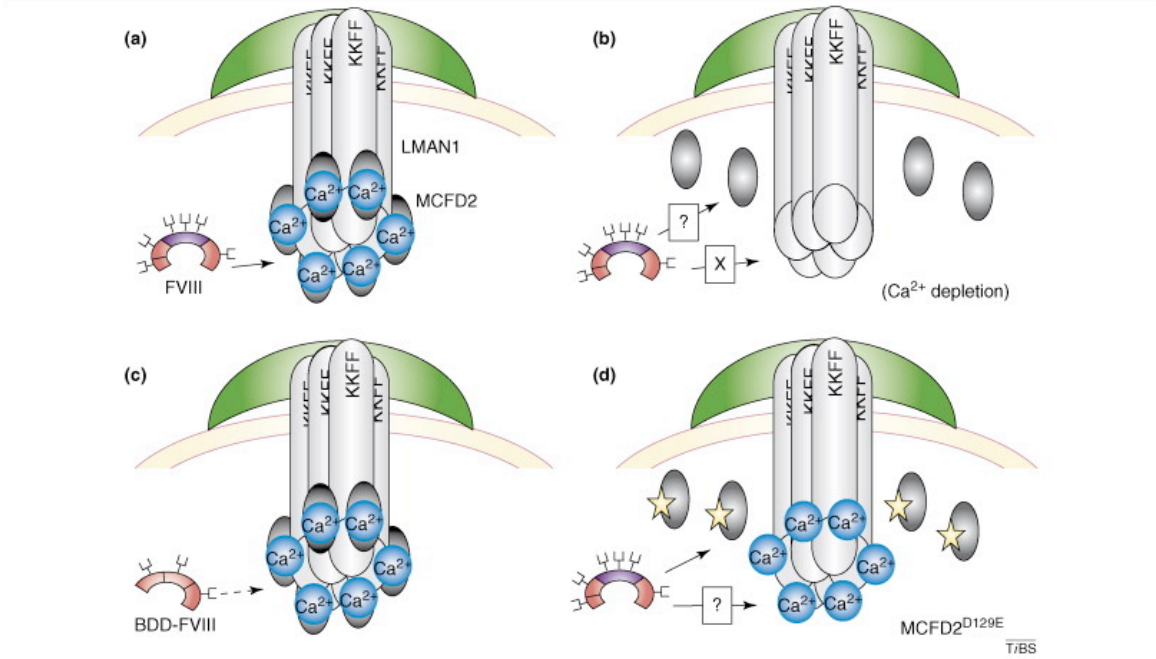


**Figure 1-1: Interaction of FVIII with LMAN1-MCFD2**

FVIII interacts with the LMAN1–MCFD2 transport-receptor complex in the ER lumen. LMAN1 hexamers (gray cylinders) are recruited to budding vesicles via a C-terminal Phe-Phe (FF) motif that binds to COPII (green). The Lys-Lys (KK) motif is an ER-retrieval signal that interacts with COPI. MCFD2 (black ovals) interacts with LMAN1 in a Ca2+-dependent manner, with 1:1 stoichiometry. (a) FVIII interacts with LMAN1 and MCFD2 in a tertiary complex. (b) Depletion of luminal Ca2+ from the ER disrupts the LMAN1–MCFD2 complex and abolishes binding of FVIII. It is not clear whether FVIII can bind to MCFD2 under the condition of luminal Ca2+depletion. (c) Binding of FVIII to LMAN1–MCFD2 is dependent on an intact central B-domain (purple segment). Binding of B-domain-deleted FVIII (BDD-FVIII) binds only weakly to LMAN1–MCFD2. (d) A mutant form of MCFD2 (starred black ovals) that fails to complex with LMAN1 retains its ability to bind FVIII, indicating that the interaction between FVIII and MCFD2 is not dependent upon LMAN1. It is not known whether LMAN1 can bind to FVIII independently of MCFD2.

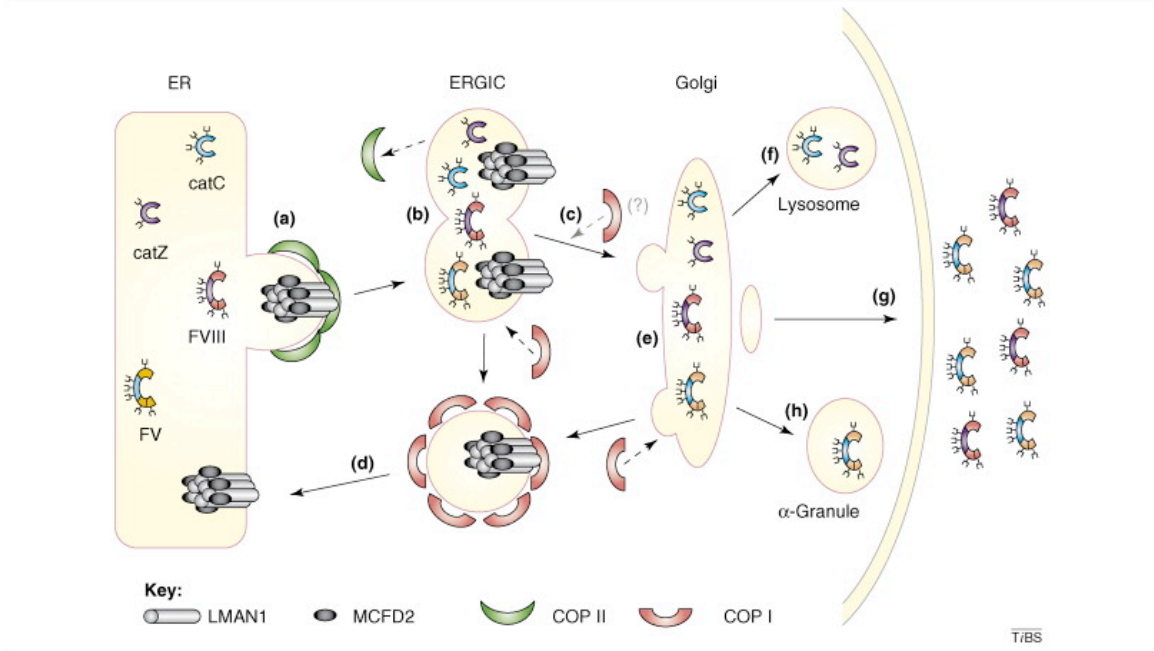


**Figure 1-2: The FV, FVIII, catC and catZ cargo-protein-transport pathways involving the LMAN1–MCFD2 complex.**

(a) Cargo proteins are packaged into COPII vesicles via interactions with the LMAN1–MCFD2 complex. LMAN1 and MCFD2 are both required for the transport of FV and FVIII, whereas MCFD2 seems to be dispensable for the transport of catC and catZ. (b) Vesicles fuse with each other to form the ERGIC, and cargo proteins are released. (c) Anterograde cargo proteins continue to move to the cis-Golgi along microtubules, probably assisted by COPI proteins. (d) The LMAN1–MCFD2 complex is recycled back to the ER from the ERGIC and cis-Golgi in COPI vesicles to participate in another round of cargo transport. (e) Cargo proteins undergo further sorting and modifications in the Golgi. (f) catC and catZ are targeted to the lysosome. (g) FV and FVIII are secreted from cells, with the exception of megakaryocytes, in which (h) FV is packaged into α-granules for storage.

## References


1. M.C. Lee MC et al., Bi-directional protein transport between the ER and Golgi, Annu. Rev. Cell Dev. Biol. 20 (2004), pp. 87–123.

2. C. Gurkan et al., The COPII cage: unifying principles of vesicle coat assembly, Nat. Rev. Mol. Cell Biol. 7 (2006), pp. 727–738.

3. E. Futai et al., GTP/GDP exchange by Sec12p enables COPII vesicle bud formation on synthetic liposomes, EMBO J. 23 (2004), pp. 4146–4155.

4. F. Supek et al., Sec16p potentiates the action of COPII proteins to bud transport vesicles, J. Cell Biol. 158 (2002), pp. 1029–1038.

5. J. Soderholm et al., The transitional ER localization mechanism of Pichia pastoris Sec12, Dev. Cell 6 (2004), pp. 649–659.

6. P. Watson et al., Sec16 defines endoplasmic reticulum exit sites and is required for secretory cargo export in mammalian cells, Traffic 7 (2006), pp. 1678–1687.

7. D. Bhattacharyya and B.S. Glick, Two mammalian Sec16 homologs have nonredundant functions in ER export and transitional ER organization, Mol. Biol. Cell 18 (2007), pp. 839–849.

8. C. Appenzeller-Herzog and H.P. Hauri, The ER–Golgi intermediate compartment (ERGIC): in search of its identity and function, J. Cell Sci. 119 (2006), pp. 2173–2183.

9. J. Lippincott-Schwartz and W. Liu, Insights into COPI coat assembly and function in living cells, Trends Cell Biol. 16 (2006), pp. e1–e4.

10. J.A. Martinez-Menarguez et al., Vesicular tubular clusters between the ER and Golgi mediate concentration of soluble secretory proteins by exclusion from COPI-coated vesicles, Cell 98 (1999), pp. 81–90.

11. K.J. Palmer et al., The role of microtubules in transport between the endoplasmic reticulum and Golgi apparatus in mammalian cells, Biochem. Soc. Symp. 72 (2005), pp. 1–13.

12. P. Watson et al., Coupling of ER exit to microtubules through direct interaction of COPII with dynactin, Nat. Cell Biol. 7 (2005), pp. 48–55.

13. J.R. Whyte and S. Munro, Vesicle tethering complexes in membrane traffic, J. Cell Sci. 115 (2002), pp. 2627–2637.

14. Oprins et al., The ER to Golgi interface is the major concentration site of secretory proteins in the exocrine pancreatic cell, Traffic 2 (2001), pp. 831–838.
15. E. Miller et al., Cargo selection into COPII vesicles is driven by the Sec24p subunit, EMBO J. 21 (2002), pp. 6105–6113.

16. E.A. Miller et al., Multiple cargo binding sites on the COPII subunit Sec24p ensure capture of diverse membrane proteins into transport vesicles, Cell 114 (2003), pp. 497–509.

17. E. Mossessova et al., SNARE selectivity of the COPII coat, Cell 114 (2003), pp. 483–495.

18. M. Aridor et al., The Sar1 GTPase coordinates biosynthetic cargo selection with endoplasmic reticulum export site assembly, J. Cell Biol. 152 (2001), pp. 213–229.

19. M.C. Stankewich et al., Human Sec31B: a family of new mammalian orthologues of yeast Sec31p that associate with the COPII coat, J. Cell Sci. 119 (2006), pp. 958–969.

20. C. Barlowe, Signals for COPII-dependent export from the ER: what's the ticket out?, Trends Cell Biol. 13 (2003), pp. 295–300.

21. O. Nufer et al., ER export of ERGIC-53 is controlled by cooperation of targeting determinants in all three of its domains, J. Cell Sci. 116 (2003), pp. 4429–4440.

22. M. Dominguez et al., gp25L/emp24/p24 protein family members of the cis-Golgi network bind both COP I and II coatomer, J. Cell Biol. 140 (1998), pp. 751–765.

23. S. Otte and C. Barlowe, The Erv41p–Erv46p complex: multiple export signals are required in trans for COPII-dependent transport from the ER, EMBO J. 21 (2002), pp. 6095–6104.

24. M.W. Wendeler et al., Role of Sec24 isoforms in selective export of membrane proteins from the endoplasmic reticulum, EMBO Rep. 8 (2007), pp. 258–264.

25. H. Farhan et al., Concentrative export from the endoplasmic reticulum of the γ-aminobutyric acid transporter 1 requires binding to SEC24D, J. Biol. Chem. 282 (2007), pp. 7679–7689.

26. N. Nishimura and W.E. Balch, A di-acidic signal required for selective export from the endoplasmic reticulum, Science 277 (1997), pp. 556–558.

27. X. Wang et al., COPII-dependent export of cystic fibrosis transmembrane conductance regulator from the ER uses a di-acidic exit code, J. Cell Biol. 167

(2004), pp. 65–74.

28. D. Ma et al., Role of ER export signals in controlling surface potassium channel numbers, Science 291 (2001), pp. 316–319.

29. C. Votsmeier and D. Gallwitz, An acidic sequence of a putative yeast Golgi membrane protein binds COPII and facilitates ER export, EMBO J. 20 (2001), pp. 6742–6750.

30. P. Malkus et al., Concentrative sorting of secretory cargo proteins into COPII-coated vesicles, J. Cell Biol. 159 (2002), pp. 915–921.

31. O. Nufer et al., Role of cytoplasmic C-terminal amino acids of membrane proteins in ER export, J. Cell Sci. 115 (2002), pp. 619–628.

32. F. Paulhe et al., A specific endoplasmic reticulum export signal drives transport of stem cell factor (Kitl) to the cell surface, J. Biol. Chem. 279 (2004), pp. 55545–55555.

33. L.H. Boyle et al., Selective export of HLA-F by its cytoplasmic tail, J. Immunol. 176 (2006), pp. 6464–6472.

34. C.G. Giraudo and H.J. Maccioni, Endoplasmic reticulum export of glycosyltransferases depends on interaction of a cytoplasmic dibasic motif with Sar1, Mol. Biol. Cell 14 (2003), pp. 3753–3766.

35. Y. Liu et al., Sec22p export from the endoplasmic reticulum is independent of SNARE pairing, J. Biol. Chem. 279 (2004), pp. 27225–27232.

36. W.J. Belden and C. Barlowe, Role of Erv29p in collecting soluble secretory proteins into ER-derived transport vesicles, Science 294 (2001), pp. 1528–1531.

37. S. Otte and C. Barlowe, Sorting signals can direct receptor-mediated export of soluble proteins into COPII vesicles, Nat. Cell Biol. 6 (2004), pp. 1189–1194.

38. C.A. Bue et al., Erv26p directs pro-alkaline phosphatase into endoplasmic reticulum-derived coat protein complex II transport vesicles, Mol. Biol. Cell 17 (2006), pp. 4780–4789.

39. J. Powers and C. Barlowe, Erv14p directs a transmembrane secretory protein into COPII-coated transport vesicles, Mol. Biol. Cell 13 (2002), pp. 880–891.

40. C. Bokel et al., Drosophila Cornichon acts as cargo receptor for ER export of the TGFα-like growth factor Gurken, Development 133 (2006), pp. 459–470.

41. H. Nakanishi et al., Erv14 family cargo receptors are necessary for ER exit during sporulation in Saccharomyces cerevisiae, J. Cell Sci. 120 (2007), pp. 908–916.

42. M. Muniz et al., The Emp24 complex recruits a specific cargo molecule into endoplasmic reticulum-derived vesicles, J. Cell Biol. 148 (2000), pp. 925–930.

43. T. Satoh et al., Structures of the carbohydrate recognition domain of Ca2+-independent cargo receptors Emp46p and Emp47p, J. Biol. Chem. 281 (2006), pp. 10410–10419.

44. K. Sato and A. Nakano, Oligomerization of a cargo receptor directs protein sorting into COPII-coated transport vesicles, Mol. Biol. Cell 14 (2003), pp. 3055–3063.

45. L.M. Welsh et al., Genetic and molecular interactions of the Erv41p–Erv46p complex involved in transport between the endoplasmic reticulum and Golgi complex, J. Cell Sci. 119 (2006), pp. 4730–4740.

46. W.C. Nichols et al., Mutations in the ER–Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII, Cell 93 (1998), pp. 61–70.

47. B. Zhang and D. Ginsburg, Familial multiple coagulation factor deficiencies: new biologic insight from rare genetic bleeding disorders, J. Thromb. Haemost. 2 (2004), pp. 1564–1572.

48. F. Vollenweider et al., Mistargeting of the lectin ERGIC-53 to the endoplasmic reticulum of HeLa cells impairs the secretion of a lysosomal enzyme, J. Cell Biol. 142 (1998), pp. 377–389.

49. C. Appenzeller et al., The lectin ERGIC-53 is a cargo transport receptor for glycoproteins, Nat. Cell Biol. 1 (1999), pp. 330–334.

50. B. Zhang et al., Bleeding due to disruption of a cargo-specific ER-to-Golgi transport complex, Nat. Genet. 34 (2003), pp. 220–225.

51. B. Zhang et al., Combined deficiency of factor V and factor VIII is due to mutations in either LMAN1 or MCFD2, Blood 107 (2006), pp. 1903–1907.

52. L.M. Velloso et al., Crystal structure of the carbohydrate recognition domain of p58/ERGIC-53, a protein involved in glycoprotein export from the endoplasmic reticulum, J. Biol. Chem. 277 (2002), pp. 15979–15984.

53. L.M. Velloso et al., The crystal structure of the carbohydrate-recognition domain of the glycoprotein sorting receptor p58/ERGIC-53 reveals an unpredicted metal-binding site and conformational changes associated with calcium ion binding, J.

Mol. Biol. 334 (2003), pp. 845–851.

54. E.P. Neve et al., Oligomerization and interacellular localization of the glycoprotein receptor ERGIC-53 is independent of disulfide bonds, J. Mol. Biol. 354 (2005), pp. 556–568.

55. B. Zhang et al., LMAN1 and MCFD2 form a cargo receptor complex and interact with coagulation factor VIII in the early secretory pathway, J. Biol. Chem. 280 (2005), pp. 25881–25886.

56. V.A. Morais et al., N-glycosylation of human nicastrin is required for interaction with the lectins from the secretory pathway calnexin and ERGIC-53, Biochim. Biophys. Acta 1762 (2006), pp. 802–810.

57. L. Mattioli et al., ER storage diseases: a role for ERGIC-53 in controlling the formation and shape of Russell bodies, J. Cell Sci. 119 (2006), pp. 2532–2541.

58. B. Nyfeler et al., Capturing protein interactions in the secretory pathway of living cells, Proc. Natl. Acad. Sci. U. S. A. 102 (2005), pp. 6350–6355.

59. B. Nyfeler et al., Cargo selectivity of the ERGIC-53/MCFD2 transport receptor complex, Traffic 7 (2006), pp. 1473–1481.

60. C. Appenzeller-Herzog et al., Carbohydrate- and conformation-dependent cargo capture for ER-exit, Mol. Biol. Cell 16 (2005), pp. 1258–1267.

61. C. Appenzeller-Herzog et al., pH-induced conversion of the transport lectin ERGIC-53 triggers glycoprotein release, J. Biol. Chem. 279 (2004), pp. 12943–12950.

62. T.E. Machen et al., pH of TGN and recycling endosomes of H+/K+-ATPase-transfected HEK-293 cells: implications for pH regulation in the secretory pathway, Am. J. Physiol. Cell Physiol. 285 (2003), pp. C205–C214.

63. L. Orci et al., Mammalian Erv46 localizes to the endoplasmic reticulum–Golgi intermediate compartment and to cis-Golgi cisternae, Proc. Natl. Acad. Sci. U. S. A. 100 (2003), pp. 4586–4591.

64. F. Chen et al., TMP21 is a presenilin complex component that modulates gamma-secretase but not epsilon-secretase activity, Nature 440 (2006), pp. 1208–1212.

65. O. Nufer et al., Profile-based data base scanning for animal L-type lectins and characterization of VIPL, a novel VIP36-like endoplasmic reticulum protein, J. Biol. Chem. 278 (2003), pp. 15886–15896.

66. S. Hara-Kuge et al., Involvement of VIP36 in intracellular transport and secretion of glycoproteins in polarized Madin–Darby canine kidney (MDCK) cells, J. Biol. Chem. 277 (2002), pp. 16332–16339.

67. E.P. Neve et al., VIPL, a VIP36-like membrane protein with a putative function in the export of glycoproteins from the endoplasmic reticulum, Exp. Cell Res. 288 (2003), pp. 70–83.

68. Y. Zhang and D.B. Williams, Assembly of MHC class I molecules within the endoplasmic reticulum, Immunol. Res. 35 (2006), pp. 151–162.

69. J.J. Ladasky et al., Bap31 enhances the endoplasmic reticulum export and quality control of human class I MHC molecules, J. Immunol. 177 (2006), pp. 6172–6181.

70. M.E. Paquet et al., Bap29/31 influences the intracellular traffic of MHC class I molecules, J. Immunol. 172 (2004), pp. 7548–7555.

71. B. Jones et al., Mutations in a Sar1 GTPase of COPII vesicles are associated with lipid absorption disorders, Nat. Genet. 34 (2003), pp. 29–31.

72. J.C. Fromme and R. Schekman, COPII-coated vesicles: flexible enough for large cargo?, Curr. Opin. Cell Biol. 17 (2005), pp. 345–352.

73. A.A. Mironov et al., ER-to-Golgi carriers arise through direct en bloc protrusion and multistage maturation of specialized ER exit domains, Dev. Cell 5 (2003), pp. 583–594.

74. S.A. Siddiqi et al., COPII proteins are required for Golgi fusion but not for endoplasmic reticulum budding of the pre-chylomicron transport vesicle, J. Cell Sci. 116 (2003), pp. 415–427.

75. S.A. Boyadjiev et al., Cranio-lenticulo-sutural dysplasia is caused by a SEC23A mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking, Nat. Genet. 38 (2006), pp. 1192–1197.

76. A.K. Gedeon et al., Identification of the gene (SEDL) causing X-linked spondyloepiphyseal dysplasia tarda, Nat. Genet. 22 (1999), pp. 400–404.

77. M. Sacher et al., TRAPP, a highly conserved novel complex on the cis-Golgi that mediates vesicle docking and fusion, EMBO J. 17 (1998), pp. 2494–2503.

78. J. Gecz et al., Human wild-type SEDL protein functionally complements yeast Trs20p but some naturally occurring SEDL mutants do not, Gene 320 (2003), pp. 137–144.

79. A.K. Gedeon et al., The molecular basis of X-linked spondyloepiphyseal dysplasia tarda, Am. J. Hum. Genet. 68 (2001), pp. 1386–1397.

80. H. Cai et al., TRAPPI tethers COPII vesicles by binding the coat subunit Sec23, Nature 445 (2007), pp. 941–944.

## CHAPTER II:

## MAMMALIAN SEC24D IS ESSENTIAL FOR EARLY EMBRYONIC DEVELOPMENT


### Abstract

SEC24 is one of the five core components of COPII vesicles and is believed to be the major subunit of COPII driving cargo selection. Mammals possess multiple isoforms of most COPII components, including four isoforms of SEC24 (SEC24A-D); however, little is known about the functions of the multiple isoforms in the processes of COPII vesicle formation and cargo recruitment. To begin to evaluate the functional contributions of the different mammalian SEC24 isoforms, we generated a mouse model deficient in *Sec24d*. Homozygous null mice lacking *Sec24d* exhibit lethality at an early stage of embryonic development. Mice heterozygous for *Sec24d* exhibit normal growth, development, fertility and survival, and have no gross morphologic abnormalities. Although the mRNAs and proteins for SEC24A-D demonstrate broad and overlapping tissue distributions, the other isoforms of SEC24 appear unable to compensate for loss of SEC24D. These results indicate a distinct and essential functional role for SEC24D in early mouse development.


### Introduction

Approximately one-third of all eukaryotic proteins transit the intracellular secretory pathway *en route* to their final destinations. This includes all secreted,

endosomal/lysosomal and plasma membrane proteins. Newly synthesized proteins destined for these cellular compartments exit the endoplasmic reticulum (ER) in COPII-coated vesicles [1-3]. In yeast, the COPII coat is composed of the small GTP-binding protein Sar1p, and two protein subunits, the heterodimeric Sec23/Sec24p complex and the heterotetrameric Sec13/Sec31p complex (Figure 2-1) [4]. Sar1p generates membrane curvature and initiates vesicle formation by inserting an N-terminal amphipathic helix into the ER membrane [5]. The active membrane-bound Sar1p-GTP recruits Sec23p/Sec24p, and Sec24p drives the selective recruitment of cargo proteins into budding vesicles [6-8]. Polymerization of the outer Sec13p/Sec31p complex completes vesicle budding [9].

In mammals, four *SEC24* orthologs (*SEC24A-D*) have been identified (Figure 2-1). Based on protein sequence identity, the four SEC24 isoforms fall into two subfamilies, a SEC24A/B subgroup and a SEC24C/D subgroup [10]. While deletion of yeast Sec24p is lethal, deletion of either of two non-essential Sec24p homologs, Lst1p and Iss1p, results in specific cargo-transport defects [11-13]. The existence of multiple SEC24 isoforms suggests that a variety of COPII vesicles could be generated in order to expand cargo specificity; however, little is known about the distribution or unique features of the individual mammalian isoforms. Recent studies provide evidence that the SEC24A/B and SEC24C/D subgroups differ in their affinity for a subset of known cargo-sorting signals [14]. However, it remains unclear whether SEC24A-D differ in their tissue- and cell type-specific expression patterns. Moreover, it remains to be determined whether they function in a distinct, overlapping, or redundant manner. In the present study, the observed embryonic lethality of *Sec24d* null mice indicates the existence of at

least some distinct functional role(s) for SEC24D, the loss of which cannot be compensated for by the other isoforms of SEC24.

## Materials and Methods

### Generation of *Sec24d* gene trap mice

A search of the International Gene Trap Consortium database (http://www.genetrap.org) identified three ES-cell clones carrying gene trap vector insertions within the *Sec24d* gene (RRT226, AR803 and RRR785). Clone AR803 was not recoverable from frozen storage. Gene-trapped E14Tg2a.4 ES cell clones RRT226 and RRR785 (Bay Genomics, San Francisco, CA) were cultured as described [15] with the use of ESGRO (Millipore), and expanded cultures were used to prepare RNA and genomic DNA. The RRT226 and RRR785 ES-cells (and mice derived from these cells) are referred to as Line 1 and Line 2, respectively. The gene trap insertion sites located within intron 8 (Line 1) and intron 20 (Line 2) of the *Sec24d* gene were confirmed by reverse transcriptase-PCR (Figures 2-2 and 2-3) using the primers RT1 A-C (Line 1) and RT2 A-C (Line 2). All primers used in this study are listed in Table 2-1. ES cell-mouse chimeras were prepared by blastocyst microinjection as described [16] and bred with C57BL/6 mice to obtain germline transmission. ES-cell derived F1 agouti offspring were genotyped for the presence of the gene trap allele using primers, Neo A and Neo B, to amplify a region of the neomycin cassette. Mice carrying the gene trap allele were maintained by backcrossing to C57BL6/J.

**Mapping of the gene trap vector insertion sites**

The exact location of the gene trap vector insertion sites in intron 8 of Line 1 and intron 20 of Line 2 were mapped using PCR amplification and DNA sequencing. A series of forward primers evenly spaced throughout the intronic sequence (I8F1-21 and I20F1-7) were combined with a reverse primer (1C or 2C) specific to the 5'end of the gene trap vector sequence. Amplicons corresponding to a specific product spanning the insertion site were confirmed by DNA sequencing.

**Genotyping mice by PCR and Southern blot**

Mice from Line 1 were genotyped using a three-primer competitive PCR assay consisting of a common forward primer, (1A) located upstream of the insertion site in intron 8, and two reverse primers, located downstream of the insertion site in intron 8, (1B) or within the gene trap vector (1C). This reaction produces products of different sizes from the wild-type (762 bp) and gene trap (666 bp) alleles, which can be resolved by agarose gel electrophoresis (Figure 2-4). Mice from Line 2 were genotyped in a similar manner, using a common forward primer (2A) located upstream of the insertion site in intron 20, and two reverse primers, located downstream of the insertion site (2B) in intron 20 or within the gene trap vector (2C). This reaction produces a 715 bp product from the wild-type allele and a 558 bp product from the gene trap allele, which can be resolved by agarose gel electrophoresis (Figure 2-5). Genotypes of 33 Line 1 mice were also confirmed by Southern blot analysis, as previously described [17], using a 371 bp probe amplified from C57BL6/J genomic DNA with the primers ApaI A and ApaI B. The probe was hybridized to ApaI-digested genomic DNA, as previously described [17]. Hybridization to the wild-type allele detected a 4.7kb ApaI restriction fragment, whereas

hybridization to the gene trap allele detected a 3.0kb ApaI restriction fragment. In mice heterozygous for the Line 1 gene trap allele, both restriction fragments are present (Figure 2-6).

**Timed mating**

Timed matings were performed by intercrossing *Sec24d* heterozygous mice. Embryos were harvested at day E10.5-11.5 for genotyping and histological analysis, or at E13.5 for the preparation of mouse embryonic fibroblasts. Mouse embryonic fibroblasts were isolated as previously described [18]. For blastocyst collection, female *Sec24d+/-* mice were superovulated by intraperitoneal injection of 0.5 IU pregnant mare serum gonadotropin (PMSG) on day 1 and 0.5 IU human chorionic gonadotropin (HCG) on day 3. Females were then mated with *Sec24d+/-* males, and copulation plugs were noted on day 4, counted as day E0.5 of embryonic development. Blastocysts were harvested by the University of Michigan Transgenic Core on day 7 (E3.5), as previously described [19], and crude lysates were genotyped by three-primer PCR.

**Tissue histology and beta-galactosidase staining**

An autopsy and histological survey was performed on adult *Sec24d+/-* mice. Tissues were submitted to the University of Michigan Histology Core for sectioning and H&E staining. Adult mouse tissue and embryos from timed matings between *Sec24d*+/- mice were fixed and stained for beta-galactosidase activity using a commercial kit (Specialty Media). Embryos and tissues were fixed for 45 minutes and 1 hour, respectively, on wet ice. Embryos were stained at 37ºC for several hours to overnight. Tissues were stained overnight at 30ºC on a rocking platform and the staining solution was supplemented with 0.02% Nonidet P-40 to improve tissue penetration.

**Electron microscopy**

Liver and pancreatic tissue were excised and emersion fixed in 2.5% glutaraldehyde in 0.1 M Sorensen's buffer (0.1M $Na_2HPO_4$, 0.1M $KH_2PO_4$), pH 7.4, overnight at 4ºC. Post-processing and transmission electron microscopy (TEM) was performed at the University of Michigan Microscopy & Image Analysis Laboratory. Tissue samples were rinsed in Sorenson's buffer and post-fixed in 1% osmium tetroxide in 0.1M Sorensen's buffer. Samples were rinsed in double-distilled water to remove phosphate and stained *en bloc* with aqueous 3% uranyl acetate for 1 hour. Stained tissues were dehydrated in ascending concentrations of ethanol, treated with propylene oxide, and embedded in Epon epoxy resin. Semi-thin sections were stained with toluidine blue for tissue identification. Ultra-thin 70nm sections were prepared from regions of interest and post stained with uranyl acetate and lead citrate.

Preparation of mouse embryonic fibroblasts for TEM was carried out as follows: cell monolayers grown to about 80% confluency were rinsed in serum free Dulbeccos' modified eagle medium (Invitrogen) and fixed for 1 hour at room temperature in 2.5% glutaraldehyde in 0.1 M Sorensen's buffer, pH 7.4. The samples were rinsed in buffer and post-fixed for 15 minutes in 1% osmium tetroxide in Sorenson's buffer. Fixed cells were scraped from the culture dishes and pelleted in 1.5 ml microfuge tubes. The samples were rapidly dehydrated by resuspending and pelleting in a graded series of ethanol, infiltrated and embedded in Epon, and polymerized. Ultra-thin sections were collected onto copper grids. The grids were post stained with uranyl acetate and lead citrate.

TEM was performed using a Philips CM100 electron microscope at 60 kV. Images were recorded digitally using a Hamamatsu ORCA-HR digital camera system operated by AMT software (Advanced Microscopy Techniques Corp.).

**Blood collection and CBC analysis**

Wild-type and *Sec24d* heterozygous littermates were anesthetized briefly with isofluorane and approximately 25 microliters of blood was collected from the retro-orbital plexus into citrate-coated anticoagulant glass capillary tubes. A complete blood count was performed on an Advia120 whole blood analyzer (Bayer).

**Real-time RT-PCR and western blotting**

Total RNA was isolated from a panel of tissues from two-month old male C57BL6/J mice using Trizol reagent (Invitrogen), followed by DNaseI digestion and purification using an RNeasy kit (Qiagen). First-strand cDNA synthesis was carried out using a Superscript kit (Invitrogen). Real-time PCR was performed on an ABI HT7300 machine using a 2x SYBR Green Master Mix (Applied Biosystems) and primers specific for each Sec24 isoform (pga.mgh.harvard.edu/primerbank/). Western blot analysis of mouse tissues was performed as previously described [20]. Briefly, tissues were homogenized in ice-cold lysis buffer (100mM Tris, pH 7.5, 1% NP-40, 0.1% Deoxycolate, 130 mM NaCl, 5mM MgCl2, 1mM $Na_3VO_4$, 10 mM NaF) supplemented with complete protease inhibitor (Roche). Cell lysates were cleared by centrifuging for 10 minutes at maximum speed in a bench-top centrifuge and passing through a 0.22 μM filter. After normalizing protein concentrations by Bradford assay (BioRad), 20ug of protein from each tissue was separated by SDS-PAGE and transferred to nitrocellulose (BioRad). Western blotting was performed with the indicated antibodies: a commercially

available antibody that recognizes both SEC23A and SEC23B (Sigma) used at a 1:1000 dilution, and anti-peptide antibodies raised against each of the four SEC24 isoforms (courtesy of Xiao-Wei Chen) used at a 1:500 dilution. An anti-tubulin antibody (Sigma) was used at a dilution of 1:2000 to assess equal protein loading. A goat anti-rabbit IgG HRP-conjugated secondary antibody (Pierce) was used at a 1:10,000 dilution, and HRP was detected using an enhanced chemiluminescence kit (Perkin-Elmer).

**Sequence alignments**

Protein sequence alignments were generated by the ClustalX multiple sequence alignment program (http://www.ebi.ac.uk/Tools/clustalw2/) using the Gonnet series protein weight matrix.

**Results**

**Generation of *Sec24d* null mice**

SEC24D is a 1032 amino acid protein encoded by a gene spanning 98,141 kilobases of genomic DNA sequence and 23 exons (Figure 2-7). Two ES-cell clones (RRT226 and RRR785, referred to as Line 1 and Line 2, respectively) carrying gene trap vector insertions in intron 8 (Line 1) or intron 20 (Line 2) of the *Sec24d* gene were obtained from the International Gene Trap Consortium (http://www.genetrap.org). Identity of these two gene trap clones was confirmed by reverse transcriptase-PCR amplification of products specific for the junction between the intron and the gene trap vector (Figures 2-2 and 2-3). Fine mapping of the gene trap vector insertion sites within Line 1 and Line 2, carried out by PCR amplification and DNA sequencing, showed that the gene trap vector insertion occurred at position 3378 of intron 8 (numbering from the start of the intron) in Line 1 (Figure 2-2) and position 639 of intron 20 in Line 2 (Figure

2-3). Injection of Line 1 and Line 2 ES-cells into C57BL6/J donor blastocysts gave rise to chimeric males, which were bred to C57BL6/J females. Germline transmission of the gene trap allele was achieved for both clones. ES-cell-derived offspring were identified by their 129P2-derived agouti coat color and all agouti pups were subsequently genotyped with primers specific for the neomycin cassette within the gene trap vector to determine the presence (Neo+) or absence (Neo-) of the gene trap allele.

In an attempt to generate *Sec24d* null animals, intercrosses were performed with Neo+ F1 agouti mice from Line 1. The expected Mendelian distribution of genotypes for a heterozygous intercross is 25% homozygous wild-type (*Sec24d*+/+), 50% heterozygous (*Sec24d*+/-) and 25% homozygous null (*Sec24d*-/-). However, when 119 F2 offspring of these crosses were genotyped at weaning, no *Sec24d*-/- mice were observed ($p < 2 \times 10^{-11}$) (Table 2-2). To determine when during development that the *Sec24d*-/- mice were being lost, four litters of Line 1 mice were harvested at mid-embryogenesis (E10.5-11.5) and twenty-eight embryos were genotyped at the *Sec24d* locus. Again, no homozygous null embryos were present at this stage of development ($p < 3 \times 10^{-3}$) (Table 2-2). Timed matings were performed again, and this time, pre-implantation embryos were harvested at the blastocyst stage (E3.5). Although no *Sec24d*-/- blastocysts were observed out of a total of twenty-seven ($p < 0.01$), seventeen 8-cell stage embryos were harvested alongside the blastocysts, and one of these 8-cell stage embryos was observed to be *Sec24d*-/-, confirmed by repeat genotyping (Table 2-2). These results indicate that deficiency of *Sec24d* in Line 1 is lethal at an early stage of embryonic development, probably prior to the blastocyst stage.

To confirm these results, a second independent line of *Sec24d* null mice were generated from the RRR785 (Line 2) gene trap ES-cell clone. Intercrosses of Neo+ F1 agouti mice generated from the Line 2 gene trap clone yielded a complete absence of *Sec24d-/-* mice when 88 offspring of the F2 generation were genotyped at 2 weeks of age ($p < 5 \times 10^{-7}$) (Table 2-3). Based on our previous findings from Line 1, timed matings of *Sec24d* heterozygous mice from Line 2 were set up to collect blastocyst-stage embryos for genotyping. However, in contrast to the Line 1, genotyping of 99 blastocysts from Line 2 intercrosses yielded *Sec24d+/+*, *Sec24d+/-* and *Sec24d-/-* embryos in the expected Mendelian ratio of 25:50:25 (Table 2-3).

**Excess of *Sec24d* heterozygous mice in Line 1**

The expected genotype distribution for an F1 intercross is 25:50:25; however, if all homozygous null embryos are lost, we would expect a ratio of 33:67 for the remaining two genotypes. In the F2 generation (offspring of F1 intercross) of Line 1, a statistically significant deviation from the expected 33:67 (homozygous wild-type:heterozygous) genotype distribution was observed (Table 2-2). Out of 119 F2 mice, 24 (20.2%) were *Sec24d+/+* and 92 (79.8%) were *Sec24d+/-*, which represents a statistically significant deviation from the expected ratio ($p < 4 \times 10^{-3}$). Repeat PCR genotyping performed on all F2 mice resulted in 100% concordance with the original genotype assignments. To further exclude the possibility of genotyping error, Southern blot analysis was performed to confirm the genotypes of 33 (5 *Sec24d+/+* and 28 *Sec24d+/-*) out of 119 F2 mice (Table 2-2 and Figure 2-6). Again, all genotypes ascertained by Southern blot agreed with the original results from PCR genotyping.

To determine whether the excess of *Sec24d* heterozygous mice would persist over multiple generations, we performed backcrosses of Line 1 to the C57BL6/J strain. Although a statistically significant excess of *Sec24d*+/- mice was observed in three generations, it was not present in consecutive generations (Table 2-4). The total numbers summed over all backcross generations show a trend towards an excess of heterozygous mice, but do not reach a statistically significant deviation from the expected 50:50 ratio. Therefore, the excess of *Sec24d* heterozygous mice in Line 1 does not seem to be consistent from one generation to the next.

**Phenotypic analysis of *Sec24d* heterozygous mice**

*Sec24d*+/- mice from both gene trap lines are viable, fertile and exhibit no obvious phenotypic abnormalities. An autopsy and histological analysis failed to identify any gross defects (data not shown). Routine CBC analysis showed no significant difference in any parameter between *Sec24d*+/- mice and *Sec24d*+/+ littermates from Line 1 (Table 2-5). Electron microscopy of pancreas and liver tissue from *Sec24d*+/- mice failed to show a consistent dilation of the ER or any other obvious defects in the cellular organization or composition of pancreatic islets (Figure 2-8), acinar cells (Figure 2-9), or hepatocytes (Figure 2-10) compared to littermate *Sec24d*+/+ controls. Likewise, electron microscopy of mouse embryonic fibroblasts derived from *Sec24d*+/- and *Sec24d*+/+ embryos failed to demonstrate any consistent difference in the degree of ER dilation (data not shown).

**Analysis of *Sec24d* expression pattern**

The gene trap vector insertion in Line 1 and Line 2 is designed to lead to splicing of the upstream exon to a strong splice acceptor in the vector sequence, resulting in a

transcript, which encodes a fusion protein consisting of the 5' portion of SEC24D fused in-frame to a beta-galactosidase/neomycin (β-geo) reporter cassette [21]. The gene trap fusion proteins from the Line 1 and Line 2 alleles are expected to contain the amino acids encoded by the first 8 exons (roughly the N-terminal third of the protein) or first 20 exons of *Sec24d*, respectively, fused to the β-geo reporter protein. Embryos from timed matings between *Sec24d*+/- mice were harvested at E10.5 and stained for β-galactosidase activity. For comparison, E10.5 embryos from a timed mating between *Sec24a*+/- mice (also generated in our laboratory) were also harvested and stained for β-galactosidase activity. While the expression pattern of SEC24A appears diffuse throughout the embryo, SEC24D expression is more restricted, with high levels of β-galactosidase activity detected in the heart, brain and spinal cord region (Figure 2-11). To determine whether the level and/or pattern of *Sec24a-d* expression might vary during later stages of development, expression of *Sec24a-d* mRNA was analyzed by real-time RT-PCR of total RNA isolated from embryos at E10.5, E14 and E18 available from the University of Michigan Center for Organogenesis (Figure 2-12). *Sec24d* expression was already apparent at E10.5, as was expression of the other three isoforms of *Sec24* (Figure 2-12). Whole-animal expression of all four *Sec24* isoforms appeared to remain at a relatively constant level throughout mid- to late-embryogenesis.

The expression pattern of SEC24 in adult mouse tissues was analyzed by real-time RT-PCR and western blot. *Sec24d* mRNA was detected in total RNA isolated from mouse liver, heart, lung, spleen, pancreas, muscle, kidney and brain, with the highest level of expression in the heart and liver (Figure 2-13). *Sec24a* and *Sec24b* were also broadly expressed, and *Sec24c* appeared to have the most tissue-specific expression

pattern. Due to moderate degradation in RNA samples prepared from the pancreas and spleen (Figure 2-13), the real-time RT-PCR signals from these samples may be artificially low; however, the signals from the *18S rRNA* control sample did not seem to be greatly affected (Figure 2-13). Western blot analysis also showed a broad, overlapping distribution of SEC24 isoforms in adult mouse tissues (Figure 2-14), with the highest levels of SEC24D detected in brain, liver, pancreas and lung.

**Sequence comparison of SEC24 isoforms across species**

The four isoforms of SEC24 fall into two subfamilies on the basis of protein sequence identity. Although divergence of the SEC24A/B and SEC24C/D subfamilies precedes the vertebrate-ascidian split, divergence within each pair occurred later, after the vertebrate-ascidian split (Alex Kondrashov, personal communication). Clear orthologs of all four SEC24 isoforms are present in zebrafish, whereas flies and worms are limited to two isoforms. Multiple sequence alignment of the mouse SEC24 proteins demonstrates a hypervariable N-terminal region (Figure 2-15) and a higher degree of sequence conservation in the C-terminal half of the protein (Figure 2-16), a region containing several predicted conserved domains (Figure 2-7).

**Discussion**

The *Sec24d* null mice described in this report were generated using gene trap ES-cell clones. Splicing of the upstream exons (exons 1-8 in Line 1 and exons 1-20 in Line 2) to the gene trap reporter is expected to generate a truncated SEC24D fusion protein that would function as a null allele of *Sec24d*. However, the possibility remains that mice generated from Line 1 and/or Line 2 could represent incomplete null alleles. If viable *Sec24d-/-* embryos were present in Line 1, RT-PCR analysis could be performed to

determine whether any residual normal *Sec24d* mRNA is produced from the gene trap allele; however, we could not obtain *Sec24d-/-* embryos due to their very early embryonic lethality. If this type of analysis were performed on samples from *Sec24d+/-* animals, it would be difficult to distinguish between mRNA produced from the wild-type allele and residual normal transcripts that may originate from the gene trap allele. One possibility would be to take advantage of the mouse strain differences from which the two alleles originated. A coding SNP between C57BL6/J (the wild-type allele) and 129P2 (the knockout allele) would allow us to distinguish between mRNA produced from each allele to facilitate a quantitative analysis. If there is a small amount of residual full-length mRNA being transcribed and translated in Line 1, the levels are not high enough to promote survival of *Sec24d-/-* embryos beyond the 8-cell embryo stage of development. It is possible that SEC24D is absolutely required for survival at the level of a single cell; however, residual maternal SEC24D from the oocyte might also allow the fertilized embryo to survive through several rounds of cell division.

The discrepancy between the two independent gene trap lines, with *Sec24d-/-* blastocysts being absent in Line 1, but present in expected numbers in Line 2, suggests that the gene trap allele in Line 2 may indeed be a hypomorphic allele that retains some function. Two possible scenarios could account for the presence of residual *Sec24d* function in homozygous null embryos from Line 2. First, a low level of normal pre-mRNA splicing could occur around the gene trap vector sequence, producing sufficient amounts of normal *Sec24d* mRNA and allowing the embryos to survive past the 8-cell stage. Second, the SEC24D/ β-geo fusion protein could retain some normal function, especially given that the Line 2 fusion transcript contains the first 20 of 23 exons of the

*Sec24d* coding sequence. Indeed, while the Line 1 allele results in truncation prior to the first of five conserved domains, preserving only the hypervariable N-terminal region, the Line 2 allele results in truncation of SEC24D just before the last of the five conserved domains, leaving the majority of SEC24D intact (Figure 2-7). It is also important to consider that the SEC24D/β-geo fusion proteins produced from the gene trap alleles could, in theory, exert a dominant-negative effect; however, there is no evidence for a detrimental effect of the fusion proteins in *Sec24d* heterozygous mice from either line.

In F1 intercrosses of heterozygous mice from Line 1, we observed a statistically significant excess of heterozygous animals that was initially puzzling. One possible explanation for the observed excess of heterozygous mice in Line 1 would be that the *Sec24d* heterozygous genotype provides a selective advantage to mice of this genotype. While it is difficult to envision a scenario in which a 50% reduction in SEC24D would result in a selective advantage over wild-type, a detrimental allele in C57BL/J closely linked to *Sec24d* could provide an explanation. The *Sec24d* gene trap allele was originally propagated in ES-cells derived from the 129P2 strain, and the chimeras and all resulting offspring were bred to the C57BL6/J strain. Therefore, in mice selected by genotyping to be heterozygous for the gene trap allele, the region at or around the *Sec24d* knockout allele will be ES-cell derived 129P2 sequence (the 'congenic' interval, which will narrow over repeated backcross generations) and the corresponding segment on the homologous chromosome will be derived from the pure C57BL6/J strain. If a detrimental recessive C57BL6/J allele were located nearby the *Sec24d* locus, then mice heterozygous for 129P2 and C57BL6/J alleles could have a selective advantage. If this hypothesis were correct, then the same skewing towards an excess of heterozygous mice

should have been observed in the second independent *Sec24d* gene trap line (Line 2). Although the genotype distributions for the F2 and N2 generations of mice from Line 2 did not show a significant deviation from the expected ratio of 1/3:2/3 ($\chi^2 = 0.8796$), the number of mice analyzed may not have been large enough to distinguish the observed ratio from the ~1/4:3/4 ratio that was observed in Line 1. Given these results, the observed excess of heterozygous mice is most likely an artifact unique to Line 1, perhaps a spontaneous mutation in the original ES-cell line that is linked to *Sec24d*.

Because of the critical role of SEC24 in the early secretory pathway, we focused our histological analysis on tissues with a high demand for protein secretion. Disruption of COPII export is expected to lead to a build-up of cargo proteins, which may result in dilation of the endoplasmic reticulum that can be visualized at the ultrastructural level. For example, in the human disorder, CLSD, which results from a missense mutation in *SEC23A,* ER dilation is observed in cultured skin fibroblasts from both homozygous affected individuals and phenotypically normal heterozygous carriers [22]. In our analysis of liver tissue, pancreas, and cultured murine embryonic fibroblasts from *Sec24d* heterozygous mice, we did not observe any consistent dilation of the ER, indicating that a modest reduction in the level of SEC24D in the cell types we examined is not sufficient to visibly disrupt the overall structure of the early secretory pathway.

Overall, our data suggest a critical early function for SEC24D in mammalian development. Lack of SEC24D may lead to a global defect in ER-to-Golgi protein transport, or it may result in a more specific deficit of a selective subset of cargo proteins. Although we detected expression of all four isoforms in total RNA isolated from E10.5 embryos, the developmental timing and expression pattern of the SEC24 proteins prior to

that stage is not known. Our initial studies provide evidence that the expression patterns of the four SEC24 isoforms in adult mouse tissues are at least partially overlapping; however, higher-resolution studies are required to determine if there is diversification of SEC24 isoform expression at the cell-type specific level within a given tissue. An important outstanding question is whether the various isoforms of SEC24 have evolved specific cargo-related functions. If so, have some isoforms become more specialized in their expression pattern or cargo-binding specificity, with others maintaining a more general role and a broader capacity for cargo recruitment? Given the absolute requirement for SEC24D at the earliest stages of development, it might represent the core isoform of SEC24, perhaps most analogous in function to yeast Sec24p, whereas the remaining isoforms could be non-essential or result in a less severe phenotype when deleted, analogous to yeast Lst1p and Iss1p.

Several human diseases have recently been attributed to defects in components of COPII. Mutations in the GTP-binding site of *SAR1B* lead to a class of lipid absorption disorders known as chylomicron retention diseases (CMRD) [23]. Missense mutations in *SEC23A* lead to cranio-lenticulo-sutural dysplasia (CLSD), characterized by the persistence of wide-open fontanelles into childhood and the development of Y-shaped cataracts [22]. Most recently, mutations in *SEC23B* were identified as a cause of congenital dyserythropoietic anemia type II (CDAII), a disorder leading to a specific defect in erythrocyte development [24,25]. Given the general role of COPII proteins in ER-to-Golgi transport, the limited nature of the phenotypes observed in these human diseases is surprising. However, a common theme in these disorders is the nature of the mutations. None are predicted to result in complete deficiency of the corresponding

COPII protein; rather they result in the production of a mutant protein that likely retains some normal function. Similarly, a more specific milder mutation in *Sec24d* might be compatible with adult viability and uncover a subtler phenotype. A hypomorphic mutation of *SEC24D* could underlie a human disease whose genetic cause is yet to be identified. Likewise, true null alleles of *SAR1B*, *SEC23A* and *SEC23B* might result in embryonic lethality. Further investigation into these human disorders combined with studies of mouse models deficient in components of COPII will better our understanding of cargo-specific protein transport in the early secretory pathway.

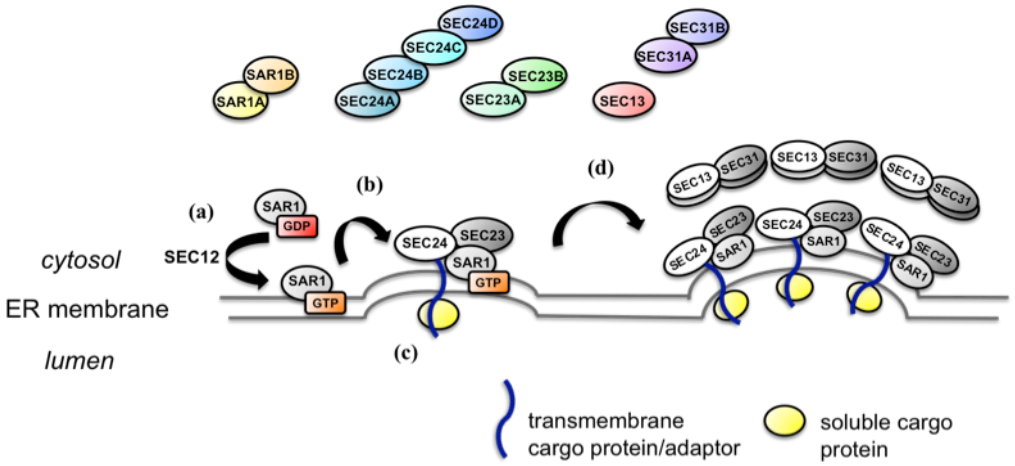


**Figure 2-1: Composition and assembly of COPII vesicles.**

Newly synthesized proteins exit the ER in COPII-coated vesicles, which are composed of the small GTP-binding protein SAR1, and two protein subunits, the heterodimeric SEC23-SEC24 complex and the heterotetrameric SEC13-SEC31 complex. Mammals possess multiple isoforms of four of the five core components (shown in color). (a) The exchange of GDP for GTP, catalyzed by the membrane bound guanine nucleotide exchange factor, SEC12, activates SAR1, which inserts into the ER membrane via an N-terminal amphipathic helix. (b) SAR1-GTP recruits the inner coat complex, SEC23-SEC24. (c) SEC24 interacts with the cytoplasmic domains of transmembrane proteins, which may in turn serve as cargo receptors for soluble proteins. (d) The outer coat complex, SEC13-SEC31, polymerizes the COPII vesicle to drive budding from the ER membrane.

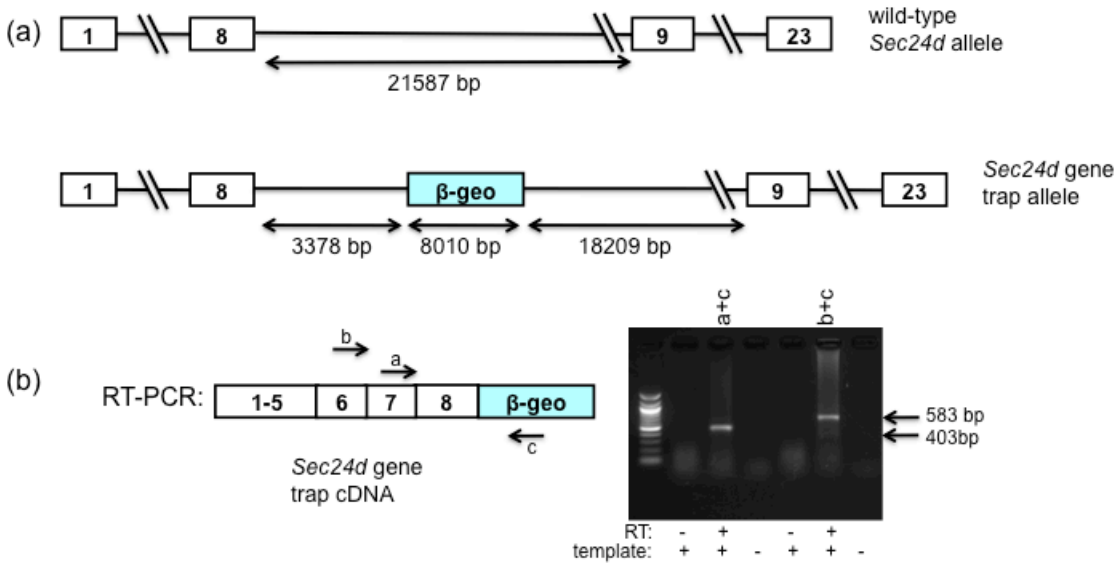


**Figure 2-2: Schematic representation and confirmation of the Line 1 gene trap allele.**

(a) Schematic representation of the wild-type and gene trap alleles at the mouse *Sec24d* locus. (b) Confirmation of the gene trap insertion by RT-PCR of RNA islolated from the RRT226 ES cell clone.

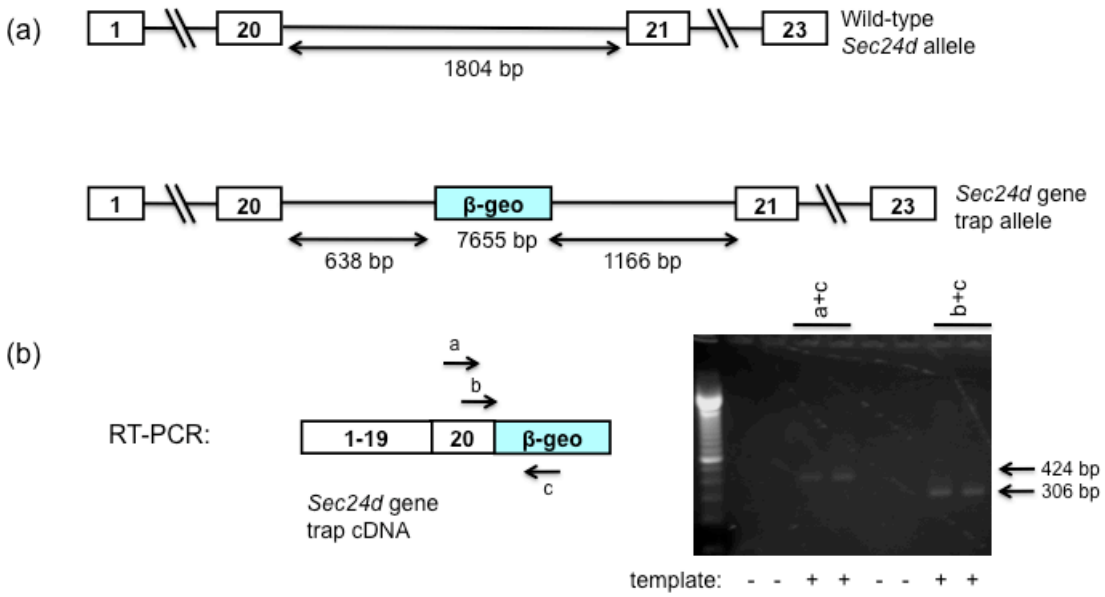


**Figure 2-3: Schematic representation and confirmation of the Line 2 gene trap allele.**

(a) Schematic representation of the wild-type and gene trap alleles at the mouse *Sec24d* locus. (b) Confirmation of the gene trap insertion by RT-PCR of RNA islolated from RRR785 ES cell clone.

**Table 2-1: Primers used in this study.**

| **Primers for RT-PCR** | | I8F13 | GGGACAAAACAGCAGCCTAC |
|---|---|---|---|
| RT1 A | CACCACCGAGTGTGTGATTC | I8F14 | CACTGGGGATATGGAACCTG |
| RT1 B | GTCTTACCCAGGAGGGTTCC | I8F15 | GGTGGGGAAGAGAACTTGTG |
| RT1 C | ATTCAGGCTGCGCAACTGTTGGG | I8F16 | CTGGCCTCTTTACACCCTTG |
| RT2 A | AGTTTTGAACCAGCCCTTGA | I8F17 | AAAGAGCGAGACCAACCTGA |
| RT2 B | CCAGATTCCATGAAGGTGCT | I8F18 | TTTTCCTGTAGGCCCATGAC |
| RT2 C | GTTTTCCCAGTCACGACGTT | I8F19 | AGCACAGGGAAGCCTAAGTG |
| **Primers for Genotyping** | | I8F20 | CCCTTTCCTCTTCCTCCACT |
| Neo A | CTTGCGCAGCTGTGCTCGACGTTG | I8F21 | GAGGTCAGAAGAGGGCATCA |
| Neo B | TCTTCGTCCAGATCATCCTGATCG | **Primers for Insertion Site Mapping (Line 2)** | |
| 1A | CATTCGTTGCTCCTCCTCTT | I20F1 | GGCAGTGGAAGGTGTAAGGA |
| 1B | TGTCCAGGAAACACGACAGA | I20F2 | GCCATGCAAGAGTCCTCAGT |
| 1C | GGGTCTCAAAGTCAGGGTC | I20F3 | GCCCCTGTCTCTAAGCCTCT |
| 2A | GCCCCTGTCTCTAAGCCTCT | I20F4 | CATCCTGTTCGTCCTCCATC |
| 2B | GCTTGTGGCAGGATTTTCTC | I20F5 | TGATCGGTTGCCACATAAAA |
| 2C | GACCTGGCTCCTATGGGATA | I20F6 | CCCTAGTCGGGCTCTTACCT |
| **Primers for Insertion Site Mapping (Line 1)** | | I20F7 | GGCCTTTCTCCCTCAAAAAG |
| I8F1 | CCATGCAGTGCTACACAAGC | **Primers for Southern Blot Probe** | |
| I8F2 | CTGCTGCCTGAAGATCAAGA | ApaI A | AATCCGTGGTTGTAGGTTGC |
| I8F3 | CATTCGTTGCTCCTCCTCTT | ApaI B | CAAAGGATCTCCCCACTCTG |
| I8F4 | AATCTGACCTGGGGAGAAGC | **Primers for Real-time RT-PCR** | |
| I8F5 | CCCAACCCTCACGACAATAA | 24A F | AGGAAGGTATTACGTCAAATGCC |
| I8F6 | GCAGAGGCTGCTATTCCATC | 24A R | TGGGGTTCTGATTAACAAGCTG |
| I8F7 | ATGGATGCTGCTGGAACTCT | 24B F | ATGTATGGCAGAGGTCCTTCC |
| I8F8 | CACAGGGAAAACGTGGAAAG | 24B R | CTGGGAGGCACTCGAATAGG |
| I8F9 | CGTGTCCTTCCCTAAACAGC | 24C F | CTGGCCGGAATGCAGATCAG |
| I8F10 | CAGGTGGGGGATGTTATGAG | 24C R | TGAGGATAGGAGCCGTATGGA |
| I8F11 | GGTGCTTTCAAATTGGTCAC | 24D F | TAGCCAAAGGTTTCTGACCGA |
| I8F12 | GATGGCGTGTAAGCTGTTGA | 24D R | GCCGCCATTTCCACATCTG |

All primer sequences listed are shown in the 5' to 3' orientation.

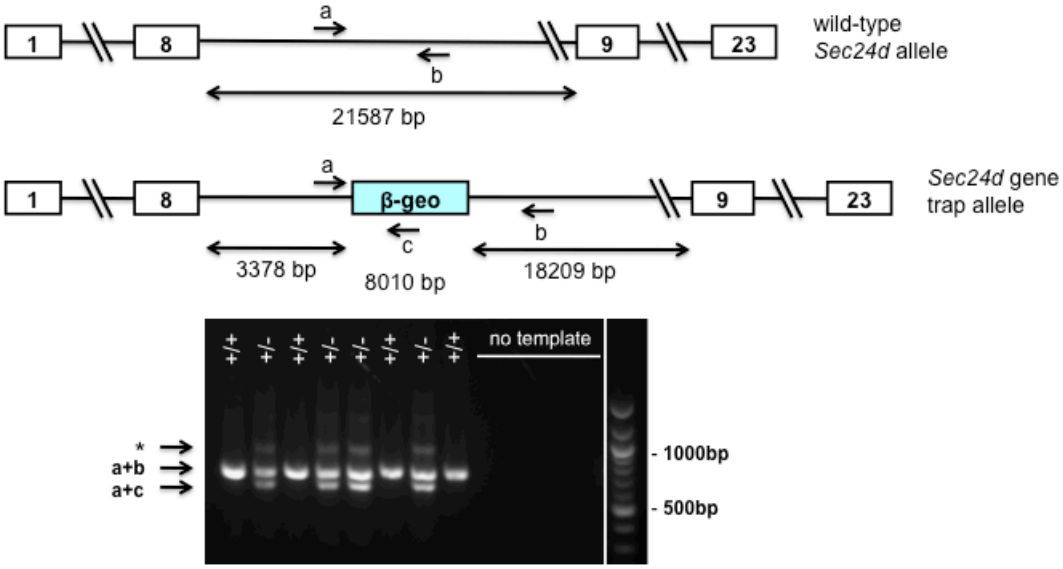


**Figure 2-4: Schematic of Line 1 (RRT226) genotyping assay.**

A three-primer competitive PCR assay was used to genotype Line 1 mice at the *Sec24d* locus. A common forward primer (primer a) lies upstream of the gene trap vector insertion site in intron 8, while two reverse primers distinguish between the wild-type (primer b) and gene trap (primer c) alleles. The a+b primer set gives rise to a 762 bp amplicon from the wild-type allele and the a+c primer set gives rise to a 666 bp amplicon from the gene trap allele. Results for 8 mice are shown above. Genotypes are denoted +/+ for homozygous wildtype and +/- for heterozygous mice. The asterisk denotes a heteroduplex band present only in the heterozygous samples.

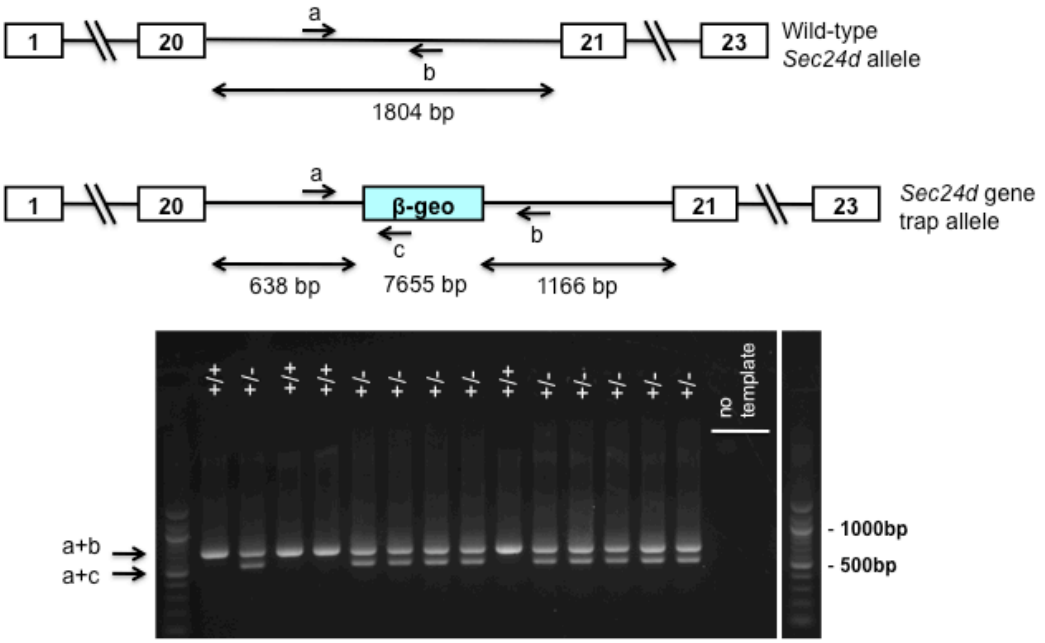


**Figure 2-5: Schematic of Line 2 (RRR785) genotyping assay.**

A three-primer competitive PCR assay was used to genotype Line 2 mice at the *Sec24d* locus. A common forward primer (primer a) lies upstream of the gene trap vector insertion site in intron 20, while two reverse primers distinguish between the wild-type (primer b) and gene trap (primer c) alleles. The a+b primer set gives rise to a 715 bp amplicon from the wild-type allele and the a+c primer set gives rise to a 558 bp amplicon from the gene trap allele. Genotyping results for 14 mice are shown above, with +/+ denoting homozygous wild-type mice and +/- denoting heterozygous mice.

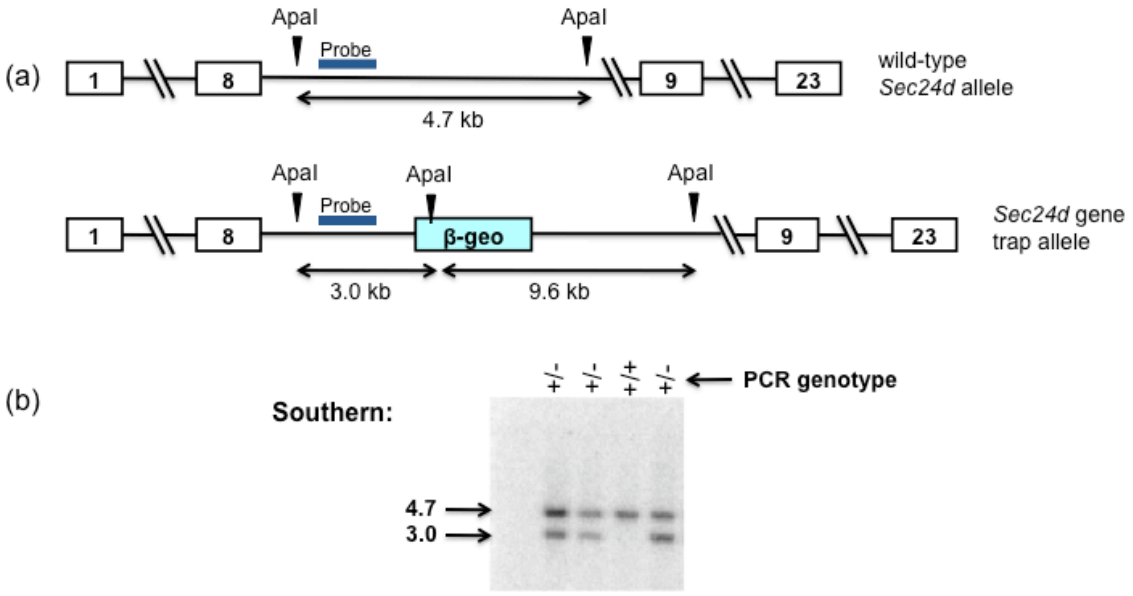


**Figure 2-6: Confirmation of Line 1 (RRT226) genotypes by Southern blot.**

(a) Schematic representation of the wild-type and gene trap alleles at the mouse Sec24d locus. (b) Confirmation of the gene trap insertion by Southern blot analysis of ApaI-digested Line 1 genomic DNA from *Sec24d*+/+ and *Sec24d*+/- mice; ApaI restriction enzyme sites and probe location are depicted in (a). Results for 4 mice are shown above, with the PCR genotypes at the top; the genotypes of 33 mice total (5 *Sec24d*+/+ and 28 *Sec24d*+/-) were confirmed by Southern blot, all in agreement with previous PCR genotypes.

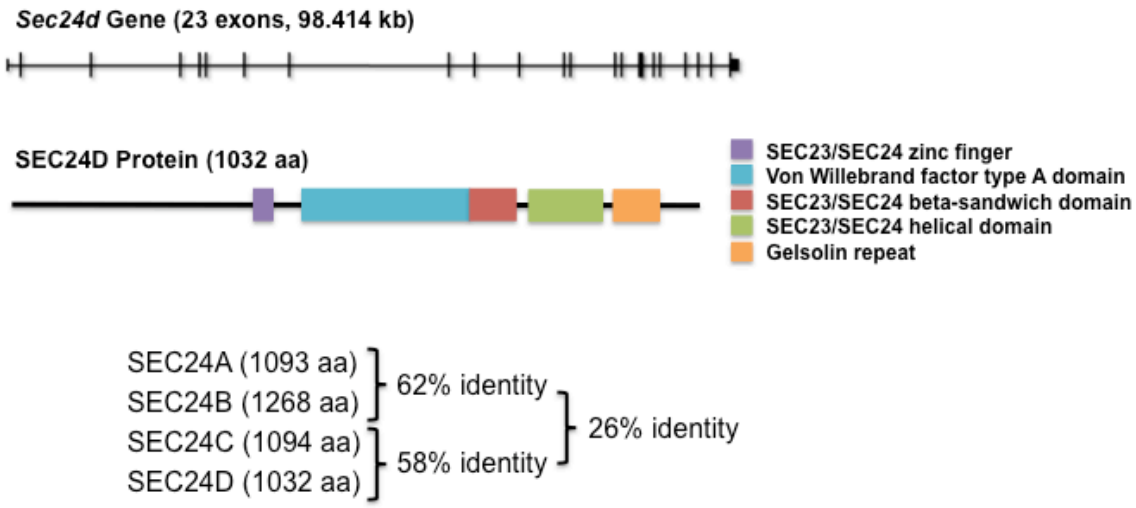


**Figure 2-7: Schematic of the *Sec24d* gene and protein product.**

The *Sec24d* gene architecture is depicted at the top with vertical and horizontal lines representing exons and introns, respectively. The SEC24D protein is shown below with the conserved domains highlighted. Though variable in overall protein length, all four isoforms of SEC24 share the same conserved domain organization and an N-terminal hypervariable region. The SEC24 isoforms can be divided into two subfamilies based on sequence identity, an A/B subfamily and a C/D subfamily.

**Table 2-2: Summary of genotypes from Line 1 (RRT226) intercrosses.**

| | Number and (%) of mice by genotype | | | | |
|---|---|---|---|---|---|
| | **+/+** | **+/-** | **-/-** | | |
| Expected (intercross) | (25) | (50) | (25) | *p*-value (25:50:25)* | *p*-value (33:67)* |
| Expected (embryonic lethal) | (33) | (67) | (0) | | |
| **Stage of development** | | | | | |
| Postnatal (n = 212) | 44 (20.8) | 168 (79.2) | 0 | 1.93E-20 | 1.02E-04 |
| F1 intercross (n = 119) | 24 (20.2)[a] | 92 (79.8)[a] | 0 | 1.54E-11 | 3.87E-03 |
| N5 intercross (n = 93) | 20 (21.5) | 73 (78.5) | 0 | 3.74E-09 | 1.55E-02 |
| E10.5-11.5 (n = 28) | 6 (21.4) | 22 (78.6) | 0 | 2.86E-03 | 1.81E-01 |
| Blastocyst (n = 27) | 8 (29.6) | 19 (70.4) | 0 | 9.94E-03 | 6.83E-01 |
| 8-cell embryo (n = 17) | 2 (11.8) | 14 (82.4) | 1 (5.8) | 2.68E-02 | 8.33E-02 |
| Total (n = 495) | 104 (21.0) | 391 (78.8)[b] | 1 (0.2) | 7.96E-46 | 6.23E-09 |

[a] 33 of 119 genotypes (5 *Sec24d*+/+ and 28 *Sec24d*+/-) were confirmed by Southern blot analysis.
[b] An excess of heterozygous mice was observed such that the ratio of *Sec24d*+/+ to *Sec24d*+/- mice differed significantly from the expected 1/3 to 2/3 ratio given an embryonic lethal *Sec24d*-/- phenotype.
* Significant *p*-values differ from the expected ratio of 25:50:25 (fifth column) or 33:67 (sixth column).

**Table 2-3: Summary of genotypes from Line 2 (RRR785) intercrosses.**

| Number of mice by genotype | | | | |
|---|---|---|---|---|
| | +/+ | +/- | -/- | p-value |
| F1 Intercross (n = 88) | 30 | 58 | 0 | |
| Expected (25:50:50) | 22 | 44 | 22 | 4.21E-07 |
| Expected (33:67) | 29.33 | 58.67 | | 8.80E-01 |
| Blastocyst (n = 99) | 26 | 47 | 26 | |
| Expected (25:50:50) | 24.75 | 49.5 | 24.75 | 8.81E-01 |

No *Sec24d* homozygous null mice were observed when the F2 generation (offspring of F1 intercrosses between *Sec24d+/-* mice) of mice from Line 2 was genotyped at 2 weeks of age. Additional F1 intercrosses were set up to obtain blastocyst-stage embryos. In contrast to Line 1, in which no null blastocysts were observed, *Sec24d* null blastocysts were present in expected numbers in Line 2.

**Table 2-4: Genotypes of mice from backcrosses of Line 1 (RRT226).**

| Backcross generation | Number and (%) of mice by genotype | | *p*-value (50:50) | *p*-value (33:67) |
|---|---|---|---|---|
| | +/+ | +/- | | |
| Expected ratio | (50) | (50) | | |
| Expected (excess of *Sec24d+/-*) | (33.3) | (66.7) | | |
| N2 (n = 35) | 11 (31.4) | 24 (68.6) | **0.0280** | 0.8111 |
| N3 (n = 44) | 26 (59.1) | 18 (40.9) | 0.2278 | **0.0003** |
| N4 (n = 45) | 27 (60.0) | 18 (40.0) | 0.1797 | **0.0001** |
| N5 (n = 51) | 17 (33.3) | 34 (66.7) | **0.0173** | 1.0000 |
| N6 (n = 47) | 25 (53.2) | 22 (46.8) | 0.6617 | **0.0039** |
| N7 (n = 73) | 25 (34.2) | 48 (65.8) | **0.0071** | 0.8685 |
| Total (n = 295) | 131 (44.4) | 164 (55.6) | 0.0547 | **0.0001** |

Backcross of a *Sec24d+/-* heterozygous mouse to a wild-type C57BL6/J mouse is expected to result in a 50:50 ratio of homozygous wild-type (*Sec24d+/+*) to heterozygous (*Sec24d+/-*) mice. Significant *p*-values (shown in bold) demonstrating a deviation from the expected 50:50 ratio are shown in the fourth column. The same backcross generations show no significant difference from an expected ratio of 33:67 (fifth column).

**Table 2-5: Complete blood count analysis of *Sec24d*+/+ and *Sec24d*+/- mice.**

| Mouse | Sex | Genotype | WBC x10$^3$ cells/ ul | RBC x10$^6$ cells/ ul | HGB g/dL | HCT % | MCV fL | MCH pg | MCHC g/dL |
|---|---|---|---|---|---|---|---|---|---|
| 1 | M | *Sec24d* +/+ | 0.66 | 0.93 | 1.4 | 4.4 | 47.6 | 15.0 | 31.6 |
| 2 | M | *Sec24d* +/+ | 0.93 | 0.77 | 1.2 | 3.6 | 47.4 | 15.0 | 31.7 |
| 3 | M | *Sec24d* +/- | 0.82 | 0.98 | 1.6 | 4.5 | 47.4 | 16.3 | 34.4 |
| 4 | M | *Sec24d* +/- | 0.70 | 0.97 | 1.6 | 4.5 | 47.3 | 16.1 | 34.0 |
| 5 | F | *Sec24d* +/+ | 0.41 | 1.00 | 1.6 | 4.7 | 46.4 | 15.8 | 34.1 |
| 6 | F | *Sec24d* +/+ | 0.58 | 0.94 | 1.7 | 4.4 | 46.8 | 17.5 | 37.4 |
| 7 | F | *Sec24d* +/- | 0.81 | 0.99 | 1.6 | 4.7 | 47.9 | 16.2 | 33.8 |
| 8 | F | *Sec24d* +/- | 0.55 | 1.03 | 1.6 | 4.8 | 46.8 | 16.0 | 34.1 |

| Mouse | Sex | Genotype | CHCM g/dL | CH pg | RDW % | HDW g/dL | PLT x10$^3$ cells/ul | MPV fL |
|---|---|---|---|---|---|---|---|---|
| 1 | M | *Sec24d* +/+ | 31.2 | 14.8 | 12.2 | 1.85 | 108 | 6.8 |
| 2 | M | *Sec24d* +/+ | 31.5 | 14.9 | 12.2 | 1.82 | 80 | 5.9 |
| 3 | M | *Sec24d* +/- | 31.2 | 14.8 | 11.9 | 1.87 | 110 | 6.7 |
| 4 | M | *Sec24d* +/- | 31.3 | 14.8 | 11.5 | 1.78 | 91 | 6.0 |
| 5 | F | *Sec24d* +/+ | 31.9 | 14.8 | 12.1 | 1.82 | 106 | 6.5 |
| 6 | F | *Sec24d* +/+ | 31.8 | 14.9 | 11.9 | 1.79 | 71 | 5.1 |
| 7 | F | *Sec24d* +/- | 31.5 | 15.1 | 11.8 | 1.79 | 102 | 6.6 |
| 8 | F | *Sec24d* +/- | 31.6 | 14.8 | 11.7 | 1.82 | 111 | 6.7 |

Whole blood was drawn from two males and two females of each genotype by retro-orbital puncture and analyzed on an Advia120 whole blood analyzer. No consistent differences were noted between *Sec24d*+/+ and *Sec24d*+/- mice in any of the thirteen routine parameters that were measured.

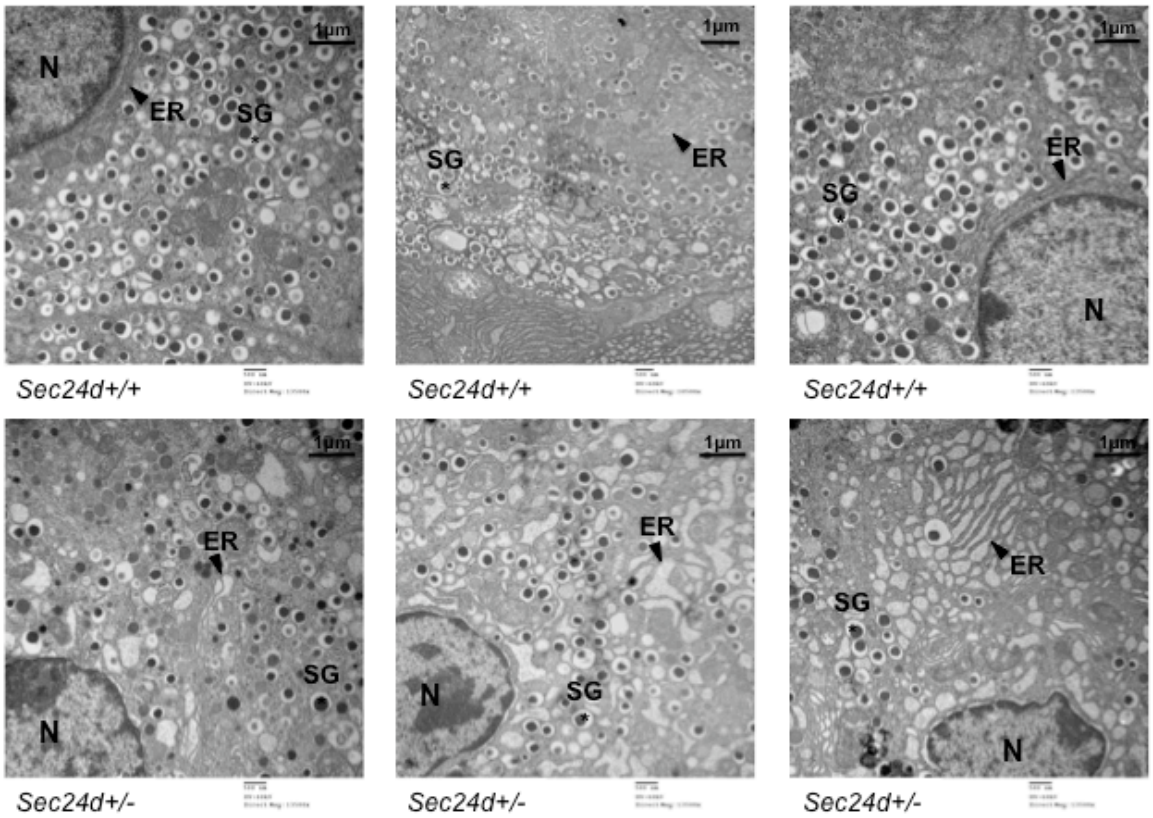


**Figure 2-8: Transmission electron micrograph of pancreatic islet cells.**

Pancreas tissue was harvested from three age-matched male mice of each genotype. Samples were viewed at 10,500-13,500x direct magnification. Islet cells are identified by the presence of specialized secretory granules. A scale bar at the top right of each image indicates a length of 1µm. No consistent ER dilation was observed in *Sec24d* heterozygous mice relative to wild-type mice. Abbreviations: N = nucleus, ER = endoplasmic reticulum (black arrowheads), SG = secretory granules.

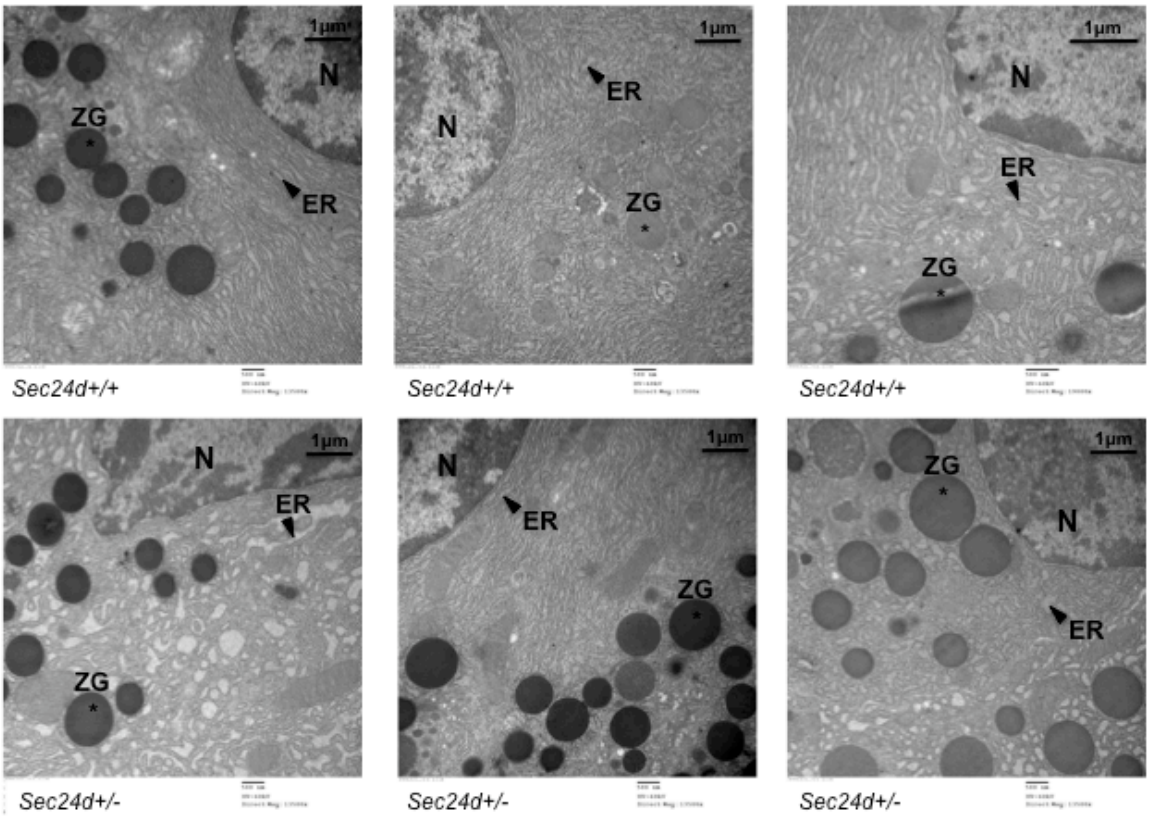


**Figure 2-9: Transmission electron micrograph of pancreatic acinar cells.**

Pancreas tissue was harvested from three age-matched male mice of each genotype. Samples were viewed at 10,500-13,500x direct magnification. Acinar cells are identified by the presence of zymogen granules. A scale bar at the top right of each image indicates a length of 1µm. No consistent ER dilation was observed in *Sec24d* heterozygous mice relative to wild-type mice. Abbreviations: N = nucleus, ER = endoplasmic reticulum (black arrowheads), ZG = zymogen granules.

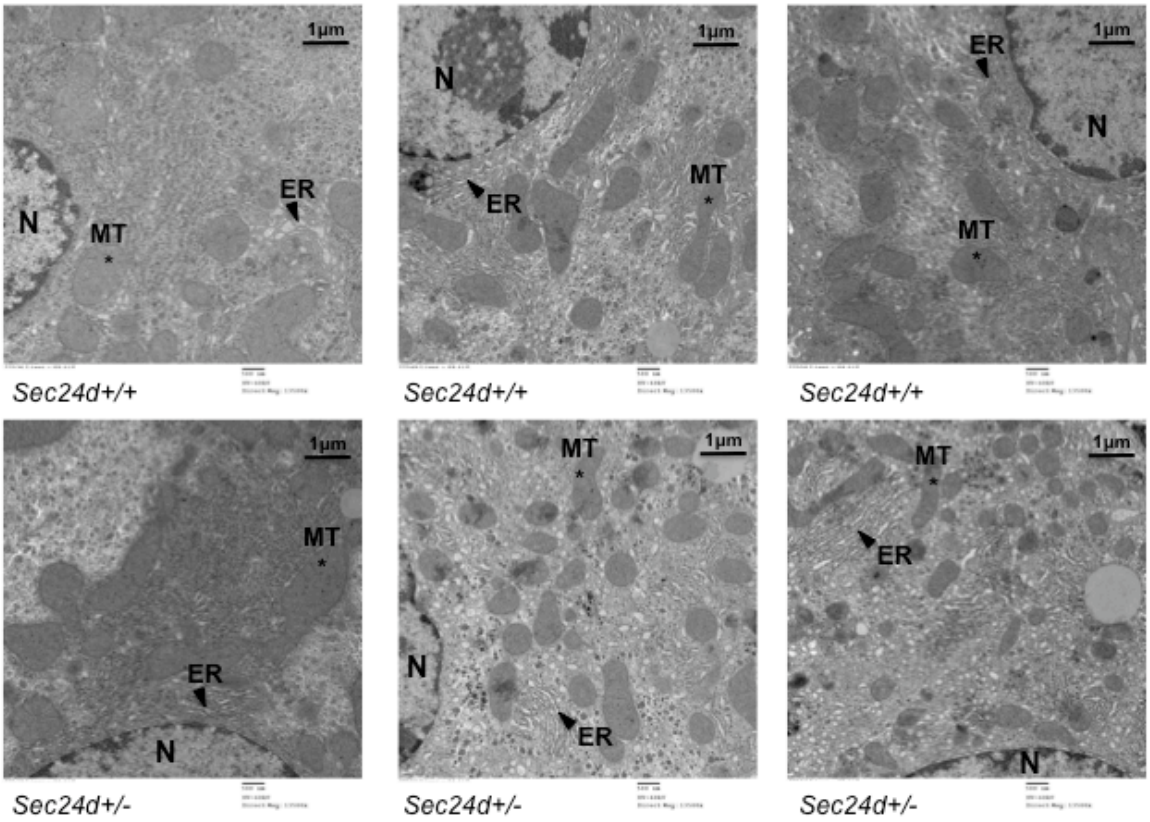


**Figure 2-10: Transmission electron micrograph of hepatocytes.**

Liver tissue was harvested from three age-matched male mice of each genotype. Samples were viewed at 10,500-13,500x direct magnification. A scale bar at the top right of each image indicates a length of 1μm. No consistent ER dilation was observed in *Sec24d* heterozygous mice relative to wild-type mice. Abbreviations: N = nucleus, ER = endoplasmic reticulum (black arrowheads), MT = mitochondria.

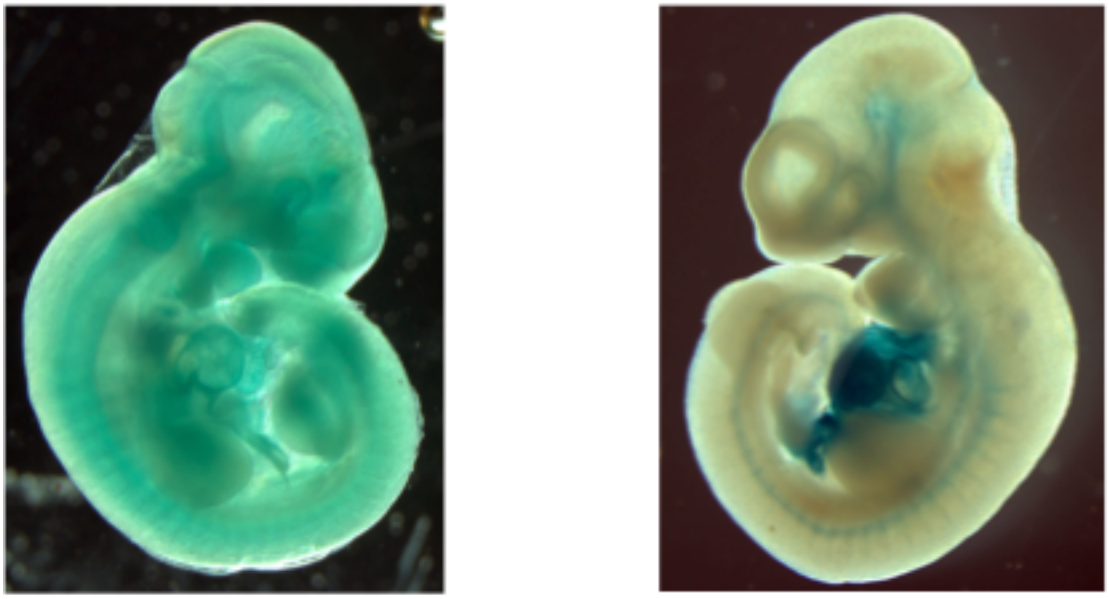


**Figure 2-11: Mouse embryos stained for β-galactosidase activity.**

Embryos were harvested at E10.5 and stained for β-galactosidase activity. The gene trap vector contains a β-geo cassette, therefore, the X-gal staining pattern serves as a proxy for the normal SEC24A and SEC24D expression patterns. It is interesting that while SEC24A shows a ubiquitous expression pattern and SEC24D appears more specific (brain, heart and spinal cord), *Sec24a* null mice are viable (B. Zhang et al., manuscript in preparation) and *Sec24d* null mice die very early in development.

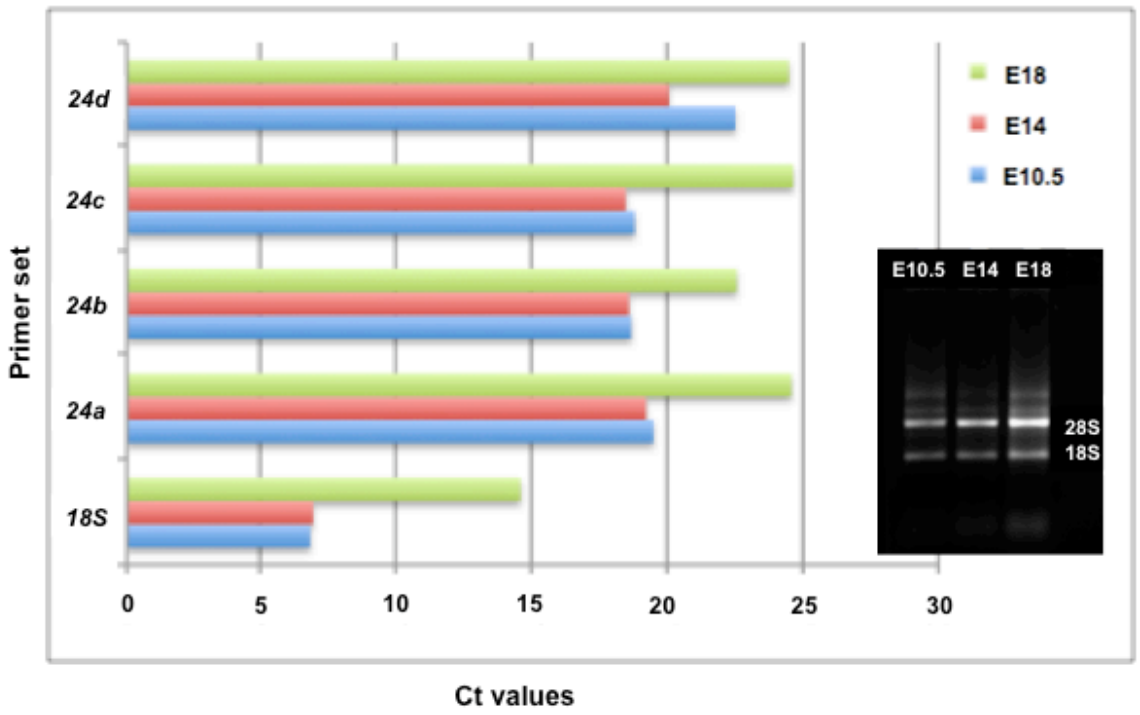


**Figure 2-12: Expression of *Sec24a-d* mRNA in developing mouse embryos.**

RNA from embryos at E10.5, E14 and E18 was assayed by real-time RT-PCR. All four isoforms of *Sec24* are expressed at mid-embryogenesis and mRNA levels remain consistent into late embryogenesis. Ct values are an average of measurements done in triplicate. The inset shows the RNA isolated from mouse embryos at E10.5, E14 and E18. RNA integrity was assessed by agarose gel electrophoresis. Minor degradation is evident in the E18 sample.

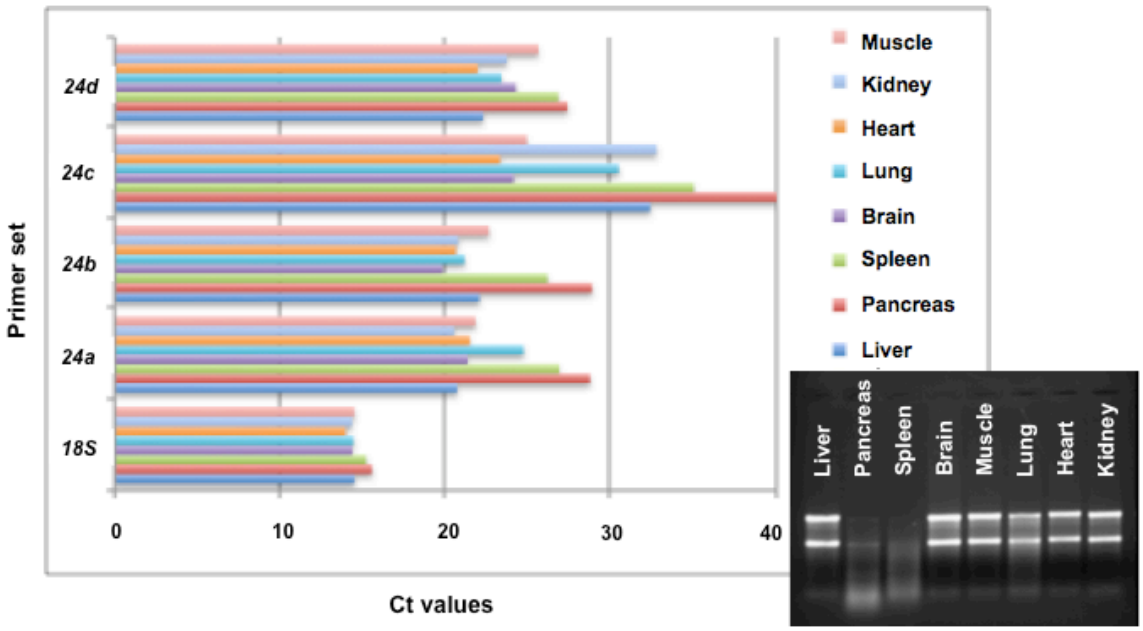


**Figure 2-13: Expression of *Sec24a-d* mRNA in adult mouse tissues.**

RNA from adult mouse tissues was assayed by real-time RT-PCR. Ct values are an average of measurements done in triplicate. Shown in the inset, moderate degradation of the RNA was visible in the samples prepared from pancreas and spleen. While there is a slight increase in the Ct values for the *18S* primer set for pancreas and spleen, it is unclear why the Ct values for pancreas and spleen were affected to a greater extent in the reactions with the *Sec24* primers.

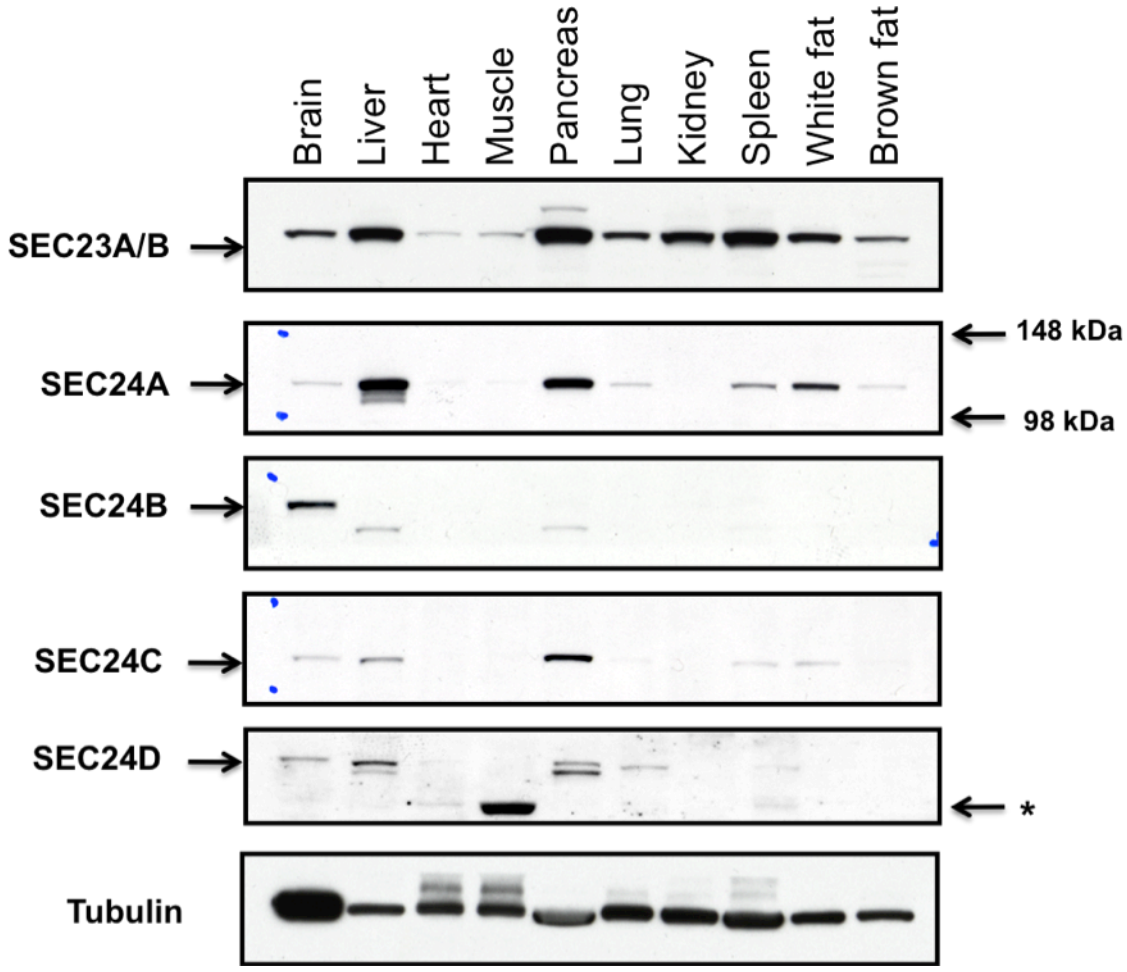


**Figure 2-14: Initial assessment of SEC24A-D tissue distribution.**

Mouse tissue lysates were separated by SDS-PAGE, transferred to nitrocellulose, and probed with the indicated antibody (left). The top and bottom of each SEC24 blot corresponds to the 148 kDa and 98 kDa markers, respectively. An asterisk marks an intense band of a smaller than expected size for SEC24D present in the muscle tissue lysate, which could be due to non-specific antibody reactivity. Doublets present in the liver and pancreas samples for SEC24D could represent alternative isoforms of SEC24D, or cross-reactivity with SEC24C. SEC24A appears to be the most widely expressed isoform, although it is possible the SEC24A antibody is more sensitive compared to the SEC24B-D antibodies. Further characterization is needed to better address these issues. (Figure courtesy of Xiao-Wei Chen)

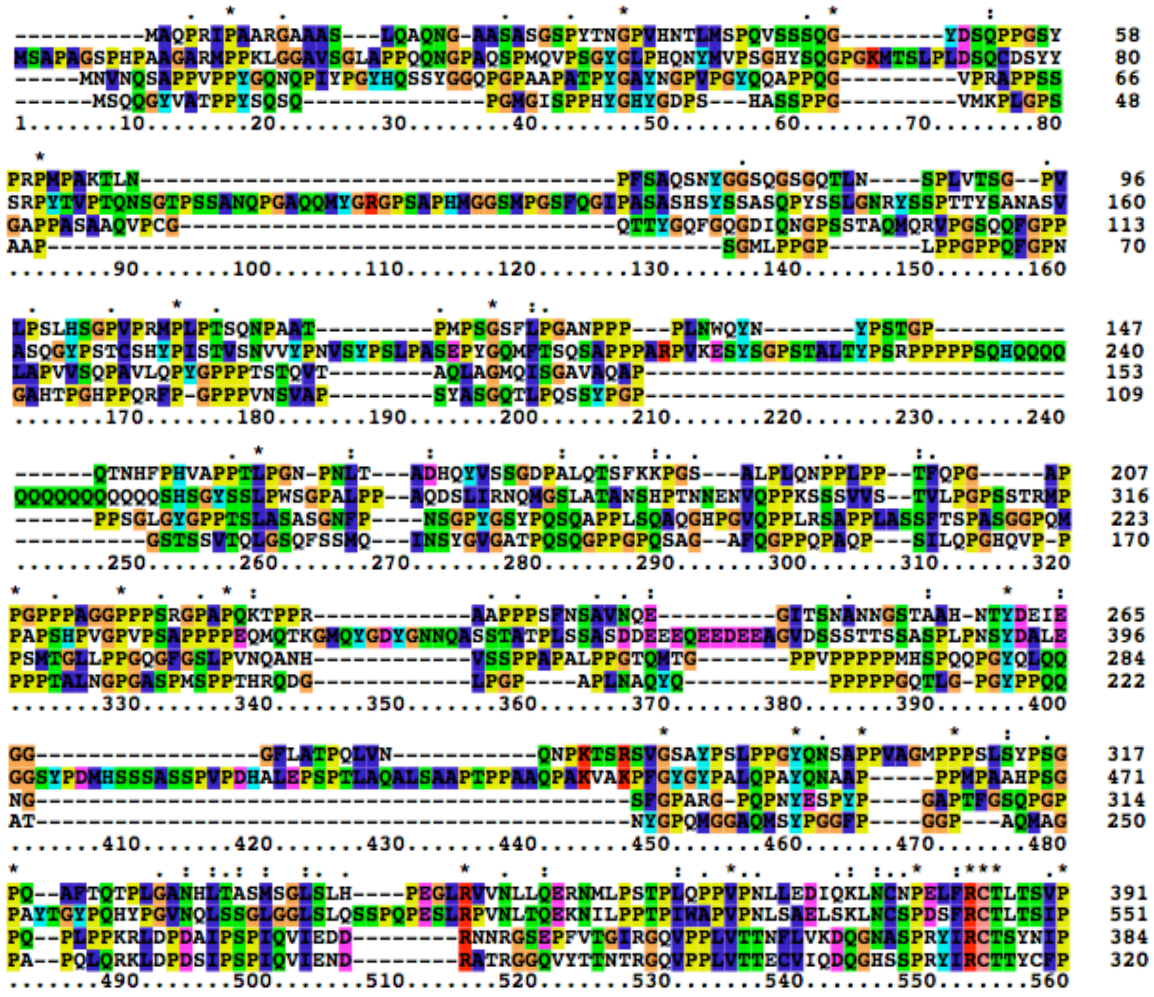

**Figure 2-15: Protein sequence alignment of SEC24A-D (N-terminal region).**

Sequence alignments were performed using ClustalX software with the Gonnet series matrix. Roughly the N-terminal third (the most variable region between SEC24 isoforms) of the protein alignment is shown.

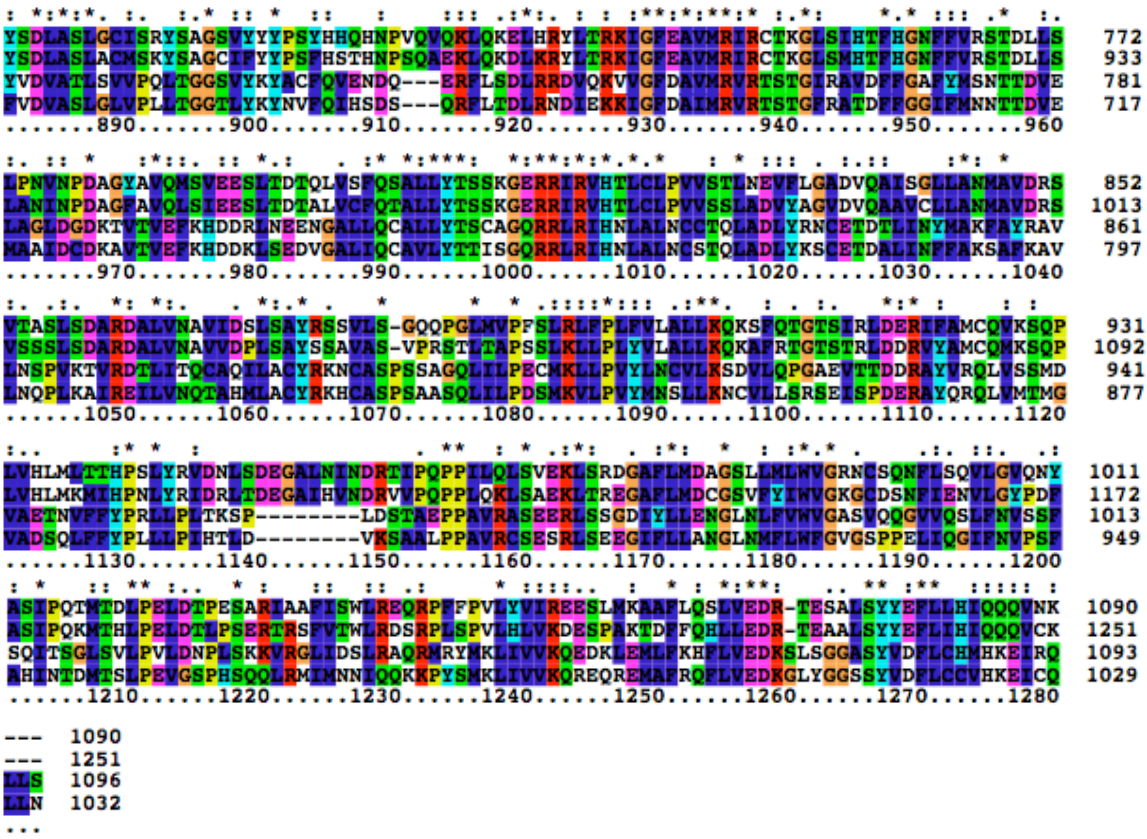

**Figure 2-16: Protein alignment of SEC24A-D (C-terminal region).**

Sequence alignments were performed using ClustalX software with the Gonnet series matrix. Approximately the C-terminal third (the most highly conserved region between SEC24 isoforms) of the protein alignment is shown.

## References


1. Bonifacino JS, Glick BS. The mechanisms of vesicle budding and fusion. Cell. 2004;116:153-66.

2. Lee MC, Miller EA, Goldberg J, Orci L, Schekman R. Bi-directional protein transport between the ER and Golgi. Annu Rev Cell Dev Biol. 2004;20:87-123.

3. Lee MC and Miller EA. Molecular mechanisms of COPII vesicle formation. Semin Cell Dev Biol. 2007;18:424-34.

4. Bickford LC, Mossessova E, Goldberg J. A structural view of the COPII vesicle coat. Curr Opin Struct Biol. 2004;14:147-53.

5. Lee MC, Orci L, Hamamoto S, Futai E, Ravazzola M, Schekman R. Sar1p N-terminal helix initiates membrane curvature and completes the fission of a COPII vesicle. Cell. 2005;122:605-17.

6. Miller E, Antonny B, Hamamoto S, Schekman R. Cargo selection into COPII vesicles is driven by the Sec24p subunit. EMBO J. 2002;21:6105-13.

7. Miller EA, Beilharz TH, Malkus PN, Lee MC, Hamamoto S, Orci L, Schekman R. Multiple cargo binding sites on the COPII subunit Sec24p ensure capture of diverse membrane proteins into transport vesicles. Cell. 2003;114:497-509.

8. Miller EA, Liu Y, Barlowe C, Schekman R. ER-Golgi transport defects are associated with mutations in the Sed5p-binding domain of the COPII coat subunit, Sec24p. Mol Biol Cell. 2005;16:3719-26.

9. Stagg SM, Gürkan C, Fowler DM, LaPointe P, Foss TR, Potter CS, Carragher B, Balch WE. Structure of the Sec13/31 COPII coat cage. Nature. 2006;439:234-8.

10. Tang BL, Kausalya J, Low DY, Lock ML, Hong W. A family of mammalian proteins homologous to yeast Sec24p. Biochem Biophys Res Commun. 1999;258:679-84.

11. Roberg KJ, Crotwell M, Espenshade P, Gimeno R, Kaiser CA. LST1 is a SEC24 homologue used for selective export of the plasma membrane ATPase from the endoplasmic reticulum. J Cell Biol. 1999;145:659-72.

12. Kurihara T, Hamamoto S, Gimeno RE, Kaiser CA, Schekman R, Yoshihisa T. Sec24p and Iss1p function interchangeably in transport vesicle formation from the endoplasmic reticulum in Saccharomyces cerevisiae. Mol Biol Cell. 2000;11:983-98.

13. Peng R, De Antoni A, Gallwitz D. Evidence for overlapping and distinct functions in protein transport of coat protein Sec24p family members. J Biol Chem. 2000;275:11521-8.

14. Mancias JD, Goldberg J. Structural basis of cargo membrane protein discrimination by the human COPII coat machinery. EMBO J. 2008;27:2918-28.

15. Skarnes WC. Gene trapping methods for the identification and functional analysis of cell surface proteins in mice. Methods Enzymol. 2000;328:592-615.

16. Kendall SK, Samuelson LC, Saunders TL, Wood RI, Camper SA. Targeted disruption of the pituitary glycoprotein hormone alpha-subunit produces hypogonadal and hypothyroid mice. Genes Dev. 1995;9:2007-19.

17. Bahou WF, Bowie EJ, Fass DN, Ginsburg D. Molecular genetic analysis of porcine von Willebrand disease: tight linkage to the von Willebrand factor locus. Blood. 1988;72:308–313.

18. Nagy A, Gertsenstein M, Vintersten K, Behringer R. Manipulating the Mouse Embryo: A Laboratory Manual. CSH Press. 2003.

19. Van Keuren ML, Gavrilina GB, Filipiak WE, Zeidler MG, Saunders TL. Generating transgenic mice from bacterial artificial chromosomes: transgenesis efficiency, integration and expression outcomes. Transgenic Res. 2009; Apr 26 [Epub ahead of print].

20. Chen XW, Leto D, Chiang SH, Wang Q, Saltiel AR. Activation of RalA is required for insulin-stimulated Glut4 trafficking to the plasma membrane via the exocyst and the motor protein Myo1c. Dev Cell. 2007;13:391-404.

21. Stryke D, Kawamoto M, Huang CC, Johns SJ et al. BayGenomics: a resource of insertional mutations in mouse embryonic stem cells. Nucleic Acids Res. 2003;31:278-81.

22. Fromme JC, Ravazzola M, Hamamoto S, Al-Balwi M, Eyaid W, Boyadjiev SA, Cosson P, Schekman R, Orci L. The genetic basis of a craniofacial disease provides insight into COPII coat assembly. Dev Cell. 2007;13:623-34.

23. Jones B, Jones EL, Bonney SA, Patel HN, et al. Mutations in a Sar1 GTPase of COPII vesicles are associated with lipid absorption disorders. Nat Genet. 2003;34:29-31.

24. Schwarz K, Iolascon A, Verissimo F, Trede NS, et al. Mutations affecting the secretory COPII coat component SEC23B cause congenital dyserythropoietic anemia type II. Nat Genet. 2009;Jun 28. [Epub ahead of print].

25. Bianchi P, Fermo E, Vercellati C, Boschetti C, et al. Congenital dyserythropoietic anemia type II (CDAII) is caused by mutations in the SEC23B gene. Hum Mutat. 2009;Jun 26. [Epub ahead of print].

# CHAPTER III:

# A PROTEOMICS APPROACH TO STUDYING DEFECTS IN ER-TO-GOLGI PROTEIN TRANSPORT


## Abstract

The LMAN1-MCFD2 complex is the only known example of a specific cargo transport receptor in mammals. In the human disorder, combined deficiency of factor V and factor VIII (F5F8D), deficiency of either LMAN1 or MCFD2 leads to a coordinate reduction in plasma levels of factor V (FV) and factor FVIII (FVIII). Although FV and FVIII are the only cargo proteins known to be affected in patients with F5F8D, the evolutionary conservation and broad expression pattern of LMAN1 and MCFD2 suggest they have a more general role in protein secretion. Mouse models of F5F8D engineered to be deficient in *Lman1* or *Mcfd2* represent an important resource for further investigation into selective cargo transport by the LMAN1-MCFD2 complex. A wide range of samples can be obtained for analysis, including plasma, tissue samples and primary cells. Here, we report an initial characterization of samples from *Lman1* null mice using a mass spectrometry-based proteomics approach, and demonstrate the use of isobaric tags (iTRAQ reagents) to quantitatively analyze proteins from plasma, pancreatic zymogen granules, and lectin-affinity purified liver tissue lysates. Although our preliminary analyses presented a range of technical challenges and failed to identify any consistent differences between wild-type and *Lman1* null mice, a global proteomics-

based approach remains a promising method by which to identify additional cargo proteins whose intracellular trafficking relies upon the LMAN1-MCFD2 transport receptor complex.

**Introduction**

Combined deficiency of factor V and factor VIII (F5F8D) is the first known human disease due to a defect in ER-to-Golgi transport of a specific subset of secretory proteins [1,2]. Mutations in either of two genes, *LMAN1* or *MCFD2*, underlie this disorder, which affects over 100 families worldwide [3]. F5F8D is inherited in an autosomal recessive fashion, is distinct from coinheritence of FV deficiency (parahemophilia) and FVIII deficiency (hemophilia A), and results in mild-to-moderate bleeding symptoms due to the critical role of FV and FVIII as co-factors in the clotting cascade (Figure 3-1). Prior to its implication in F5F8D, LMAN1 was well known as ERGIC-53, a 53-kDa marker protein of the ER-Golgi intermediate compartment (ERGIC). LMAN1 is a type I transmembrane protein that spans the ER membrane and contains both a di-phenylalanine ER-exit motif and a di-lysine ER-retrieval signal, which promote its recycling between the ER and ERGIC (Figure 3-2). In contrast, MCFD2 is a 16-kDa soluble protein, which contains two EF-hand domains, but lacks any identifiable sorting signals (Figure 3-2). MCFD2 is retained in the early secretory pathway via an interaction with LMAN1. LMAN1 and MCFD2 normally form a cargo receptor complex that is required for efficient transport of factor V (FV) and factor VIII (FVIII) from the ER to the Golgi (Figure 3-3) [4,5]. A deficiency of either component, or a disruption of LMAN1-MCFD2 complex formation, results in inadequate transport and decreased secretion of FV and FVIII. Patients with F5F8D have levels of FV and FVIII that are 5-

30% normal, with the mean levels being lower in the subgroup of patients with *MCFD2* mutations [6]. It is unknown whether the residual secretion of FV and FVIII in the absence of the LMAN1-MCFD2 cargo receptor complex represents an alternative receptor-mediated transport pathway or a basal level of secretion by bulk flow.

Although patients with F5F8D do not exhibit any overt phenotype apart from the mild-to-moderate bleeding disorder due to deficiency of FV and FVIII, it is unknown whether the transport of other cargo proteins is affected. Recent reports have suggested two lysosomal proteins, cathepsin C and cathepsin Z, and alpha-1-antitrypsin as potential cargo proteins for LMAN1 (Figure 3-3) [7-9]. It is not certain whether transport of these candidate cargo proteins is dependent on MCFD2, and defects in the sorting/secretion of these proteins has not been confirmed in patients with F5F8D. The evolutionary conservation and ubiquitous expression pattern of LMAN1 and MCFD2 suggests they may facilitate protein trafficking and secretion of a broader range of cargo proteins. Initial attempts by our laboratory to characterize defects in protein secretion in EBV-immortalized lymphoblast cell lines derived from LMAN1- or MCDF2-deficient patients by two-dimensional gel-based proteomics were unsuccessful and failed to identify any differences in protein secretion.

In order to identify additional cargo proteins whose transport relies upon the LMAN1–MCFD2 pathway, we decided to take a proteomics-based approach using mouse models recently generated in our laboratory that are deficient in *Lman1* (also referred to as *Lman1* null or *Lman1-/-* mice) or *Mcfd2* (Bin Zhang, manuscript in preparation). The ability to study any tissue type of interest and to obtain samples from null mice and control wild-type mice that are otherwise genetically identical are a few of

the advantages of working with a mouse model. The challenge in this situation is the fact that the defect in protein secretion in F5F8D deficiency is not an all-or-nothing effect. Given the coordinate reduction in levels of FV and FVIII in this disorder, it is reasonable to assume that the levels of other cargo proteins would be affected to a similar or lesser extent. Therefore, a precise method to quantify protein abundance is necessary to detect decreases in protein levels at the range of 5-30% of normal.

Although mass spectrometry is not an inherently quantitative method, several approaches exist to perform tandem mass spectrometry analysis in a quantitative manner, including the incorporation of stable non-radioactive isotopes such as $^{13}C$, $^{15}N$, or through the addition of reporter 'tags' via chemical modification of tryptic peptides [10]. The isobaric tag for relative and absolute quantitation (iTRAQ) method utilizes a reaction with primary amine groups to label the N-termini and lysine side chains of tryptic peptides with a series of tags of equal mass (total mass of 145). Each tag is composed of a balancer (ranging in mass from 31 to 28) and a reporter group (ranging in mass from 114 to 117), and the tags yield a series of signature ions in the fragmentation mode of tandem mass spectrometry (MS/MS) analysis that differ by one m/z unit (Figures 3-4 and 3-5) [11]. The availability of four different tags (114, 115, 116 and 117), and more recently eight different tags (113-119 and 121) allows samples to be multiplexed for MS/MS analysis. Because the same peptide present in multiple differentially labeled samples will produce an identical fragmentation spectrum, the matching spectra are combined and relative quantification is obtained by comparing the ratios of the iTRAQ reporter ions in the MS/MS spectrum for that particular peptide (Figure 3-5). Importantly, the availability of four to eight different isobaric tags makes it possible to

include both 'biological' replicates, samples from individual mice (i.e. two mice of the same genotype), and 'analytical' or 'technical' replicates, duplicates of individual samples (to assess the degree of variability that might arise in the technical aspects of sample preparation, handling, and analysis). Here we report the use of the iTRAQ method and tandem mass spectrometry-based proteomics to perform a quantitative comparison of complex protein mixtures from the plasma, liver, and pancreas of wild-type versus *Lman1* null mice.

## Materials and Methods

### Lman1 null mice

*Lman1* null mice were generated using a gene-trapped ES-cell line (XST010) from BayGenomics [12]. These mice carry an insertional mutation in intron 10 of the *Lman1* gene and do not express any detectable levels of *Lman1* mRNA or protein (Bin Zhang, manuscript in preparation). All experiments utilized adult mice that had been backcrossed to C57BL6/J for at least six generations.

### Preparation of mouse plasma and plasma immunodepletion

Whole blood was withdrawn from mice anesthetized with isofluorane by cardiac puncture. Following centrifugation to pellet RBCs, WBCs and platelets, plasma was aliquoted and immediately frozen in liquid nitrogen. Plasma samples from four 2 to 3-month old male mice of each genotype (*Lman1*+/+ and *Lman1*-/-) were pooled and subjected to immunoaffinity chromatography with a multiple affinity removal system column (Agilent) designed to deplete the three most abundant mouse plasma proteins (albumin, transferrin and IgG). The efficiency of plasma immunodepletion was assessed

by SDS-PAGE on a 4-20% Tris-glycine polyacrylamide gel (Novex) and separated proteins were visualized by staining with Gel Code Blue Reagent (Pierce).

**Purification of zymogen granules from mouse pancreas**

Pancreatic zymogen granules were isolated by subcellular fraction using a previously established Percoll gradient procedure [13]. Briefly, mouse pancreata were homogenized in ice-cold homogenization buffer (10mM MOPS pH7.0, 300mM sucrose, 0.1mM $MgSO_4$, 1mM EGTA, 0.1mM PMSF). Homogenates were centrifuged at 300 x *g* and 1,200 x *g* to generate a crude particulate and the particulate was resuspended in equal volumes of homogenization buffer and Percoll. Following ultracentrifugation at 45,000 x *g*, the dense white zymogen granule band was collected, washed with homogenization buffer, and frozen at -80C. Zymogen granules were purified from four individual female mice (two 2-month old mice of each genotype, *Lman1*+/+ and *Lman1*-/-) and submitted to the University of Michigan Proteomics Core Facility for quantitative MS/MS analysis. In preparation for iTRAQ-labeling and MS/MS analysis, zymogen granules were lysed in 6M urea/0.5% SDS. A portion of each tryptic digest was used for iTRAQ-labeling, and the remainder was stored at -80ºC.

**Lectin-affinity purification of liver tissue lysates**

Liver tissue lysates were prepared as follows from two individual age- and sex-matched mice of each genotype, *Lman1*+/+ and *Lman1*-/-. Mice were anesthetized with isofluorane, and liver tissue was harvested and immediately frozen in liquid nitrogen. Lectin-affinity purification was performed using the Qproteome sialic glycoprotein kit (Qiagen), according to the manufacturer's instructions. Briefly, liver tissue was homogenized in binding buffer and subjected to lectin-affinity chromatography using the

wheat germ agglutinin (WGA) spin-columns supplied in the kit. Bound proteins were washed, eluted, and precipitated with acetone prior to submitting them to the University of Michigan Proteomics Facility for quantitative MS/MS analysis. One quarter of each sample was used for iTRAQ-labeling and the remainder was stored at -80ºC.

**Preparation of tryptic peptides and iTRAQ labeling reaction**

Preparation of samples for quantitative MS/MS analysis was carried out at the University of Michigan Proteomics Core Facility. Tryptic peptides were labeled with isobaric tags (iTRAQ reagents, Applied Biosystems) according to the manufacturer's protocol. Briefly, samples were resuspended in 0.5M triethylammonium bicarbonate (TEAB, pH 8.5). Disulfide bonds were reduced by incubation with 50mM tris-(2-carboxyethyl) phosphine (TCEP) and free cysteines were blocked by the addition of 200mM methyl methanethiosulfonate (MMTS). Reduced and blocked protein samples were precipitated with acetone, resuspended in 0.5 M TEAB/0.1% SDS, and digested with 10μg modified porcine trypsin (Promega) overnight at 40ºC. Tryptic peptides were labeled by incubating with the appropriate iTRAQ 4-plex (113, 114, 115 or 116) or 8-plex (113, 114, 115, 116, 117, 118, 119 or 121) reagent for two hours. After labeling, all samples were combined and stored at -20ºC until proceeding with fractionation by two-dimensional liquid chromatography.

**Two-dimensional liquid chromatography and MS/MS**

Liquid chromatography and MS/MS analysis were also carried out in the University of Michigan Proteomics Core facility (by Eric Simon and John Strahler). Multiplexed iTRAQ-labeled tryptic peptides were fractionated via two rounds of liquid chromatography prior to MS/MS analysis. Fractionation on a strong-cation exchange

(SCX) column yielded eight fractions (samples were eluted with increasing salt concentrations: 75, 100, 125, 150, 175, 200, 350 and 500 mM KCl). Each SCX fraction was reconstituted in 0.1% Trifluoroacetic Acid (TFA) and further separated by reverse-phase liquid chromatography (Zorbax 300SB-C18, 3.5 um, 150 x 0.1 mm; flow rate 0.4ul/min, 42min/fraction). Eluted fractions were directly mixed with CHCA matrix supplemented with 10mM ammonium phosphate and spotted onto a plate for matrix-assisted laser-depolarization ionization (MALDI) mass spectrometry (192 spots/fraction). Tandem mass spectrometry analysis was performed on a 4800 MALDI TOF/TOF Analyzer (Applied Biosystems), and peptide and protein identifications were made by searching the resulting tandem mass spectra against the mouse International Protein Index (IPI) protein database using the GPS Explorer 3.6 search tool and the Mascot 2.1 search engine (Matrix Science).

## Results

### Analysis of plasma from *Lman1* null mice

Our initial characterization focused on proteomic analysis of pooled plasma samples from wild-type and *Lman1* null mice. Samples from four individuals of each genotype were pooled in an effort to average out subtle variation in plasma protein levels among individual mice (Figure 3-6). Plasma is a highly complex protein mixture that represents a particular challenge for proteomic analysis [14,15]. The top ten most abundant plasma proteins account for 90% of the total plasma protein mass, and the top twenty-two proteins account for 99.9% of the total protein mass, leaving a small fraction representing an enormous variety of lower abundance plasma proteins (Table 3-1). Immunoaffinity chromatography has been employed to selectively remove highly

abundant plasma proteins in order to increase the probability of detecting proteins of lower abundance [16,17]. Pooled plasma samples from wild-type and *Lman1* null mice were subjected to immuno-affinity chromatography using a column designed to deplete the three most abundant mouse plasma proteins (albumin, transferring and IgG). Absorbance at 280nm was monitored and yielded a profile with two peaks; an initial peak corresponding to the proteins which did not bind to the column (the immunodepleted sample) and, once an elution buffer was added to release bound proteins and clear the column, a second peak corresponding to the proteins that were bound by the depletion column. Aliquots of plasma were reserved before and after immunodepletion in order to assess the efficiency of immunodepletion. Three concentrations (1.0, 2.5 and 5.0 micrograms total protein) of crude wild-type mouse plasma, immunodepleted wild-type plasma and immunodepleted *Lman1* null plasma were separated by SDS-PAGE and visualized by staining with Coomassie dye (Figure 3-7). The most intense band, corresponding to albumin was diminished in the depleted sample; however the results showed that a significant amount of this highly abundant protein was still present in our sample. Because the bound fraction was not collected, we cannot calculate the fraction of protein that was successfully depleted. A Bradford assay was performed to measure protein concentration in the immunodepleted plasma samples, and both samples were split into duplicates to serve as analytical replicates. Tryptic peptides from duplicate wild-type samples were labeled with the 114 and 115 iTRAQ reagents and the duplicate *Lman1* null samples were labeled with the 116 and 117 iTRAQ reagents (Table 3-2). Three independent MS/MS analyses were performed on the same set of *Lman1*+/+ and *Lman1*-/- mouse plasma samples. A summary of these analyses is provided in Table 3-3.

There was substantial variability between separate analyses in the total number of peptide and protein identifications, with a range of 61-258 protein identifications with a greater than 95% confidence score. The lists of identified proteins and iTRAQ ratios also varied. For example, Apoliprotein-A1 (ApoA1) was identified in multiple experiments with a >95% confidence score; however, in one experiment, it had a *Lman1* null:wild-type iTRAQ ratio of 0.49 and in another experiment, it had a *Lman1* null:wild-type ratio of 2.10. Western blot analysis of *Lman1*+/+ and *Lman1*-/- plasma with an anti-ApoA1 antibody showed no difference in the plasma levels of ApoA1 in wild-type and *Lman1* null mice (Figure 3-8). Although FV and FVIII could potentially serve as internal controls, each protein was only identified once, in separate experiments, and only FVIII showed an iTRAQ ratio consistent with a decrease in levels in the *Lman1*-/- plasma relative to wild-type (an iTRAQ ratio of 0.58). Not surprisingly, highly abundant plasma proteins (Table 3-1) were well represented among highly confident protein identifications, including fibrinogen, alpha-2-macroglobulin, alpha-1-antitrypsin, complement C3 and haptoglobin. Interestingly, albumin, transferrin and IgG (along with the other classes of immunoglobulins) were not among the proteins we identified by MS/MS, suggesting that the immunodepletion of these top three most abundant plasma proteins was effective in reducing the amount of these three proteins present in our samples. However, a substantial proportion of proteins identified were not expected to be present in mouse plasma, such as transcription factors, nuclear pore proteins, and other intracellular proteins. The results of three independent MS/MS analyses of the same wild-type and *Lman1* null plasma samples are shown in Tables 3-4 through 3-6. In addition, a histogram plots illustrating the distribution of iTRAQ ratios for each of the

three analyses are shown in Figures 3-9 through 3-11. A log-transformation was performed such that an iTRAQ ratio of 2 is equal to 1 and an iTRAQ ratio of 0.5 is equal to -1, with the iTRAQ ratios showing no difference clustering around zero. Although we did not identify any proteins that were consistently different between *Lman1* null and wild-type mouse plasma, the distribution of ratios is much narrower for the wild-type:wild-type (115:114) and null:null (117:116) ratios and broader for the cross-comparisons of null:wild-type (116:114 and 117:114), suggesting that there is variation in protein abundance in *Lman1* null mice relative to wild-type mice.

**Analysis of purified zymogen granules from the pancreas of *Lman1* null mice**

Zymogen granules (ZG) are a specialized organelle of pancreatic acinar cells (exocrine pancreas) that serve to store inactive pancreatic enzymes (zymogens) for regulated secretion. Because LMAN1 and MCFD2 are both expressed in the pancreas, we wanted to determine whether ZG content proteins are altered in the absence of LMAN1. ZG can be isolated from the mouse pancreas using sub-cellular fractionation and density-gradient centrifugation, resulting in a highly homogeneous sample that is relatively free of contaminants from other sub-cellular compartments (Figure 3-12). Moreover, an iTRAQ-based proteomics approach has already demonstrated success in cataloging zymogen granule membrane proteins isolated from rat pancreas [13].

In pilot experiments, we determined that 100-200 micrograms of protein could be obtained from ZG purified from 100 milligrams of mouse pancreas tissue. Therefore, pooling of multiple samples is unnecessary because sufficient protein can be obtained to perform iTRAQ-labeling (50-100 micrograms of protein is needed) and MS/MS analysis on samples from individual mice. Because the activity of endogenous ZG proteases was

a concern, the ZG pellets were lysed in a buffer containing 6M urea and 0.5% SDS. Samples from two individual *Lman1-/-* mice were labeled with iTRAQ reagents 114 and 116, and samples from two individual wild-type mice were labeled with iTRAQ reagents 115 and 117. Therefore, this experiment included biological replicates (two individual mice per genotype), but no analytical replicates (Table 3-7).

Quantitative MS/MS analysis of ZG from wild-type and *Lman1-/-* mice yielded a high number of MS spectra (5075) that were selected for MS/MS fragmentation. GPS explorer-directed searches of the MS/MS spectra against the mouse IPI database with trypsin-specificity (one missed cleavage allowed; fixed N-terminal iTRAQ, lysine-iTRAQ and MMTS-cysteine modifications) resulted in a surprisingly small number of peptide identifications (220) with a greater than 95% confidence score. Subsequent Mascot database searches performed with relaxed stringency (semi-tryptic and no enzyme specificity) yielded substantial increases in the number of peptide identifications (tryptic: 167 peptides; semi-tryptic: 733 peptides; no enzyme specificity: 992 peptides). These results strongly suggest that there was endogenous protease activity that was not inhibited by the lysis buffer containing 6M urea and 0.5% SDS. A database search performed with no enzyme specificity and no requirement for an N-terminal iTRAQ modification produced very few peptide identifications indicating that nearly all peptides were efficiently labeled with the iTRAQ reagents.

Several recent proteomic characterizations have added to the list of known zymogen granule proteins [13,18]. With tryptic cleavage specified, 16 proteins were identified based on two or more non-redundant unique peptides, 14 of which are known zymogen granule proteins, including amylase, lipase, elastase and chymotrypsinogen

(Table 3-8). An additional 37 proteins were identified with greater than or equal to 95% confidence; however, many are not known to be present in zymogen granules (Table 3-9). The distribution of iTRAQ ratios for all 53 protein identifications is shown in a histogram analysis in Figure 3-13. A log-transformation was performed such that an iTRAQ ratio of 2 is equal to 1 and an iTRAQ ratio of 0.5 is equal to -1, with the iTRAQ ratios showing no difference clustering around zero. Greater variation is observed when comparing wild-type to null iTRAQ ratios versus comparing iTRAQ ratios for the two mice of the same genotype (wild-type vs. wild-type and null vs. null), such that the control samples cluster more tightly around zero and the cross-comparisons have a broader distribution. These results are suggestive of biological variation between the zymogen granule content proteins of *Lman1* null mice and wild-type mice.

**Analysis of lectin-affinity purified liver tissue lysates from *Lman1* null mice**

Whole-cell liver tissue lysates were affinity-purified on spin-columns with wheat germ agglutinin (WGA). WGA is a lectin that specifically recognizes glycoproteins containing terminal GlcNAc or sialic acid residues present in complex-type N-linked oligosaccharides found on proteins in the *medial*- and *trans*-Golgi (Figure 3-14) [19]. Bound proteins were eluted and an acetone precipitation was performed in order to remove any remaining *N*-acetylglucosamine that would have been present in the elution buffer. Because it is a primary amine, residual *N*-acetylglucosamine present in the sample could interfere with the iTRAQ-labeling process. Liver tissue lysates were purified from two mice of each genotype (two wild-type and two *Lman1* null mice) for a total of four samples. In the meantime, it became possible to perform 8-plex iTRAQ experiments; therefore, in addition to having a biological replicate for each genotype,

each of the four individual mouse samples was divided into two samples to serve as analytic replicates, and each of the eight samples was labeled with a different isobaric tag (Table 3-10).

Overall, quantitative MS/MS analysis of the WGA-affinity purified liver tissue lysates resulted in very few protein identifications. 8,347 tandem mass spectra were obtained, but less than half (31.4%) of these spectra resulted in peptide identifications with greater than 95% confidence using the Protein Pilot search engine. A search performed with no N-terminal iTRAQ modification specified returned very few protein identifications (1.95% of spectra), indicating that we achieved nearly complete labeling of the peptide mixture with the iTRAQ reagents. Even fewer spectra (13.8%) were identified when a search was performed on the same data set using the GPS Explorer search engine. In addition, many peptide identifications did not terminate in a lysine or arginine residue, as expected for a tryptic digest. Of note, the peptide spots on the MADLI plate had an atypical glassy appearance and the power of the laser had to be raised in order to detect a signal in the mass spectrometer. The physical appearance of the spots combined with the presence of a series of 5-6 peaks separated by 288 Da in the mass spectra for all eight samples suggested that an interfering substance might be present. Because residual N-acetylglucosamine from the WGA-affinity column elution buffer is a potential cause of these problems, future experiments should include an additional ultrafiltration step to further purify the eluted glycoproteins.

Of the 235 proteins that were identified with greater than or equal to 95% confidence, 115 had at least two unique peptides. These 115 proteins identifications and their corresponding iTRAQ ratios (columns with ratios comparing *Lman1* null to wild-

type are shaded in gray) are reported in Table 3-11. In this set of 115 proteins, the majority of iTRAQ ratios were close to 1, and we did not observe any consistent differences between wild-type and *Lman1 null* mice.

**Discussion**

Although the technology and instrumentation for quantitative MS/MS is progressing at a rapid pace, the analysis of complex biological protein mixtures remains a formidable challenge. In our preliminary analysis of blood plasma from *Lman1* null mice, we validated the use of MS/MS for the identification of a complex mixture of proteins. Both FV and FVIII were identified by MS/MS analysis, and a decrease of FVIII in *Lman1*-/- plasma relative to wild-type plasma was observed using quantitative mass spectrometry. However, we failed to identify other proteins that were consistently different between *Lman1*-/- and wild-type plasma samples. This could reflect the overall complexity of plasma proteins, and the limitations of analyzing complex samples by MS/MS. Subsequent efforts were directed at analyzing samples of reduced complexity, including pancreatic zymogen granules isolated by sub-cellular fractionation and liver glycoproteins purified by lectin-affinity chromatography. Minor adjustments in the sample preparation for these two sets of samples are likely to yield additional higher quality protein identifications. Based on the high level of variability and degree of difficulty in analyzing complex protein mixtures using quantitative MS/MS analysis, it will be essential to confirm any potential cargo proteins by other methods, such as western blot analysis. As techniques in mass spectrometry based proteomics continue to improve, similar analyses of samples from wild-type and *Lman1* null mice should allow

us to identify a specific subset of cargo proteins whose transport is dependent upon the LMAN1-MCFD2 pathway.

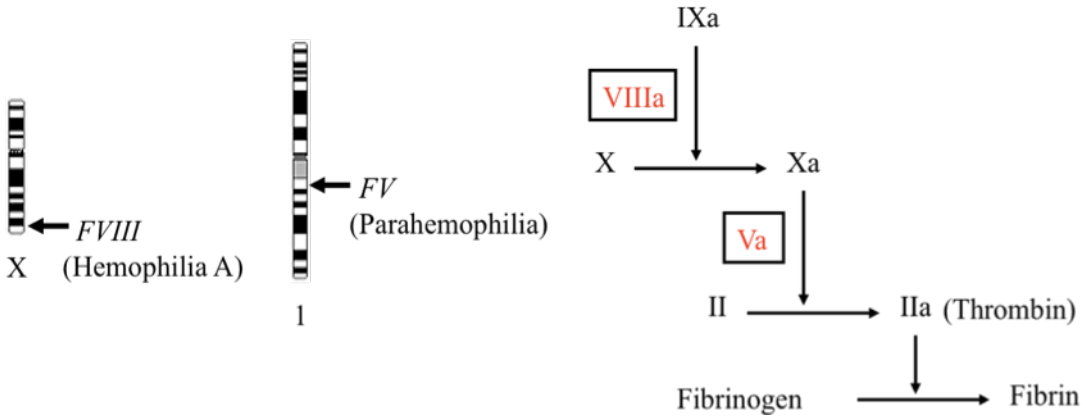


**Figure 3-1: F5F8D and the role of factor V (FV) and factor VIII (FVIII) in the coagulation cascade.**

Combined deficiency of factor V and factor VIII (F5F8D) is an autosomal recessive disorder that is distinct from co-inheritance of FV deficiency (parahemophilia) and FVIII deficiency (hemophilia A). FV and FVIII are two critical co-factors in the clotting cascade. Deficiency of FV and FVIII results in mild to moderate bleeding symptoms similar to deficiency of either factor alone.

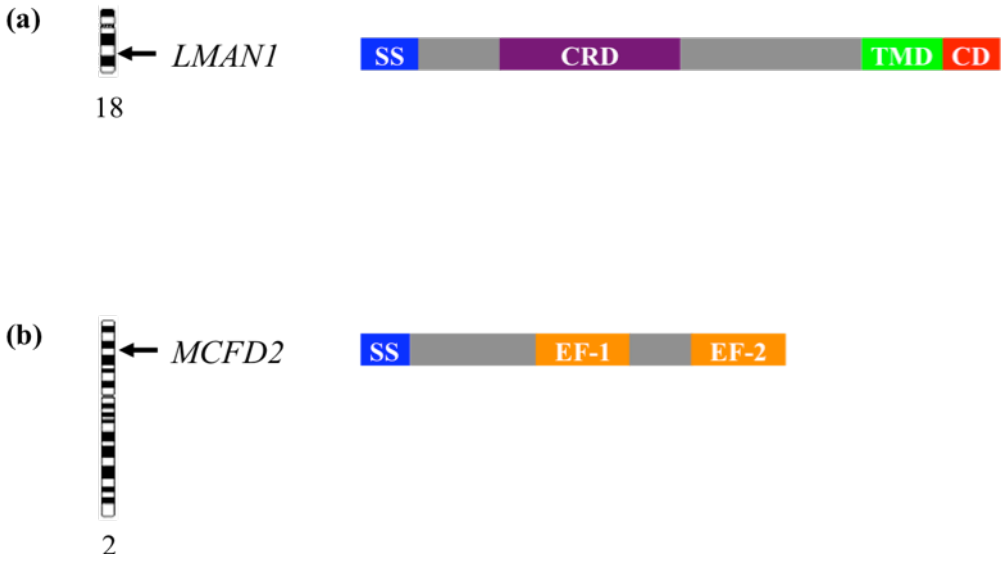


**Figure 3-2: Schematic of LMAN1 and MCFD2 proteins.**

(a) LMAN1 is a 53-kDa type-I transmembrane protein. Features include a signal sequence (SS), a carbohydrate recognition domain (CRD) that binds mannose, a transmembrane domain (TMD), and a cytoplasmic domain (CD) containing a di-phenylalanine ER exit motif and a di-lysine ER-retrieval signal. (b) MCFD2 is a 145 amino acid 16-kDa soluble protein. Features include a signal sequence (SS) and two calcium-binding EF-hand domains (EF-1 and EF-2).

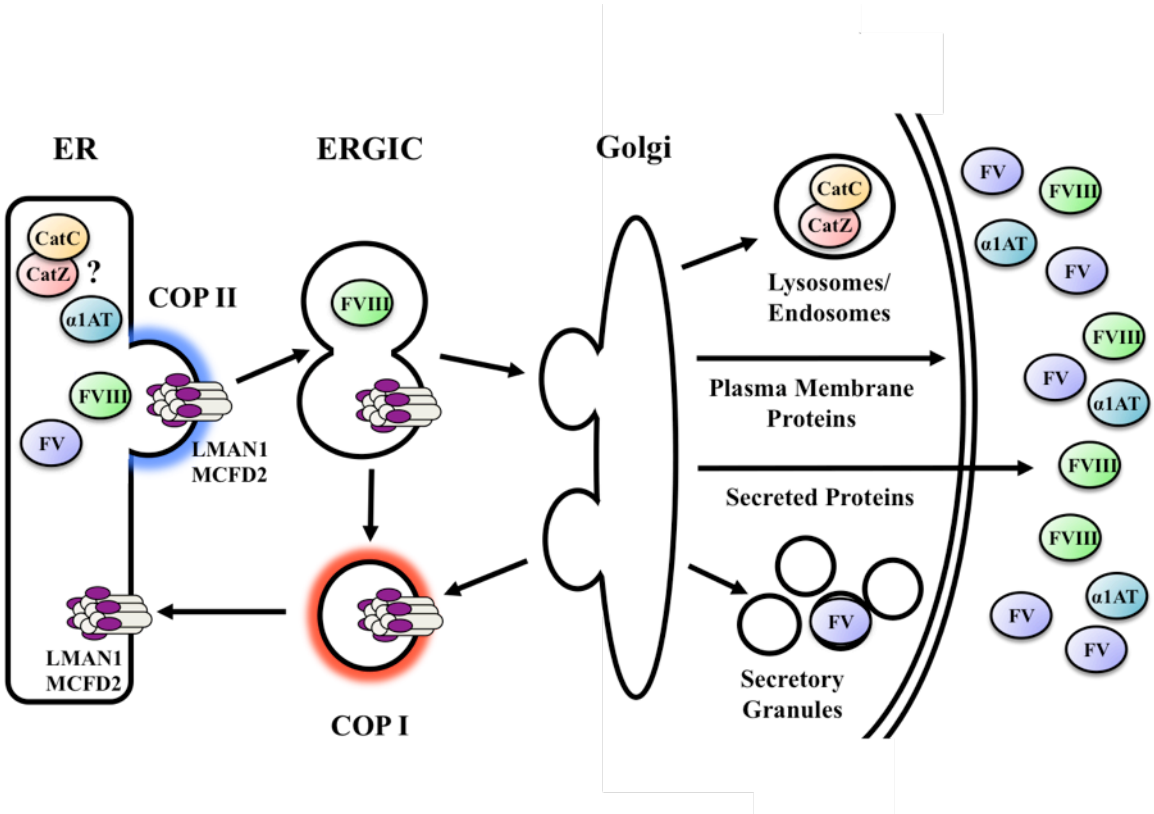


**Figure 3-3: Model for ER-to-Golgi cargo transport by the LMAN1-MCFD2 complex.**

LMAN1 and MCFD2 form a complex that cycles between the endoplasmic reticulum (ER) and ER-Golgi intermediate compartment (ERGIC) and serves as a transport receptor for factor V (FV) and factor VIII (FVIII), and possibly cathepsin C (CatC), cathepsin Z (CatZ) and α-1-antitrypsin (α1AT).

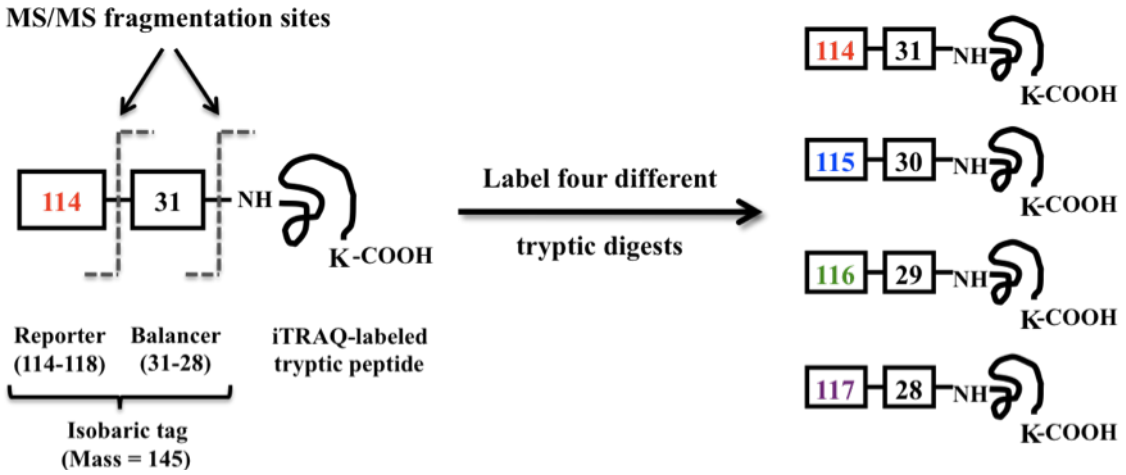


**Figure 3-4: Overview of iTRAQ labeling.**

iTRAQ (isobaric tag for relative and absolute quantitation) utilizes a set of four (shown here) or eight tags of equal mass to differentially label individual protein samples via an amine reactive group . Labeling of the same peptide found in four different samples is depicted here. In mass spectrometry (MS), tryptic peptides are analyzed, yielding a peptide mass fingerprint (PMF) spectrum. The identical tryptic peptides from each of the four samples will maintain equivalent masses. In tandem mass spectrometry (MS/MS), selected tryptic peptides are fragmented and subjected to a second round of MS analysis, producing spectra from which the amino acid sequence of the peptide can be determined (Figure 3-5). The bonds of the iTRAQ tag are labile, and subject to fragmentation during MS/MS. Neutral loss of the balancer occurs, with the positive charge remaining on the reporter (known as the iTRAQ reporter ion).

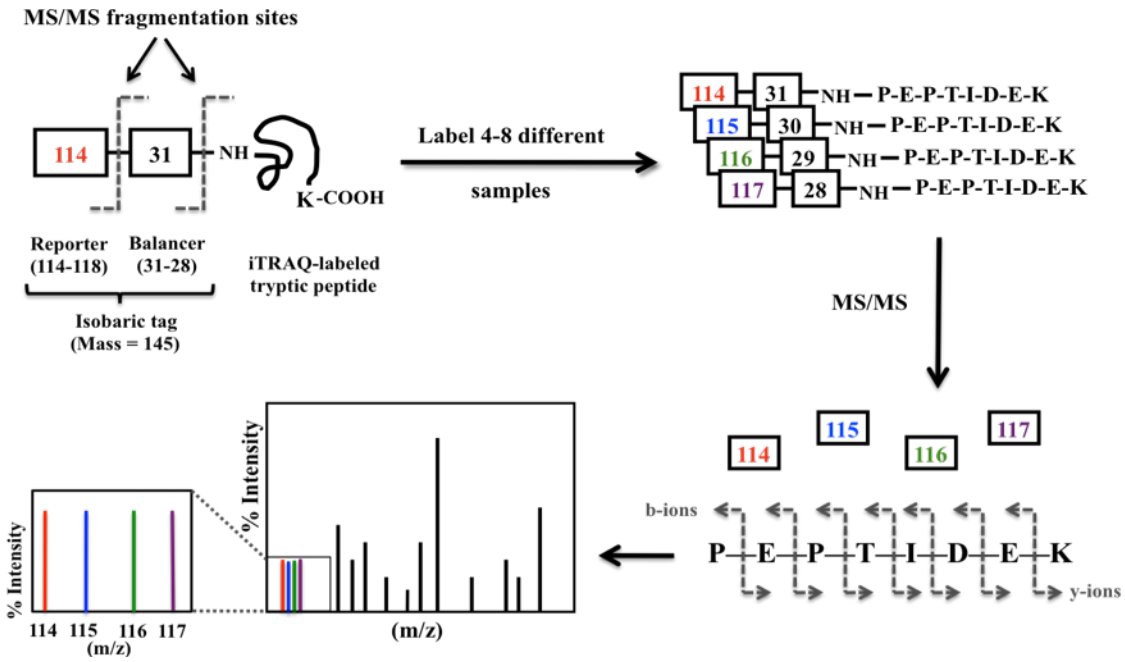


**Figure 3-5: Overview of quantitative tandem mass spectrometry analysis with iTRAQ.**

iTRAQ-labeled tryptic digests are mixed together (multiplexed) for tandem mass spectrometry analysis (MS/MS). Depicted here, a single peptide present in all four samples is selected for fragmentation, yielding a series of overlapping ions from which the amino acid sequence can be deduced. Identical peptides yield the same fragmentation spectrum. Relative peptide abundance is measured by comparing the relative intensities of the iTRAQ reporter ion peaks, found at the low m/z end of the spectrum.

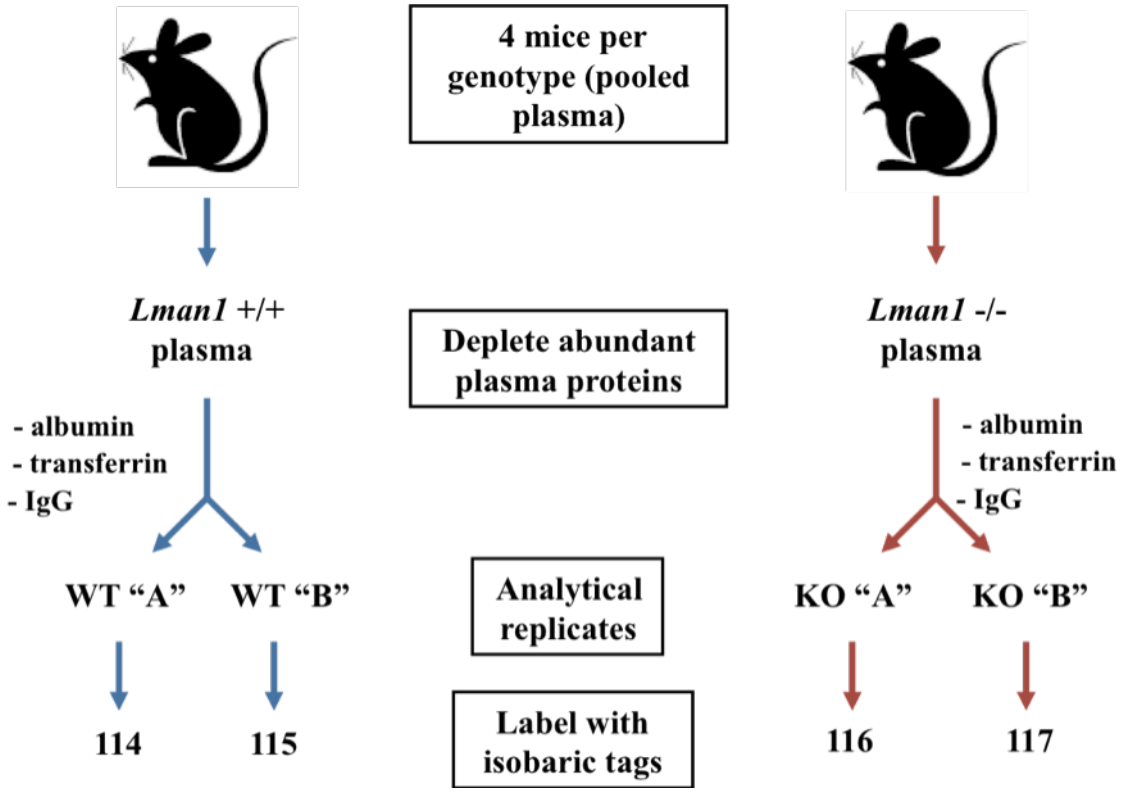


**Figure 3-6: Overview of mouse plasma sample preparation for proteomic analysis.**

Pooled plasma samples from wild-type mice were labeled with 114 and 115 (analytic replicates); pooled plasma samples from *Lman1* null (KO) mice were labeled with 116 and 117.

**Table 3-1: Abundant plasma proteins.**

| **10 most abundant plasma proteins (90% of total plasma protein mass)** | **Top 11-22 plasma proteins (account for 99% of remaining 10% from column A)** |
|---|---|
| Albumin | Apolipoprotein A1 |
| IgGs | Apolipoprotein B |
| Transferrin | Acid-1-Glycoprotein |
| Fibrinogen | Lipoprotein (a) |
| IgAs | Factor H |
| Alpha-2-Macroglobulin | Ceruloplasmin |
| IgMs | Complement C4 |
| Alpha-1-Antitrypsin | Complement Factor B |
| Complement C3 | Prealbumin |
| Haptoglobin | Complement C9 |
| | Complement C1q |
| | Complement C8 |

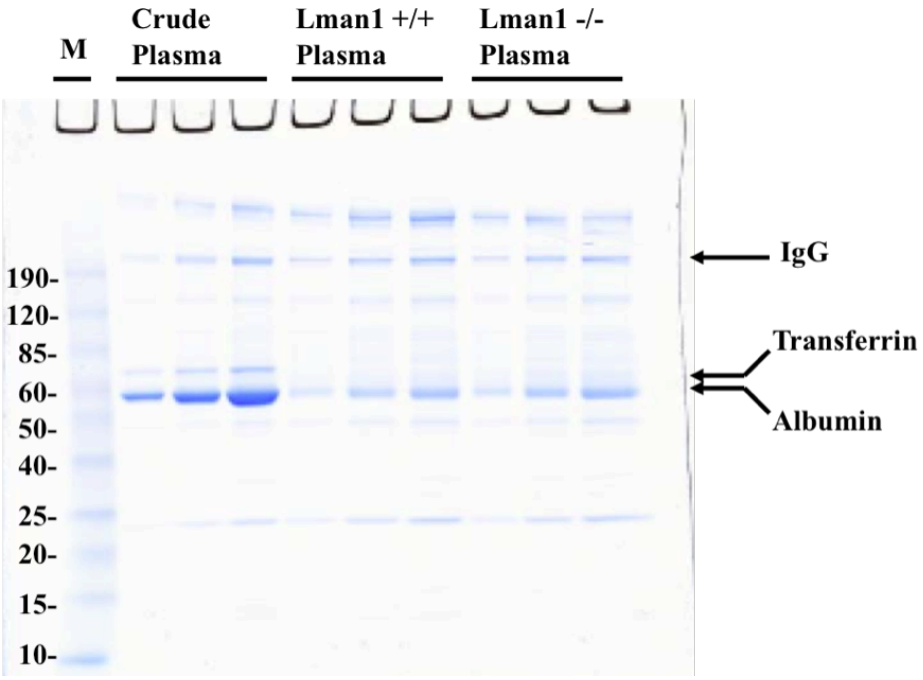


**Figure 3-7: SDS-PAGE analysis of plasma immunodepletion.**

Pooled plasma samples from wild-type mice and *Lman1* null mice were passed over a monoclonal antibody-based immunodepletion column in order to deplete the three most abundant plasma proteins (albumin, transferrin and IgG). Proteins from crude plasma and immunodepleted plasma from wild-type and *Lman1* null mice were electrophoresed on a 4-20% tris-glycine gel and visualized with Coomassie blue staining.

**Table 3-2: iTRAQ labeling of samples for MS/MS analysis of mouse plasma.**

| Mouse sample | Biological replicates | iTRAQ labels |
|---|---|---|
| 1 | *Lman1-/-* A | 114 |
| 2 | *Lman1+/+* A | 115 |
| 3 | *Lman1-/-* B | 116 |
| 4 | *Lman1+/+* B | 117 |

**Table 3-3: Summary of MS/MS analyses of mouse plasma.**

| Analysis | Protein Identifications (>95% confidence) | Protein Identifications (>95% confidence, >1 unique peptide) | FV or FVIII Identified? |
|---|---|---|---|
| 1 | 61 | 36 | no |
| 2 | 258 | 251 | FVIII only |
| 3 | 93 | 83 | no |
| 4 | >1000 | 991 | FV only |

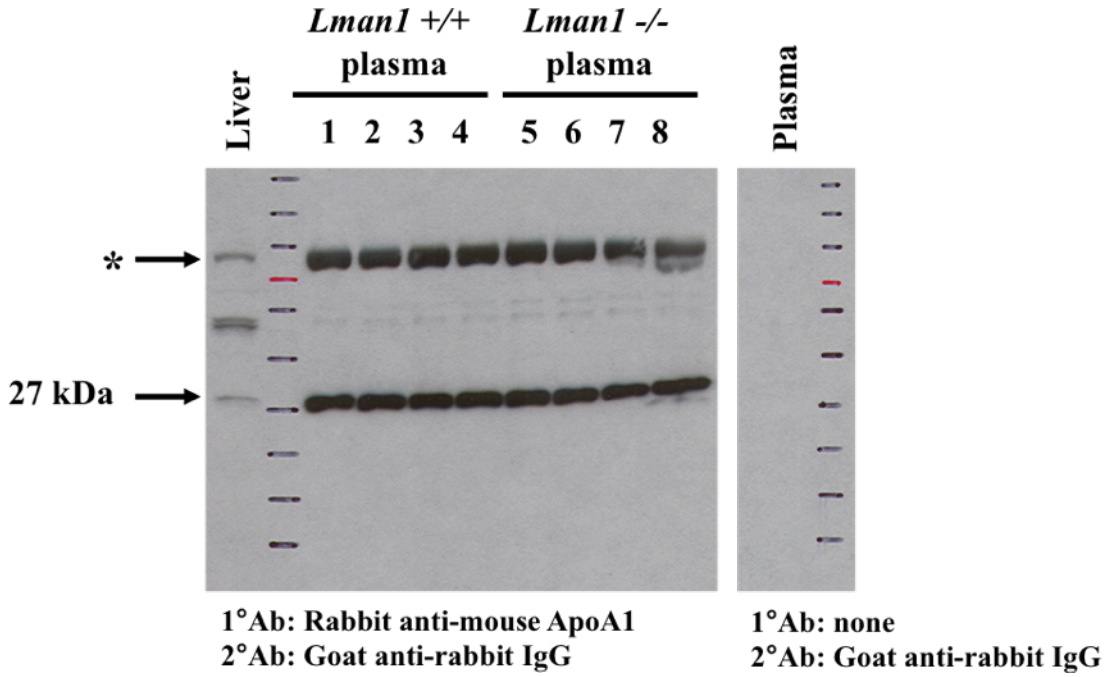


**Figure 3-8: Western blot analysis of ApoA1 in *Lman1+/+* and *Lman1-/-* mouse plasma.**

Plasma samples from each of the 4 *Lman1*+/+ mice and 4 *Lman1*-/- mice whose plasma was originally pooled for proteomic analysis were assayed for ApoA1 with a polyclonal anti-ApoA1 antibody. A whole-cell liver lysate from a wild-type mouse was assayed in parallel. A 27 kDa band, corresponding to ApoA1 was detected at similar levels in both *Lman1*+/+ and *Lman1*-/- mice. A non-specific band (noted by the asterisk) was consistent with equal sample loading. A second blot was probed in the absence of primary antibody, demonstrating that the non-specific bands observed are recognized by the anti-ApoI antibody.

**Table 3-4: Mouse plasma MS/MS analysis #1.**

| Rank | Protein Name | Accession Number | Peptide Count | Total Ion Score C.I. % | 115:114 (WT:WT) | 116:114 (KO:WT) | 117:114 (KO:WT) |
|---|---|---|---|---|---|---|---|
| 23 | Pzp protein | IPI00624663 | 7 | 100 | 0.81 | 1.00 | 0.89 |
| 94 | TiTin isoform N2-A | IPI00604969 | 7 | 100 | 0.85 | 0.80 | 0.73 |
| 7 | 186 kDa protein | IPI00624305 | 6 | 100 | 0.98 | 0.84 | 0.75 |
| 12 | Alpha-1-antitrypsin 1-4 | IPI00123924 | 6 | 100 | 0.78 | 0.98 | 0.79 |
| 215 | PREDICTED: similar to Microtubule-actin crosslinking factor 1, isoform 4 | IPI00659312 | 6 | 99 | 0.95 | 0.75 | 0.70 |
| 1 | Apolipoprotein A-I | IPI00121209 | 5 | 100 | 0.90 | 0.40 | 0.36 |
| 8 | Alpha-1-antitrypsin 1-1 | IPI00406302 | 5 | 100 | 0.86 | 0.86 | 0.75 |
| 9 | Serpina1b protein | IPI00475157 | 5 | 100 | 0.87 | 0.89 | 0.79 |
| 11 | Alpha-1-antitrypsin 1-2 | IPI00129755 | 5 | 100 | 0.78 | 0.93 | 0.76 |
| 16 | PREDICTED: similar to murinoglobulin 1 | IPI00677290 | 5 | 100 | 0.86 | 0.85 | 0.65 |
| 4 | Serine protease inhibitor A3K | IPI00131830 | 4 | 100 | 0.96 | 0.78 | 0.92 |
| 5 | Murinoglobulin 1 | IPI00330719 | 4 | 100 | 0.90 | 0.80 | 0.70 |
| 15 | Alpha-1-antitrypsin 1-5 | IPI00123927 | 4 | 100 | 0.75 | 0.71 | 0.59 |
| 17 | Hemoglobin, beta adult major chain | IPI00126208 | 4 | 100 | 0.83 | 0.90 | 0.84 |
| 21 | Fibrinogen beta chain | IPI00279079 | 4 | 100 | 0.72 | 0.91 | 0.79 |
| 6 | Fibrinogen, alpha polypeptide | IPI00115522 | 3 | 100 | 0.88 | 0.82 | 0.94 |
| 19 | Hemopexin | IPI00128484 | 3 | 100 | 0.76 | 0.87 | 0.59 |
| 20 | Vitamin D-binding protein | IPI00126184 | 3 | 100 | 0.90 | 0.78 | 0.65 |
| 50 | Chromodomain-helicase-DNA-binding protein 1 | IPI00107999 | 3 | 98 | 0.89 | 0.98 | 1.08 |
| 381 | Proteosome 26S Subunit, ATPase 2 | IPI00270326 | 3 | 96 | 0.86 | 0.72 | 0.48 |
| 81 | Fn1, Putative uncharacterized protein | IPI00352163 | 3 | 96 | 0.92 | 1.02 | 0.80 |
| 78 | Ciliary dynein heavy chain 5 | IPI00128278 | 3 | 95 | 0.90 | 0.79 | 0.82 |
| 3 | Alpha-2-HS-glycoprotein | IPI00128249 | 2 | 100 | 0.89 | 0.76 | 0.78 |
| 10 | Serum paraoxonase/arylesterase 1 | IPI00317356 | 2 | 100 | 0.76 | 0.89 | 0.65 |
| 14 | Splice Isoform LMW of Kininogen-1 | IPI00129225 | 2 | 100 | 0.72 | 0.62 | 0.59 |
| 18 | PREDICTED: similar to murinoglobulin 1 | IPI00458628 | 2 | 100 | 0.76 | 0.93 | 0.63 |
| 22 | Liver carboxylesterase N | IPI00138342 | 2 | 100 | 0.93 | 0.63 | 0.37 |
| 24 | Serpinf2, Serine proteinase inhibitor, clade F, member 2 | IPI00649122 | 2 | 100 | 0.75 | 0.67 | 0.58 |
| 31 | Fibrinogen gamma chain | IPI00122312 | 2 | 100 | 0.67 | 1.03 | 0.80 |
| 25 | PREDICTED: serine (or cysteine) proteinase inhibitor, clade A, member 3G | IPI00461418 | 2 | 100 | 0.92 | 0.88 | 0.83 |
| 41 | PREDICTED: cubilin (intrinsic factor-cobalamin receptor) isoform 1 | IPI00124058 | 2 | 100 | 0.90 | 0.62 | 0.74 |
| 38 | Splice Isoform 2 of FYVE, RhoGEF and PH domain containing protein 4 | IPI00221525 | 2 | 99 | 0.88 | 0.67 | 0.72 |
| 46 | PREDICTED: apolipoprotein B isoform 2 | IPI00666034 | 2 | 99 | 1.07 | 0.90 | 0.89 |
| 102 | Aminopeptidase puromycin sensitive | IPI00608097 | 2 | 99 | 1.09 | 0.87 | 0.85 |
| 45 | BetaIV-spectrin sigma1 | IPI00270149 | 2 | 97 | 0.70 | 1.18 | 1.25 |
| 133 | Thyrotropin receptor | IPI00111260 | 2 | 96 | 1.10 | 1.95 | 1.85 |

Protein identifications with >95% confidence and at least two unique peptides are reported. WT:WT = wild-type:wild-type, KO:WT = *Lman1* null:wild-type.

**Table 3-5: Mouse plasma MS/MS analysis #2.**

| Rank | Protein Name | Accession Number | Peptide Count | Total Ion Score C.I. % | 115:114 (WT:WT) | 116:114 (KO:WT) | 117:114 (KO:WT) |
|---|---|---|---|---|---|---|---|
| 26 | PREDICTED: similar to Nebulin | IPI00755239 | 48 | 100 | 1.03 | 1.12 | 1.09 |
| 31 | 3906 kDa protein | IPI00756257 | 38 | 100 | 1.00 | 1.07 | 1.09 |
| 1 | Isoform Long of Complement C3 | IPI00323624 | 28 | 100 | 1.04 | 0.96 | 0.93 |
| 47 | PREDICTED: similar to Microtubule-actin crosslinking factor 1 | IPI00668611 | 25 | 100 | 1.01 | 0.93 | 0.92 |
| 3 | Fibrinogen, alpha polypeptide | IPI00115522 | 21 | 100 | 1.02 | 0.91 | 0.88 |
| 46 | PREDICTED: RNA binding motif protein 16 isoform 1 | IPI00749698 | 19 | 100 | 1.07 | 1.20 | 1.17 |
| 2 | Pregnancy zone protein | IPI00624663 | 16 | 100 | 1.00 | 1.12 | 1.14 |
| 5 | Fibronectin 1 | IPI00652813 | 16 | 100 | 1.02 | 1.01 | 0.97 |
| 43 | PREDICTED: similar to very large inducible GTPase 1 isoform A | IPI00757457 | 16 | 100 | 1.03 | 0.67 | 0.69 |
| 54 | PREDICTED: similar to triple functional domain | IPI00753510 | 15 | 100 | 1.09 | 1.28 | 1.25 |
| 51 | PREDICTED: similar to Zinc finger protein Rlf | IPI00752488 | 14 | 100 | 1.00 | 1.25 | 1.21 |
| 53 | PREDICTED: similar to AT rich interactive domain 1B | IPI00750125 | 14 | 100 | 0.99 | 1.07 | 1.11 |
| 79 | PREDICTED: similar to very large inducible GTPase 1 isoform A | IPI00752883 | 14 | 100 | 1.00 | 1.03 | 1.01 |
| 6 | Murinoglobulin 1 | IPI00330719 | 13 | 100 | 0.99 | 1.14 | 1.10 |
| 55 | Synaptic nuclear envelope 1 | IPI00403993 | 13 | 100 | 0.93 | 0.94 | 0.96 |
| 78 | PREDICTED: similar to vacuolar protein sorting 13D isoform 1 | IPI00664030 | 13 | 100 | 1.06 | 1.22 | 1.22 |
| 83 | PREDICTED: similar to cubilin | IPI00750127 | 13 | 100 | 0.99 | 1.59 | 1.62 |
| 4 | Isoform LMW of Kininogen-1 | IPI00129225 | 12 | 100 | 1.00 | 0.79 | 0.79 |
| 9 | Complement component factor H | IPI00130010 | 12 | 100 | 1.06 | 1.00 | 0.96 |
| 7 | Apolipoprotein A-I | IPI00121209 | 11 | 100 | 1.00 | 2.11 | 2.09 |
| 8 | Alpha-1-antitrypsin 1-1 | IPI00406302 | 11 | 100 | 1.01 | 1.24 | 1.20 |
| 44 | PREDICTED: similar to FLJ44048 protein | IPI00352763 | 11 | 100 | 1.06 | 1.32 | 1.30 |
| 52 | PREDICTED: similar to RIKEN cDNA 5730590G19-like isoform 2 | IPI00750581 | 11 | 100 | 1.04 | 1.21 | 1.18 |
| 116 | PREDICTED: similar to | IPI00754789 | 11 | 100 | 1.08 | 1.04 | 1.04 |
| 149 | PREDICTED: similar to Probable ATP-dependent RNA helicase | IPI00754203 | 11 | 99 | 0.97 | 1.12 | 1.19 |
| 12 | Alpha-1-antitrypsin 1-5 | IPI00123927 | 10 | 100 | 0.98 | 1.26 | 1.22 |
| 16 | Murinoglobulin-2 | IPI00271262 | 10 | 100 | 0.97 | 1.12 | 1.10 |
| 66 | PREDICTED: similar to Nebulin | IPI00349292 | 10 | 100 | 0.99 | 1.10 | 1.01 |
| 77 | Absent, small, or homeotic discs 1 | IPI00553465 | 10 | 100 | 1.05 | 1.03 | 1.03 |
| 10 | Plasminogen | IPI00322936 | 9 | 100 | 0.97 | 1.01 | 1.08 |
| 11 | Alpha-1-antitrypsin 1-4 | IPI00123924 | 9 | 100 | 0.95 | 1.14 | 1.05 |
| 14 | Ceruloplasmin isoform b | IPI00775908 | 9 | 100 | 0.97 | 0.90 | 0.83 |
| 15 | Hemopexin | IPI00128484 | 9 | 100 | 1.02 | 1.40 | 1.41 |
| 17 | Fibrinogen gamma chain | IPI00122312 | 9 | 100 | 1.03 | 1.22 | 1.22 |
| 18 | Apolipoprotein A-IV | IPI00775913 | 9 | 100 | 1.04 | 1.02 | 0.99 |
| 57 | PREDICTED: similar to Midasin | IPI00751364 | 9 | 100 | 1.00 | 1.14 | 1.20 |
| 59 | PREDICTED: similar to nucleoporin 214kDa isoform 2 | IPI00673327 | 9 | 100 | 1.11 | 1.67 | 1.58 |
| 64 | PREDICTED: golgi autoantigen, golgin subfamily b | IPI00663843 | 9 | 100 | 0.98 | 0.80 | 0.73 |
| 70 | PREDICTED: similar to | IPI00625070 | 9 | 100 | 1.04 | 1.12 | 1.06 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Microtubule-actin crosslinking factor 1 | | | | | | |
| 99 | PREDICTED: similar to chromodomain helicase DNA binding protein 8 | IPI00757045 | 9 | 100 | 0.98 | 1.20 | 1.14 |
| 125 | PREDICTED: dynein heavy chain domain 3 isoform 2 | IPI00551194 | 9 | 100 | 1.03 | 0.97 | 0.97 |
| 134 | PREDICTED: myeloid/lymphoid or mixed-lineage leukemia isoform 1 | IPI00315032 | 9 | 99 | 0.99 | 0.83 | 0.89 |
| 152 | PREDICTED: similar to remodeling and spacing factor 1 isoform 12 | IPI00676351 | 9 | 99 | 1.05 | 0.74 | 0.73 |
| 243 | PREDICTED: similar to zinc finger transcription factor TRPS1 | IPI00749846 | 9 | 96 | 1.12 | 1.13 | 1.06 |
| 13 | Serpina1b protein | IPI00475157 | 8 | 100 | 1.05 | 1.22 | 1.15 |
| 19 | Alpha-1-antitrypsin 1-2 | IPI00129755 | 8 | 100 | 1.00 | 1.22 | 1.12 |
| 21 | Fibrinogen beta chain | IPI00279079 | 8 | 100 | 1.09 | 1.33 | 1.25 |
| 22 | Isoform 2 of Gelsolin | IPI00759948 | 8 | 100 | 1.02 | 0.88 | 0.86 |
| 23 | PREDICTED: similar to murinoglobulin 1 | IPI00756358 | 8 | 100 | 1.02 | 1.23 | 1.17 |
| 25 | Complement factor B | IPI00114065 | 8 | 100 | 1.01 | 1.00 | 0.98 |
| 42 | Apolipoprotein B | IPI00666034 | 8 | 100 | 1.00 | 1.16 | 1.09 |
| 62 | PREDICTED: similar to CG40351-PA.3 | IPI00756045 | 8 | 100 | 1.03 | 0.87 | 0.86 |
| 88 | PREDICTED: microtubule-actin crosslinking factor 1 isoform 2 | IPI00663180 | 8 | 100 | 1.03 | 1.10 | 1.03 |
| 93 | Isoform 3 of Usherin | IPI00742431 | 8 | 100 | 1.04 | 1.45 | 1.47 |
| 108 | PREDICTED: similar to THO complex subunit 2 (Tho2) isoform 1 | IPI00357672 | 8 | 100 | 1.11 | 1.10 | 1.03 |
| 109 | PREDICTED: similar to CG18437-PA | IPI00753249 | 8 | 100 | 1.06 | 1.42 | 1.51 |
| 126 | PREDICTED: similar to dynein heavy chain domain 3 | IPI00116310 | 8 | 100 | 1.02 | 0.95 | 0.95 |
| 148 | PREDICTED: similar to AMY-1-associating protein | IPI00749833 | 8 | 99 | 1.03 | 0.76 | 0.77 |
| 197 | PREDICTED: similar to spectrin repeat containing, nuclear envelope 2 isoform a | IPI00753411 | 8 | 98 | 0.91 | 0.83 | 0.78 |
| 258 | PREDICTED: similar to Ubiquitin-protein ligase EDD1 | IPI00753466 | 8 | 95 | 1.09 | 1.19 | 1.20 |
| 56 | Hydin protein | IPI00762600 | 7 | 100 | 1.00 | 0.83 | 0.82 |
| 68 | PREDICTED: similar to CG18437-PA | IPI00380056 | 7 | 100 | 1.19 | 0.91 | 1.07 |
| 86 | PREDICTED: similar to PDZ domain-containing RING finger protein 4 | IPI00752560 | 7 | 100 | 0.92 | 0.91 | 0.86 |
| 94 | PREDICTED: similar to Spectrin alpha chain | IPI00678951 | 7 | 100 | 0.95 | 0.71 | 0.66 |
| 130 | Zinc-finger homeodomain protein 4 | IPI00122043 | 7 | 100 | 1.08 | 1.02 | 1.05 |
| 133 | Chromosome fragility associated gene 1 | IPI00408664 | 7 | 99 | 1.06 | 1.19 | 1.14 |
| 142 | Plectin 1 | IPI00399735 | 7 | 99 | 1.06 | 1.08 | 1.03 |
| 145 | PREDICTED: similar to centromere protein F (350/400kD) isoform 4 | IPI00459137 | 7 | 99 | 1.00 | 0.90 | 0.90 |
| 151 | PREDICTED: similar to Spectrin beta chain, brain 4 | IPI00678555 | 7 | 99 | 1.16 | 1.08 | 1.07 |
| 158 | PREDICTED: similar to PHD finger protein 20 | IPI00755799 | 7 | 99 | 0.99 | 1.01 | 1.02 |
| 200 | HECT, UBA and WWE domain containing 1 | IPI00463909 | 7 | 98 | 1.10 | 0.86 | 0.89 |
| 213 | PREDICTED: zinc finger, SWIM domain containing 6 | IPI00457484 | 7 | 97 | 1.21 | 1.77 | 1.78 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 219 | PREDICTED: similar to Down-regulated in metastasis homolog isoform 4 | IPI00672476 | 7 | 97 | 0.85 | 0.99 | 0.89 |
| 241 | PREDICTED: similar to PHD finger protein 20 | IPI00751585 | 7 | 96 | 0.96 | 1.08 | 1.13 |
| 20 | Serine protease inhibitor A3K | IPI00131830 | 6 | 100 | 0.97 | 0.96 | 0.96 |
| 24 | Vitamin D-binding protein | IPI00126184 | 6 | 100 | 1.08 | 1.83 | 1.89 |
| 41 | PREDICTED: similar to Complement C4 precursor | IPI00672438 | 6 | 100 | 1.11 | 1.09 | 1.17 |
| 91 | PREDICTED: hypothetical protein LOC74041 | IPI00761400 | 6 | 100 | 1.14 | 1.02 | 1.10 |
| 100 | Adult male testis cDNA, RIKEN full-length enriched library, clone:4930516O21 | IPI00263587 | 6 | 100 | 1.14 | 1.66 | 1.52 |
| 102 | PREDICTED: similar to DNA excision repair protein ERCC-6 | IPI00756866 | 6 | 100 | 1.00 | 1.38 | 1.31 |
| 110 | A-kinase anchor protein 9 | IPI00396787 | 6 | 100 | 0.98 | 0.94 | 0.91 |
| 112 | PF6 | IPI00421240 | 6 | 100 | 1.12 | 1.34 | 1.26 |
| 117 | PREDICTED: similar to insulin receptor substrate 4 | IPI00757034 | 6 | 100 | 1.14 | 0.64 | 0.62 |
| 120 | PREDICTED: similar to Serine/threonine-protein kinase Nek1 | IPI00750013 | 6 | 100 | 1.05 | 1.23 | 1.15 |
| 131 | Syne2 protein | IPI00785250 | 6 | 100 | 0.94 | 1.06 | 1.00 |
| 153 | PREDICTED: jumonji domain containing 1C | IPI00749954 | 6 | 99 | 0.92 | 0.97 | 0.99 |
| 174 | PREDICTED: dachsous 1 | IPI00469095 | 6 | 99 | 0.94 | 1.56 | 1.38 |
| 180 | Zinc finger protein ZAS3 | IPI00757213 | 6 | 98 | 1.10 | 1.34 | 1.28 |
| 182 | Isoform IB of Formin-1 isoforms I/II/III | IPI00229709 | 6 | 98 | 0.92 | 0.91 | 0.90 |
| 187 | Talin-1 | IPI00465786 | 6 | 98 | 1.02 | 0.98 | 0.92 |
| 189 | PREDICTED: similar to BCL6 co-repressor-like 1 | IPI00756834 | 6 | 98 | 0.99 | 0.90 | 0.93 |
| 196 | Phf20 protein | IPI00752298 | 6 | 98 | 0.99 | 1.35 | 1.33 |
| 201 | PREDICTED: hypothetical protein LOC381522 | IPI00137280 | 6 | 97 | 0.98 | 0.85 | 0.88 |
| 214 | PREDICTED: similar to NACHT, leucine rich repeat and PYD containing 4A | IPI00751387 | 6 | 97 | 0.99 | 1.11 | 1.22 |
| 239 | PREDICTED: beaded filament structural protein in lens-CP94 | IPI00749974 | 6 | 96 | 1.01 | 1.15 | 1.07 |
| 240 | Similar to CTCL TUMOR ANTIGEN SE2-2 | IPI00762098 | 6 | 96 | 0.96 | 0.59 | 0.56 |
| 247 | PREDICTED: similar to Afadin | IPI00753111 | 6 | 95 | 0.98 | 1.82 | 1.83 |
| 253 | Isoform 1 of Monogenic audiogenic seizure susceptibility protein 1 | IPI00754998 | 6 | 95 | 0.97 | 1.35 | 1.16 |
| 257 | PREDICTED: similar to NACHT, leucine rich repeat and PYD containing 4A | IPI00752466 | 6 | 95 | 0.91 | 0.97 | 1.09 |
| 38 | Complement C5 | IPI00330833 | 5 | 100 | 1.14 | 0.70 | 0.74 |
| 39 | PK-120 | IPI00312711 | 5 | 100 | 1.05 | 0.90 | 0.85 |
| 40 | Vitronectin | IPI00129240 | 5 | 100 | 1.03 | 0.89 | 0.88 |
| 67 | PREDICTED: similar to remodeling and spacing factor 1 isoform 2 | IPI00664790 | 5 | 100 | 1.04 | 1.35 | 1.37 |
| 71 | PREDICTED: similar to jumonji domain containing 1B | IPI00754680 | 5 | 100 | 1.19 | 0.86 | 0.89 |
| 75 | PREDICTED: similar to CG15270-PA, isoform A | IPI00756988 | 5 | 100 | 0.96 | 1.30 | 1.27 |
| 92 | PREDICTED: similar to furry CG32045-PC, isoform C | IPI00761381 | 5 | 100 | 1.05 | 1.41 | 1.40 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 95 | 215 kDa protein | IPI00776221 | 5 | 100 | 0.91 | 0.69 | 0.70 |
| 96 | Spectrin beta chain, brain 1 | IPI00319830 | 5 | 100 | 0.99 | 1.37 | 1.40 |
| 104 | PREDICTED: similar to TATA-binding protein-associated factor 172 | IPI00676717 | 5 | 100 | 0.97 | 0.57 | 0.63 |
| 106 | PREDICTED: similar to TRAF2 and NCK-interacting kinase isoform 10 | IPI00665134 | 5 | 100 | 1.20 | 0.95 | 0.93 |
| 107 | PREDICTED: similar to Interferon-activatable protein 204 | IPI00345576 | 5 | 100 | 1.08 | 0.97 | 1.03 |
| 118 | Polypeptide N-acetylgalactosaminyltransferase 5 | IPI00223910 | 5 | 100 | 1.02 | 0.95 | 0.91 |
| 121 | MKIAA0687 protein | IPI00460720 | 5 | 100 | 1.16 | 0.92 | 0.93 |
| 124 | PREDICTED: similar to loss of heterozygosity, 11, chromosomal region 2, gene A homolog | IPI00757684 | 5 | 100 | 0.94 | 1.10 | 1.04 |
| 129 | PREDICTED: similar to kinesin-like motor protein C20orf23 | IPI00749917 | 5 | 100 | 1.10 | 1.55 | 1.44 |
| 135 | SMC6 protein | IPI00165850 | 5 | 99 | 1.01 | 0.85 | 0.85 |
| 136 | 437 kDa protein | IPI00751838 | 5 | 99 | 1.06 | 1.39 | 1.19 |
| 147 | Kinesin family member 15 | IPI00115366 | 5 | 99 | 0.87 | 0.95 | 0.99 |
| 150 | PREDICTED: similar to sodium channel 25 isoform 4 | IPI00662587 | 5 | 99 | 1.06 | 1.15 | 1.11 |
| 156 | Isoform 1 of Intersectin-2 | IPI00762547 | 5 | 99 | 0.96 | 0.89 | 0.84 |
| 160 | PREDICTED: similar to SGT1, suppressor of G2 allele of SKP1 | IPI00758198 | 5 | 99 | 1.13 | 1.20 | 1.19 |
| 162 | PREDICTED: erythrocyte protein band 4.1-like 4b | IPI00752248 | 5 | 99 | 0.98 | 0.92 | 0.87 |
| 165 | Alpha-1D adrenergic receptor | IPI00128833 | 5 | 99 | 1.20 | 1.44 | 1.29 |
| 168 | FYVE zinc finger phosphatase | IPI00129408 | 5 | 99 | 1.05 | 0.85 | 0.84 |
| 173 | PREDICTED: similar to protein tyrosine phosphatase, receptor type, Q isoform 1 | IPI00674433 | 5 | 99 | 1.10 | 0.81 | 0.79 |
| 186 | PREDICTED: CLASP1 isoform 10 | IPI00664157 | 5 | 98 | 1.15 | 1.23 | 1.22 |
| 188 | PREDICTED: similar to tetratricopeptide repeat domain 21B | IPI00751783 | 5 | 98 | 1.01 | 1.06 | 0.97 |
| 191 | Expressed sequence AU020772 | IPI00674192 | 5 | 98 | 0.99 | 1.83 | 1.77 |
| 205 | Parafibromin | IPI00170345 | 5 | 97 | 1.03 | 0.87 | 0.93 |
| 209 | Putative splicing factor, arginine/serine-rich 14 | IPI00405760 | 5 | 97 | 0.95 | 0.99 | 0.85 |
| 211 | Breast cancer type 2 susceptibility protein homolog | IPI00314969 | 5 | 97 | 1.04 | 1.69 | 1.66 |
| 220 | Isoform 1 of Myosin light chain kinase, smooth muscle | IPI00125493 | 5 | 97 | 0.97 | 1.36 | 1.33 |
| 221 | PREDICTED: similar to myosin IXA isoform 2 | IPI00675346 | 5 | 97 | 1.03 | 1.41 | 1.33 |
| 225 | Nuclear pore complex-associated intranuclear coiled-coil protein TPR | IPI00177059 | 5 | 96 | 0.96 | 0.62 | 0.62 |
| 235 | Low-density lipoprotein receptor-related protein 1B | IPI00119787 | 5 | 96 | 1.20 | 1.59 | 1.42 |
| 27 | Histidine-rich glycoprotein HRG | IPI00322304 | 4 | 100 | 1.05 | 0.76 | 0.73 |
| 29 | Antithrombin-III | IPI00136642 | 4 | 100 | 0.89 | 1.18 | 1.19 |
| 36 | Haptoglobin | IPI00409148 | 4 | 100 | 1.10 | 0.51 | 0.55 |
| 65 | Serine protease inhibitor A3F | IPI00330683 | 4 | 100 | 0.98 | 1.04 | 0.97 |
| 74 | Apolipoprotein H | IPI00648243 | 4 | 100 | 1.04 | 1.30 | 1.36 |
| 81 | Methionyl-tRNA synthetase | IPI00461469 | 4 | 100 | 0.95 | 0.43 | 0.44 |
| 87 | PREDICTED: similar to Zinc finger protein 469 | IPI00354300 | 4 | 100 | 0.90 | 0.69 | 0.70 |
| 89 | PREDICTED: similar to junctophilin 3 | IPI00757557 | 4 | 100 | 1.26 | 1.15 | 1.25 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 98 | Rotatin | IPI00379692 | 4 | 100 | 1.02 | 1.46 | 1.40 |
| 101 | PREDICTED: similar to T-cell receptor alpha chain V region PHDS58 | IPI00755734 | 4 | 100 | 1.12 | 0.88 | 0.85 |
| 103 | PREDICTED: maltase-glucoamylase | IPI00753696 | 4 | 100 | 0.89 | 0.95 | 1.01 |
| 114 | Myh2 protein | IPI00553454 | 4 | 100 | 0.99 | 0.95 | 0.92 |
| 115 | Isoform 1 of Splicing factor, arginine/serine-rich 7 | IPI00222763 | 4 | 100 | 1.40 | 1.54 | 1.31 |
| 122 | Pericentrin | IPI00457623 | 4 | 100 | 1.03 | 0.69 | 0.69 |
| 123 | Sialidase-1 precursor | IPI00315576 | 4 | 100 | 1.02 | 1.03 | 1.04 |
| 127 | Isoform 2 of Tyrosine-protein phosphatase non-receptor type 23 | IPI00606716 | 4 | 100 | 1.42 | 0.66 | 0.74 |
| 144 | PREDICTED: BAT2 domain containing 1 isoform 5 | IPI00605041 | 4 | 99 | 1.18 | 1.15 | 1.17 |
| 154 | PREDICTED: similar to polymerase (RNA) III | IPI00753896 | 4 | 99 | 1.09 | 0.67 | 0.60 |
| 155 | Isoform 2 of Ubiquitin ligase protein MIB2 | IPI00467304 | 4 | 99 | 1.09 | 1.04 | 1.07 |
| 157 | 120 kDa protein | IPI00649811 | 4 | 99 | 0.96 | 1.71 | 1.76 |
| 159 | PREDICTED: similar to latrophilin 2 | IPI00755967 | 4 | 99 | 1.10 | 0.82 | 0.85 |
| 161 | Extracellular matrix protein FRAS1 | IPI00223699 | 4 | 99 | 0.94 | 0.71 | 0.72 |
| 163 | Isoform 1 of Beta klotho | IPI00118044 | 4 | 99 | 1.07 | 0.67 | 0.67 |
| 164 | Ifi203, interferon activated gene 203 isoform 1 | IPI00127322 | 4 | 99 | 1.03 | 1.12 | 1.09 |
| 167 | GCN1 general control of amino-acid synthesis 1-like 1 | IPI00457499 | 4 | 99 | 1.05 | 1.57 | 1.60 |
| 170 | Isoform CBF1 of CCAAT/enhancer-binding protein zeta | IPI00752710 | 4 | 99 | 1.00 | 1.25 | 1.30 |
| 177 | Intraflagellar transport protein 140 | IPI00551223 | 4 | 98 | 1.04 | 1.03 | 0.94 |
| 184 | Dis3, Exosome complex exonuclease RRP44 | IPI00407764 | 4 | 98 | 0.89 | 0.71 | 0.60 |
| 198 | DNA topoisomerase 2-alpha | IPI00122223 | 4 | 98 | 0.95 | 0.86 | 0.79 |
| 203 | Zinc finger protein GLI1 | IPI00112331 | 4 | 97 | 0.92 | 0.81 | 0.75 |
| 204 | PREDICTED: similar to TBC1 domain family member 4 | IPI00667930 | 4 | 97 | 1.11 | 0.84 | 0.82 |
| 207 | PREDICTED: phosphatidylinositol 3 kinase, regulatory subunit | IPI00668264 | 4 | 97 | 1.09 | 0.98 | 0.97 |
| 210 | PREDICTED: similar to Ankyrin repeat domain protein 11 | IPI00750990 | 4 | 97 | 0.94 | 1.45 | 1.51 |
| 229 | Inositol 1,4,5-trisphosphate receptor type 3 | IPI00137400 | 4 | 96 | 1.19 | 1.36 | 1.44 |
| 230 | Isoform 4 of Protein piccolo | IPI00752192 | 4 | 96 | 1.01 | 1.12 | 1.00 |
| 236 | Alanyl-tRNA synthetase | IPI00321308 | 4 | 96 | 0.95 | 0.81 | 0.84 |
| 245 | ATP-binding cassette, sub-family D, member 3 | IPI00553576 | 4 | 96 | 1.00 | 1.44 | 1.66 |
| 246 | Guanylyl cyclase GC-E | IPI00130729 | 4 | 95 | 1.18 | 1.19 | 1.08 |
| 256 | Myosin, heavy polypeptide 3, skeletal muscle, embryonic | IPI00380895 | 4 | 95 | 0.95 | 0.96 | 0.97 |
| 28 | Alpha-2-antiplasmin | IPI00118924 | 3 | 100 | 1.30 | 1.45 | 1.28 |
| 32 | Serine protease inhibitor A3G | IPI00669817 | 3 | 100 | 0.98 | 0.97 | 0.93 |
| 34 | Plasma kallikrein | IPI00113057 | 3 | 100 | 0.99 | 1.14 | 1.22 |
| 35 | Hemoglobin beta adult major chain | IPI00126208 | 3 | 100 | 1.12 | 1.70 | 1.72 |
| 37 | Apolipoprotein E | IPI00323571 | 3 | 100 | 1.00 | 1.81 | 1.59 |
| 48 | Alpha-1-acid glycoprotein 1 | IPI00118130 | 3 | 100 | 1.00 | 0.92 | 0.90 |
| 50 | Complement factor I | IPI00320675 | 3 | 100 | 0.84 | 1.05 | 1.02 |
| 63 | Alpha-2-macroglobulin | IPI00454052 | 3 | 100 | 0.93 | 1.53 | 1.46 |
| 72 | Laminin beta-3 chain | IPI00117093 | 3 | 100 | 1.06 | 1.17 | 1.06 |
| 73 | Sex-limited protein | IPI00551206 | 3 | 100 | 1.22 | 1.08 | 1.24 |
| 76 | Isoform 2 of Probable global | IPI00314654 | 3 | 100 | 0.91 | 0.96 | 0.94 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | transcription activator SNF2L1 | | | | | | |
| 80 | PREDICTED: similar to very large inducible GTPase 1 isoform A | IPI00458896 | 3 | 100 | 1.07 | 0.84 | 0.85 |
| 82 | MKIAA0969 protein (Fragment) | IPI00625001 | 3 | 100 | 1.00 | 0.73 | 0.62 |
| 85 | 531 kDa protein | IPI00381244 | 3 | 100 | 1.08 | 1.45 | 1.47 |
| 105 | Myosin IIIA | IPI00170086 | 3 | 100 | 1.22 | 1.22 | 1.11 |
| 111 | Serum paraoxonase/arylesterase 1 | IPI00317356 | 3 | 100 | 1.06 | 0.71 | 0.92 |
| 119 | Myosin VIIa isoform 2 | IPI00551515 | 3 | 100 | 1.00 | 0.90 | 0.89 |
| 138 | Poly (ADP-ribose) polymerase family, member 6 | IPI00671644 | 3 | 99 | 0.98 | 0.98 | 0.98 |
| 139 | Myosin-9 | IPI00123181 | 3 | 99 | 0.97 | 0.70 | 0.75 |
| 140 | 1200016B10Rik protein | IPI00453902 | 3 | 99 | 1.04 | 1.66 | 1.66 |
| 143 | 52 kDa protein | IPI00754372 | 3 | 99 | 0.96 | 1.12 | 1.12 |
| 146 | Stabilin-2 | IPI00467600 | 3 | 99 | 1.05 | 1.40 | 1.37 |
| 166 | PREDICTED: similar to cyclin fold protein 1 isoform 2 | IPI00125016 | 3 | 99 | 1.15 | 1.09 | 1.26 |
| 169 | Human immunodeficiency virus type I enhancer-binding protein 2 homolog | IPI00115900 | 3 | 99 | 1.04 | 0.91 | 0.86 |
| 171 | Hypothetical Pleckstrin homology (PH) domain containing protein | IPI00227244 | 3 | 99 | 1.10 | 0.88 | 1.06 |
| 181 | Clathrin, heavy polypeptide | IPI00648173 | 3 | 98 | 0.92 | 1.25 | 1.12 |
| 183 | Immediate early response gene 5 protein | IPI00131538 | 3 | 98 | 1.08 | 1.23 | 1.27 |
| 185 | Cullin-4A | IPI00321407 | 3 | 98 | 0.99 | 1.01 | 1.01 |
| 190 | Isoform 1 of Protein C13orf7 homolog | IPI00322413 | 3 | 98 | 1.08 | 1.07 | 1.05 |
| 192 | PREDICTED: similar to FLJ46154 protein | IPI00750482 | 3 | 98 | 0.95 | 1.09 | 1.03 |
| 193 | PREDICTED: similar to Probable ubiquitin carboxyl-terminal hydrolase FAF-X | IPI00752867 | 3 | 98 | 0.96 | 0.67 | 0.62 |
| 194 | Ccdc33, Isoform 1 of Coiled-coil domain-containing protein 33 | IPI00651987 | 3 | 98 | 0.95 | 1.35 | 1.31 |
| 195 | PREDICTED: similar to microtubule associated serine/threonine kinase | IPI00762779 | 3 | 98 | 0.93 | 1.18 | 1.15 |
| 199 | Gpsm2, LGN protein | IPI00165828 | 3 | 98 | 1.09 | 1.34 | 1.22 |
| 206 | PREDICTED: similar to cyclin fold protein 1 | IPI00754210 | 3 | 97 | 1.12 | 0.97 | 1.08 |
| 212 | PREDICTED: similar to Hephaestin | IPI00749896 | 3 | 97 | 1.04 | 1.34 | 1.34 |
| 217 | Fpgt protein | IPI00420488 | 3 | 97 | 0.98 | 0.85 | 0.81 |
| 223 | Weakly Similar to Chlamydomonas Reinhardtii Strain 1132D- Flagellar Protofilament Ribbon Protein | IPI00132357 | 3 | 97 | 1.20 | 2.00 | 1.35 |
| 227 | Testicular microtubules-related protein 4 | IPI00321122 | 3 | 96 | 1.07 | 0.97 | 0.93 |
| 228 | PREDICTED: hypothetical protein LOC70802 isoform 1 | IPI00649452 | 3 | 96 | 1.10 | 0.94 | 1.03 |
| 234 | Ankrd22, Ankyrin repeat domain-containing protein 22 | IPI00319036 | 3 | 96 | 1.09 | 1.29 | 1.24 |
| 237 | Potassium voltage-gated channel subfamily H member 5 | IPI00223584 | 3 | 96 | 1.03 | 0.88 | 0.86 |
| 249 | Coiled-coil alpha-helical rod protein | IPI00400252 | 3 | 95 | 1.04 | 0.84 | 0.86 |
| 250 | Rpap1, Isoform 1 of RNA polymerase II-associated protein 1 | IPI00377618 | 3 | 95 | 0.94 | 0.79 | 0.83 |
| 254 | Isoform 2 of Myosin-14 | IPI00608077 | 3 | 95 | 0.97 | 0.35 | 0.32 |
| 30 | Alpha-2-HS-glycoprotein | IPI00128249 | 2 | 100 | 0.80 | 0.93 | 1.06 |
| 33 | Hemoglobin subunit alpha | IPI00469114 | 2 | 100 | 0.94 | 0.91 | 0.83 |
| 45 | Complement component C8 alpha | IPI00453977 | 2 | 100 | 0.95 | 0.70 | 0.77 |
| 49 | Carboxypeptidase N, polypeptide 2 | IPI00119522 | 2 | 100 | 1.16 | 1.19 | 1.05 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | homolog | | | | | | |
| 58 | Glycosylphosphatidylinositol specific phospholipase D1 | IPI00225715 | 2 | 100 | 0.84 | 1.12 | 1.00 |
| 60 | Liver carboxylesterase N | IPI00138342 | 2 | 100 | 0.83 | 0.90 | 0.92 |
| 61 | Zinc-alpha-2-glycoprotein | IPI00137987 | 2 | 100 | 0.93 | 0.70 | 0.73 |
| 84 | GUGU alpha | IPI00469387 | 2 | 100 | 0.95 | 1.19 | 1.05 |
| 97 | Isoform 2 of Alkylated repair protein alkB homolog 4 | IPI00674394 | 2 | 100 | 0.98 | 1.51 | 1.55 |
| 113 | 5-aminolevulinate synthase, erythroid-specific, mitochondrial | IPI00135065 | 2 | 100 | 1.04 | 0.56 | 0.60 |
| 128 | Isoform 2 of REST corepressor 3 | IPI00742420 | 2 | 100 | 1.06 | 0.87 | 0.84 |
| 137 | Ankar, Ankyrin and armadillo repeat-containing protein | IPI00223907 | 2 | 99 | 1.45 | 2.03 | 2.01 |
| 141 | 94 kDa protein | IPI00475247 | 2 | 99 | 0.93 | 1.05 | 1.04 |
| 172 | Erythroid differentiation-related factor 1 | IPI00454066 | 2 | 99 | 1.11 | 0.99 | 0.97 |
| 175 | Isoform 1 of Complement component C8 beta chain | IPI00121274 | 2 | 99 | 1.19 | 1.66 | 1.67 |
| 179 | Isoform 1 of PHD finger protein 8 | IPI00109676 | 2 | 98 | 1.04 | 1.35 | 1.35 |
| 208 | PREDICTED: similar to Rab3 GTPase-activating protein non-catalytic subunit (Rab3 GTPase-activating | IPI00756920 | 2 | 97 | 1.03 | 2.24 | 2.16 |
| 215 | Epidermal growth factor receptor isoform 2 | IPI00411099 | 2 | 97 | 1.06 | 1.32 | 1.19 |
| 218 | Organic cation transporter 2 | IPI00119459 | 2 | 97 | 0.95 | 1.15 | 1.11 |
| 222 | Coagulation factor X | IPI00331551 | 2 | 97 | 1.04 | 0.86 | 0.81 |
| 224 | PREDICTED: similar to proteasome 26S non-ATPase subunit 2 isoform 1 | IPI00675436 | 2 | 96 | 0.98 | 0.70 | 0.73 |
| 226 | Ribonuclease P protein subunit p20 | IPI00120909 | 2 | 96 | 1.10 | 1.51 | 1.34 |
| 231 | PREDICTED: similar to complement factor H-related protein B | IPI00671376 | 2 | 96 | 1.04 | 0.62 | 0.59 |
| 232 | Kinase non-catalytic C-lobe domain (KIND) containing 1 | IPI00380321 | 2 | 96 | 1.23 | 1.13 | 0.98 |
| 233 | Interleukin-6 | IPI00133702 | 2 | 96 | 1.24 | 2.12 | 2.00 |
| 238 | ATP-dependent RNA helicase DDX51 | IPI00396728 | 2 | 96 | 1.11 | 1.46 | 1.42 |
| 242 | PREDICTED: similar to PRAME family member 9 | IPI00678250 | 2 | 96 | 0.89 | 1.42 | 1.24 |
| 244 | Carboxypeptidase N catalytic chain | IPI00128260 | 2 | 96 | 0.98 | 0.90 | 0.83 |
| 248 | PREDICTED: similar to DnaJ | IPI00756157 | 2 | 95 | 1.05 | 1.13 | 1.09 |
| 251 | Adenylate cyclase 5 | IPI00343381 | 2 | 95 | 1.45 | 2.03 | 2.01 |
| 252 | PREDICTED: similar to immunity-related GTPase family, cinema 1 isoform 3 | IPI00667234 | 2 | 95 | 0.95 | 1.22 | 1.18 |
| 255 | MKIAA2002 protein | IPI00136917 | 2 | 95 | 1.12 | 0.96 | 0.92 |

Protein identifications with >95% confidence and at least two unique peptides are reported. WT:WT = wild-type:wild-type, KO:WT = *Lman1* null:wild-type.

**Table 3-6: Mouse plasma MS/MS analysis #3.**

| Rank | Protein Identification (Name) | Accession Number | Peptide Count | Total Ion Score C.I. % | 115:114 (WT:WT) | 116:114 (KO:WT) | 117:114 (KO:WT) |
|---|---|---|---|---|---|---|---|
| 1 | Isoform Long of Complement C3 | IPI00323624 | 16 | 100 | 1.04 | 0.99 | 0.95 |
| 22 | 3906 kDa protein | IPI00756257 | 16 | 100 | 0.99 | 0.98 | 0.92 |
| 29 | Similar to Nebulin | IPI00755239 | 16 | 100 | 1.06 | 1.06 | 0.96 |
| 2 | Apolipoprotein A-I | IPI00121209 | 10 | 100 | 1.00 | 0.99 | 0.93 |
| 4 | Pregnancy zone protein | IPI00624663 | 9 | 100 | 0.98 | 0.97 | 0.93 |
| 32 | Similar to very large inducible GTPase 1 isoform A | IPI00752883 | 9 | 100 | 0.99 | 0.98 | 1.02 |
| 5 | Fibrinogen, alpha polypeptide | IPI00115522 | 8 | 100 | 1.07 | 1.00 | 0.98 |
| 8 | Alpha-2-macroglobulin | IPI00126194 | 8 | 100 | 0.99 | 0.99 | 0.96 |
| 3 | Fibrinogen beta chain | IPI00279079 | 7 | 100 | 1.01 | 0.92 | 0.96 |
| 9 | Alpha-1-antitrypsin 1-1 | IPI00406302 | 7 | 100 | 1.02 | 1.07 | 1.03 |
| 36 | Similar to Zinc finger protein Rlf | IPI00752488 | 7 | 100 | 1.00 | 0.68 | 0.57 |
| 50 | Similar to THO complex subunit 2 (Tho2) isoform 1 | IPI00357672 | 7 | 100 | 0.99 | 0.98 | 0.94 |
| 6 | Apolipoprotein E | IPI00323571 | 6 | 100 | 1.02 | 1.03 | 0.97 |
| 10 | Hemopexin | IPI00128484 | 6 | 100 | 1.04 | 1.08 | 1.20 |
| 12 | Murinoglobulin-1 | IPI00123223 | 6 | 100 | 1.00 | 1.03 | 0.84 |
| 40 | Similar to insulin receptor substrate 4 | IPI00757034 | 6 | 100 | 1.04 | 0.97 | 0.98 |
| 42 | Similar to furry CG32045-PC, isoform C | IPI00761381 | 6 | 100 | 1.07 | 1.00 | 0.92 |
| 46 | 437 kDa protein | IPI00751838 | 6 | 100 | 1.05 | 1.02 | 1.08 |
| 54 | Similar to Ubiquitin-protein ligase EDD1 | IPI00753466 | 6 | 100 | 1.10 | 1.00 | 0.90 |
| 78 | RNA binding motif protein 16 isoform 1 | IPI00749698 | 6 | 97 | 0.93 | 0.91 | 0.83 |
| 11 | Liver carboxylesterase N | IPI00138342 | 5 | 100 | 1.07 | 0.99 | 0.98 |
| 13 | Alpha-1-antitrypsin 1-4 | IPI00123924 | 5 | 100 | 1.06 | 1.06 | 1.03 |
| 17 | Similar to murinoglobulin 1 | IPI00756358 | 5 | 100 | 1.02 | 1.00 | 0.86 |
| 48 | Similar to dynein heavy chain domain 3 | IPI00661771 | 5 | 100 | 1.04 | 0.98 | 0.96 |
| 57 | Similar to remodeling and spacing factor 1 isoform 12 | IPI00676351 | 5 | 99 | 1.01 | 1.05 | 0.86 |
| 64 | Ciliary dynein heavy chain 5 | IPI00128278 | 5 | 99 | 1.20 | 1.20 | 0.97 |
| 77 | Similar to Microtubule-actin crosslinking factor 1 | IPI00668611 | 5 | 97 | 1.02 | 0.95 | 0.93 |
| 7 | Serine protease inhibitor A3K | IPI00131830 | 4 | 100 | 1.12 | 1.02 | 0.86 |
| 14 | Ceruloplasmin isoform b | IPI00775908 | 4 | 100 | 1.09 | 1.27 | 1.02 |
| 16 | Antithrombin-III precursor | IPI00136642 | 4 | 100 | 0.93 | 1.15 | 0.84 |
| 18 | Fibrinogen gamma chain | IPI00122312 | 4 | 100 | 1.13 | 1.39 | 1.20 |
| 26 | Vitamin D-binding protein | IPI00126184 | 4 | 100 | 0.99 | 1.06 | 1.02 |
| 35 | Isoform 3 of Vacuolar protein sorting-associated protein 13C | IPI00661059 | 4 | 100 | 1.01 | 1.03 | 0.99 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 58 | Similar to Laminin alpha-3 chain | IPI00750350 | 4 | 99 | 1.09 | 0.95 | 0.81 |
| 65 | Similar to PDZ domain-containing RING finger protein 4 | IPI00752560 | 4 | 99 | 1.03 | 1.00 | 0.81 |
| 67 | Isoform 1 of Citron Rho-interacting kinase | IPI00406762 | 4 | 98 | 1.07 | 0.92 | 0.85 |
| 15 | Fn1, Putative uncharacterized protein | IPI00352163 | 3 | 100 | 0.79 | 0.69 | 0.71 |
| 19 | 6 kDa protein | IPI00762198 | 3 | 100 | 1.04 | 0.84 | 0.88 |
| 20 | Alpha-2-antiplasmin | IPI00118924 | 3 | 100 | 1.01 | 0.57 | 0.48 |
| 23 | Murinoglubulin 4 | IPI00420891 | 3 | 100 | 1.11 | 1.08 | 0.84 |
| 24 | Alpha-2-macroglobulin | IPI00454052 | 3 | 100 | 0.96 | 0.87 | 0.78 |
| 28 | Plasminogen | IPI00322936 | 3 | 100 | 1.00 | 1.00 | 1.04 |
| 30 | Complement component factor H | IPI00130010 | 3 | 100 | 0.98 | 1.02 | 1.00 |
| 31 | Isoform 2 of Gelsolin | IPI00759948 | 3 | 100 | 0.97 | 1.09 | 0.73 |
| 37 | LEK1 | IPI00404359 | 3 | 100 | 1.03 | 0.92 | 0.91 |
| 38 | 215 kDa protein | IPI00776221 | 3 | 100 | 1.10 | 1.19 | 1.06 |
| 41 | Vitronectin | IPI00129240 | 3 | 100 | 1.02 | 0.93 | 0.94 |
| 44 | Similar to CG4329-PA, isoform A | IPI00468360 | 3 | 100 | 1.07 | 0.93 | 0.90 |
| 45 | Isoform 1 of Fanconi anemia group M protein homolog | IPI00471330 | 3 | 100 | 0.92 | 0.94 | 0.97 |
| 47 | Hypothetical protein LOC105841 isoform 1 | IPI00464312 | 3 | 100 | 1.18 | 0.94 | 0.83 |
| 55 | Apolipoprotein A-IV | IPI00775913 | 3 | 100 | 1.09 | 1.34 | 1.32 |
| 66 | Cytoskeleton-associated protein 4 | IPI00223047 | 3 | 98 | 1.02 | 1.05 | 1.02 |
| 72 | Similar to THO complex subunit 2 | IPI00665626 | 3 | 98 | 0.88 | 0.88 | 0.69 |
| 76 | Similar to Kinesin-like protein KIF14 | IPI00757807 | 3 | 97 | 1.07 | 0.97 | 1.02 |
| 80 | Protocadherin-12 precursor | IPI00115860 | 3 | 96 | 1.14 | 1.07 | 0.98 |
| 82 | 126 kDa protein | IPI00761746 | 3 | 96 | 1.02 | 0.94 | 0.87 |
| 85 | MKIAA0890 protein | IPI00403301 | 3 | 96 | 1.00 | 1.04 | 0.82 |
| 92 | Rearranged L-myc fusion sequence | IPI00750196 | 3 | 95 | 1.01 | 1.11 | 1.06 |
| 93 | Similar to spectrin repeat containing, nuclear envelope 2 isoform a | IPI00753411 | 3 | 95 | 0.92 | 0.85 | 0.87 |
| 87 | Similar to CG4841-PA | IPI00649142 | 3 | 95 | 1.11 | 1.01 | 0.99 |
| 21 | Serpina3g protein | IPI00762685 | 2 | 100 | 1.01 | 0.99 | 0.80 |
| 25 | Complement factor B | IPI00114065 | 2 | 100 | 0.98 | 0.86 | 0.79 |
| 27 | Histidine-rich glycoprotein | IPI00322304 | 2 | 100 | 0.94 | 0.89 | 0.74 |
| 33 | Epidermal growth factor receptor isoform 2 | IPI00411099 | 2 | 100 | 1.24 | 1.23 | 0.85 |
| 34 | Isoform LMW of Kininogen-1 | IPI00129225 | 2 | 100 | 1.19 | 1.20 | 1.22 |
| 49 | Similar to Filaggrin | IPI00338382 | 2 | 100 | 1.00 | 0.50 | 0.47 |
| 51 | Similar to Kif19A CG9913-PB, isoform B | IPI00461313 | 2 | 100 | 1.25 | 1.09 | 0.88 |
| 52 | 4932418E24Rik, Putative uncharacterized protein | IPI00226099 | 2 | 100 | 0.98 | 1.01 | 0.94 |
| 53 | WD repeat protein 60 | IPI00226499 | 2 | 100 | 1.05 | 0.93 | 0.94 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 56 | Kng1 protein | IPI00656329 | 2 | 100 | 0.97 | 1.19 | 0.71 |
| 60 | Protein tyrosine phosphatase, receptor type, f polypeptide | IPI00674306 | 2 | 99 | 0.94 | 0.79 | 0.73 |
| 62 | Similar to neonatal submandibular gland protein C | IPI00755346 | 2 | 99 | 1.01 | 0.80 | 0.78 |
| 63 | BAT2 domain containing 1 isoform 2 | IPI00669060 | 2 | 99 | 1.02 | 0.48 | 0.43 |
| 68 | Similar to myosin IXA isoform 2 | IPI00675346 | 2 | 98 | 1.02 | 1.07 | 0.98 |
| 69 | Protein | IPI00751643 | 2 | 98 | 1.08 | 1.22 | 1.11 |
| 71 | Origin recognition complex subunit 1 isoform B | IPI00556829 | 2 | 98 | 1.20 | 0.98 | 0.95 |
| 75 | Uncharacterized protein KIAA1666 | IPI00354151 | 2 | 97 | 1.05 | 0.59 | 0.54 |
| 79 | Hypothetical Glutamine-rich region profile containing protein | IPI00761895 | 2 | 97 | 1.06 | 0.92 | 0.91 |
| 81 | Uncharacterized protein C18orf34 homolog | IPI00228949 | 2 | 96 | 1.10 | 0.85 | 0.95 |
| 83 | Flotillin-1 | IPI00117181 | 2 | 96 | 1.15 | 1.32 | 1.27 |
| 84 | MFLJ00279 protein | IPI00788324 | 2 | 96 | 1.07 | 1.02 | 1.01 |
| 86 | Fc epsilon receptor II | IPI00127664 | 2 | 96 | 0.94 | 0.92 | 0.84 |
| 88 | Similar to FLJ44048 protein | IPI00352763 | 2 | 96 | 1.08 | 0.94 | 1.05 |

Protein identifications with >95% confidence and at least two unique peptides are reported. WT:WT = wild-type:wild-type, KO:WT = *Lman1* null:wild-type.

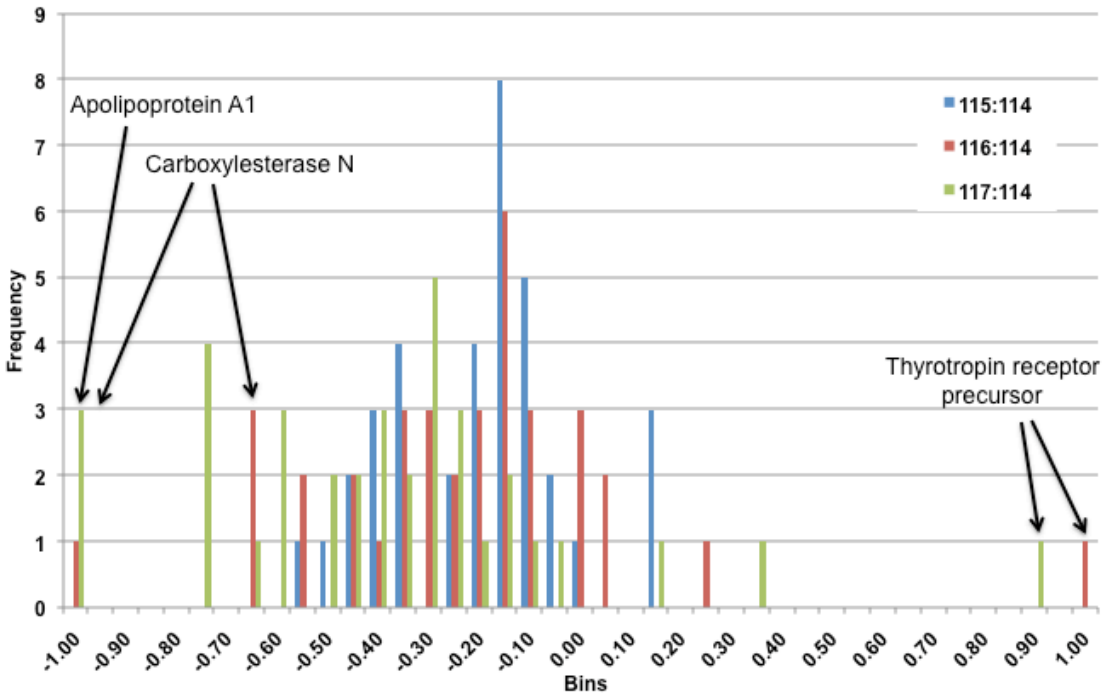


**Figure 3-9: iTRAQ ratio distribution for analysis of mouse plasma (#1)**

Pooled plasma samples from wild-type mice were labeled with 114 and 115 (analytic replicates); pooled plasma samples from *Lman1* null mice were labeled with 116 and 117. The 115:114 and 117:116 ratios represent duplicate comparisons for the wild-type and null samples, respectively. Cross comparisons (null:wild-type) are shown as the 116:114 and 117:114 ratios. The numbers on the x-axis represent a log transformation of raw iTRAQ ratios. iTRAQ ratios are shown for all protein identifications with >95% confidence and at least 2 unique peptides. A few of the corresponding protein identifications are depicted on the graph (see Table 3-4).

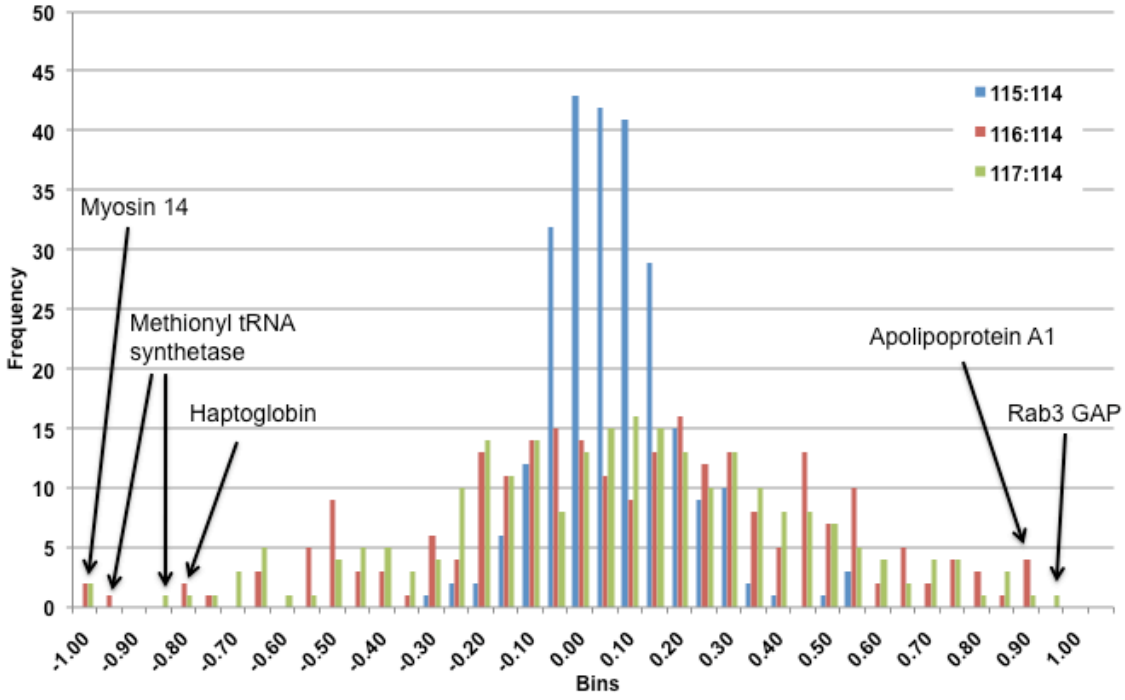


**Figure 3-10: iTRAQ ratio distribution for analysis of mouse plasma (#2)**

Pooled plasma samples from wild-type mice were labeled with 114 and 115 (analytic replicates); pooled plasma samples from *Lman1* null mice were labeled with 116 and 117. The 115:114 and 117:116 ratios represent duplicate comparisons for the wild-type and null samples, respectively. Cross comparisons (null:wild-type) are shown as the 116:114 and 117:114 ratios. The numbers on the x-axis represent a log transformation of raw iTRAQ ratios. iTRAQ ratios are shown for all protein identifications with >95% confidence and at least 2 unique peptides. A few of the corresponding protein identifications are depicted on the graph (see Table 3-5).

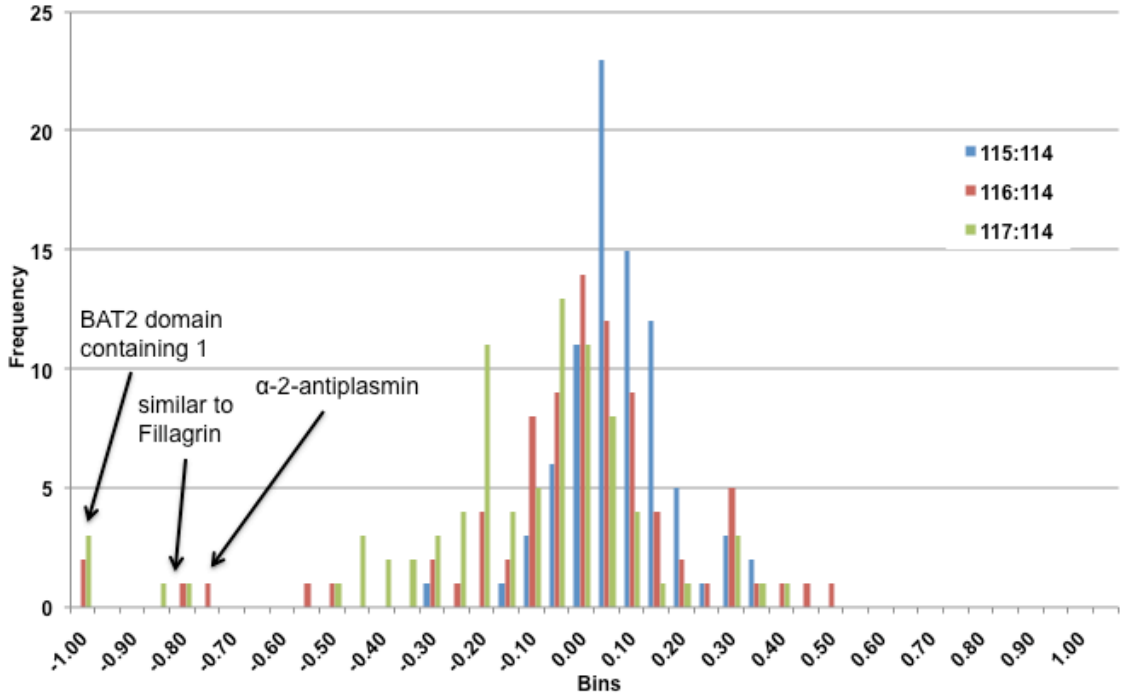


**Figure 3-11: iTRAQ ratio distribution for analysis of mouse plasma (#3)**

Pooled plasma samples from wild-type mice were labeled with 114 and 115 (analytic replicates); pooled plasma samples from *Lman1* null mice were labeled with 116 and 117. The 115:114 and 117:116 ratios represent duplicate comparisons for the wild-type and null samples, respectively. Cross comparisons (null:wild-type) are shown as the 116:114 and 117:114 ratios. The numbers on the x-axis represent a log transformation of raw iTRAQ ratios. iTRAQ ratios are shown for all protein identifications with >95% confidence and at least 2 unique peptides. A few of the corresponding protein identifications are depicted on the graph (see Table 3-6).

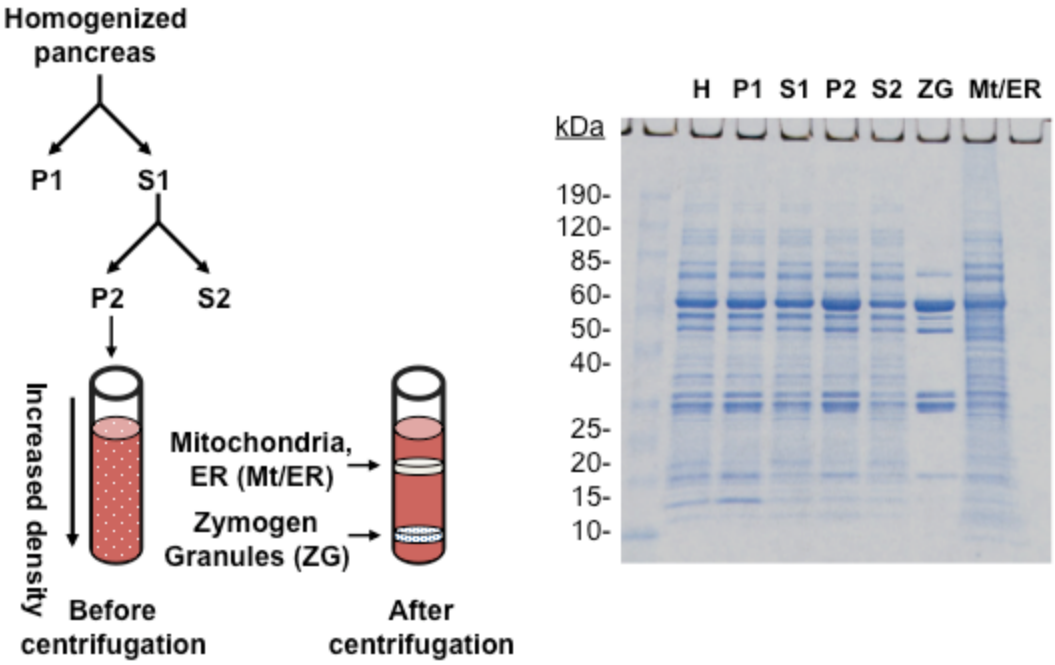


**Figure 3-12: Purification of zymogen granules from mouse pancreas for proteomic analysis.**

A schematic of the purification procedure is shown on the left. On the right, a Coomassie-stained gel shows the reduced protein complexity of the purified zymogen granules.

**Table 3-7: iTRAQ labeling of samples for MS/MS analysis of mouse zymogen granules.**

| Mouse sample | Biological replicates | iTRAQ labels |
|---|---|---|
| 1 | *Lman1-/-* A | 114 |
| 2 | *Lman1+/+* A | 115 |
| 3 | *Lman1-/-* B | 116 |
| 4 | *Lman1+/+* B | 117 |

**Table 3-8: MS/MS analysis of pancreatic zymogen granules.**

| Rank | Protein Name | Accession Number | Peptide Count | Total Ion Score C.I. % | 115:114 (WT:KO) | 116:114 (KO:KO) | 117:114 (WT:KO) |
|---|---|---|---|---|---|---|---|
| 1 | Amylase 2, pancreatic | IPI00756078 | 12 | 100 | 1.24 | 1.10 | 1.10 |
| 2 | Amylase 2, pancreatic | IPI00775996 | 12 | 100 | 1.25 | 1.07 | 1.13 |
| 6 | Alpha-amylase 1 precursor | IPI00315893 | 6 | 100 | 0.81 | 0.96 | 0.73 |
| 4 | Pancreatic lipase | IPI00420441 | 5 | 100 | 1.16 | 0.87 | 1.54 |
| 3 | Pancreatic lipase-related protein 2 | IPI00224868 | 4 | 100 | 1.00 | 0.86 | 0.63 |
| 5 | Pancreatic lipase-related protein 1 precursor | IPI00118355 | 4 | 100 | 1.06 | 0.92 | 0.91 |
| 8 | Elastase 3B, pancreatic | IPI00132043 | 3 | 100 | 1.21 | 1.02 | 1.14 |
| 9 | Chymotrypsinogen B | IPI00315696 | 3 | 100 | 1.76 | 1.14 | 1.78 |
| 13 | Bile salt-activated lipase precursor | IPI00132296 | 3 | 100 | 0.55 | 0.80 | 0.41 |
| 7 | Carboxypeptidase A1 precursor | IPI00356448 | 2 | 100 | 1.10 | 1.01 | 1.24 |
| 10 | Trypsinogen 9 | IPI00406464 | 2 | 100 | 1.08 | 0.81 | 1.24 |
| 11 | Glandular kallikrein K1 precursor | IPI00323869 | 2 | 100 | 1.43 | 0.93 | 1.04 |
| 16 | Syncollin | IPI00120271 | 2 | 100 | 0.83 | 0.88 | 0.64 |
| 18 | Carboxypeptidase B1 | IPI00351150 | 2 | 100 | 1.32 | 1.36 | 1.63 |
| 25 | Solh protein | IPI00461727 | 2 | 100 | 0.83 | 0.88 | 0.65 |
| 26 | BAP28 protein | IPI00411022 | 2 | 100 | 1.11 | 1.10 | 0.91 |

Protein identifications with >95% confidence and at least two unique peptides are reported. KO:KO = *Lman1* null:*Lman1* null, WT:KO = wild-type:*Lman1* null.

**Table 3-9: MS/MS analysis of pancreatic zymogen granules (>95% confidence).**

| Rank | Protein Name | Accession Number | Peptide Count | Total Ion Score C.I. % | 115:114 (WT:KO) | 116:114 (KO:KO) | 117:114 (WT:KO) |
|---|---|---|---|---|---|---|---|
| 1 | Amylase 2, pancreatic | IPI00756078 | 12 | 100 | 1.24 | 1.10 | 1.10 |
| 2 | Amylase 2, pancreatic | IPI00775996 | 12 | 100 | 1.25 | 1.07 | 1.13 |
| 3 | Pancreatic lipase-related protein 2 | IPI00224868 | 4 | 100 | 1.00 | 0.86 | 0.63 |
| 4 | Pancreatic lipase | IPI00420441 | 5 | 100 | 1.16 | 0.87 | 1.54 |
| 5 | Pancreatic lipase-related protein 1 precursor | IPI00118355 | 4 | 100 | 1.06 | 0.92 | 0.91 |
| 6 | Alpha-amylase 1 precursor | IPI00315893 | 6 | 100 | 0.81 | 0.96 | 0.73 |
| 7 | Carboxypeptidase A1 precursor | IPI00356448 | 2 | 100 | 1.10 | 1.01 | 1.24 |
| 8 | Elastase 3B, pancreatic | IPI00132043 | 3 | 100 | 1.21 | 1.02 | 1.14 |
| 9 | CHYMOTRYPSINOGEN B | IPI00315696 | 3 | 100 | 1.76 | 1.14 | 1.78 |
| 10 | Trypsinogen 9 | IPI00406464 | 2 | 100 | 1.08 | 0.81 | 1.24 |
| 11 | Glandular kallikrein K1 precursor | IPI00323869 | 2 | 100 | 1.43 | 0.93 | 1.04 |
| 12 | Trypsinogen 7 | IPI00131674 | 1 | 100 | 1.51 | 0.93 | 1.61 |
| 13 | Bile salt-activated lipase precursor | IPI00132296 | 3 | 100 | 0.55 | 0.80 | 0.41 |
| 14 | Lithostathine 1 precursor | IPI00117239 | 1 | 100 | 2.58 | 1.45 | 2.36 |
| 15 | Cathepsin B precursor | IPI00113517 | 1 | 100 | 1.75 | 1.13 | 1.51 |
| 16 | Syncollin | IPI00120271 | 2 | 100 | 0.83 | 0.88 | 0.64 |
| 17 | Anionic trypsin-2 precursor | IPI00403650 | 1 | 100 | 2.23 | 0.84 | 1.96 |
| 18 | Carboxypeptidase B1 | IPI00351150 | 2 | 100 | 1.32 | 1.36 | 1.63 |
| 19 | Elastase 1, pancreatic | IPI00171983 | 1 | 100 | 1.59 | 0.95 | 1.68 |
| 20 | Zymogen granule membrane protein 16 precursor | IPI00169543 | 1 | 100 | 1.62 | 1.30 | 1.51 |
| 21 | Serpin I2 precursor | IPI00122405 | 1 | 100 | 1.46 | 1.17 | 0.95 |
| 22 | Protein disulfide-isomerase precursor | IPI00133522 | 1 | 100 | 0.48 | 0.86 | 0.56 |
| 23 | Trypsinogen 10 | IPI00719897 | 1 | 100 | 1.67 | 0.79 | 1.86 |
| 24 | Chymotrypsin-like | IPI00318376 | 1 | 100 | 2.10 | 1.47 | 2.01 |
| 25 | Solh protein | IPI00461727 | 2 | 100 | 0.83 | 0.88 | 0.65 |
| 26 | BAP28 protein | IPI00411022 | 2 | 100 | 1.11 | 1.10 | 0.91 |
| 27 | Hypoxia up-regulated 1 | IPI00123342 | 1 | 100 | 0.54 | 0.61 | 0.51 |
| 28 | Polycystic kidney disease 1-like 2 | IPI00318010 | 1 | 100 | 0.76 | 1.05 | 0.77 |
| 29 | Cathepsin D | IPI00404551 | 1 | 100 | 0.85 | 0.75 | 0.84 |
| 30 | 16 kDa protein | IPI00751569 | 1 | 100 | 0.58 | 0.99 | 0.68 |
| 31 | Isoform 1 of Isocitrate dehydrogenase [NADP], mitochondrial precursor | IPI00798455 | 1 | 100 | 0.93 | 0.74 | 1.17 |
| 32 | Dynein heavy chain domain 3 isoform 2 | IPI00551194 | 1 | 100 | 2.80 | 0.99 | 2.05 |
| 34 | Isoform 3 of NK1 transcription factor-related protein 2 | IPI00229979 | 1 | 100 | 3.08 | 1.58 | 1.71 |
| 33 | Phospholipase A2 precursor | IPI00129638 | 1 | 100 | 1.58 | 1.15 | 1.05 |
| 35 | 37 kDa protein | IPI00625574 | 1 | 99 | 0.49 | 0.79 | 0.51 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 36 | Sphingosine-1-phosphate lyase 1 | IPI00320480 | 1 | 99 | 0.49 | 0.79 | 0.51 |
| 37 | Protocadherin beta 13 | IPI00129344 | 1 | 99 | 0.24 | 0.91 | 0.28 |
| 39 | TRANSCRIPTION FACTOR HES-2 (HAIRY AND ENHANCER OF SPLIT 2) | IPI00554874 | 1 | 98 | 0.94 | 0.90 | 0.66 |
| 40 | Isoform 2 of Uncharacterized protein C1orf26 homolog | IPI00453902 | 1 | 98 | 0.94 | 0.90 | 0.66 |
| 41 | Retinoblastoma-binding protein 1 isoform 5 | IPI00658848 | 1 | 98 | 0.73 | 0.82 | 0.61 |
| 42 | Isoform 2 of Otoferlin | IPI00223385 | 1 | 98 | 0.43 | 1.10 | 0.44 |
| 43 | Isoform 6 of Bullous pemphigoid antigen 1, isoforms 6/7 (Fragment) | IPI00277410 | 1 | 98 | 0.40 | 0.80 | 0.63 |
| 38 | 914 kDa protein | IPI00750701 | 1 | 98 | 1.06 | 1.13 | 0.78 |
| 44 | TPX2, microtubule-associated protein homolog | IPI00420481 | 1 | 98 | 1.20 | 1.38 | 1.08 |
| 45 | Similar to potassium channel tetramerisation domain containing 1 isoform 3 | IPI00752528 | 1 | 98 | 0.55 | 0.85 | 0.60 |
| 46 | Synaptonemal complex protein 2 | IPI00228891 | 1 | 97 | 1.58 | 1.57 | 1.47 |
| 47 | Cdc42 GTPase-activating protein | IPI00125505 | 1 | 97 | 0.97 | 1.25 | 1.00 |
| 48 | 215 kDa protein | IPI00776221 | 1 | 97 | 0.31 | 0.55 | 0.41 |
| 49 | Similar to jumonji domain containing 1B | IPI00754680 | 1 | 96 | 1.08 | 1.08 | 1.19 |
| 50 | Isoform 2 of Dual specificity testis-specific protein kinase 1 | IPI00316791 | 1 | 96 | 0.67 | 1.04 | 0.82 |
| 51 | Pyrroline-5-carboxylate reductase 2 | IPI00123278 | 1 | 95 | 0.52 | 0.93 | 0.56 |
| 52 | Novel protein | IPI00515169 | 1 | 95 | 1.15 | 0.91 | 0.89 |
| 53 | Membrane-bound transcription factor protease, site 1 | IPI00473288 | 1 | 95 | 1.27 | 1.16 | 1.45 |

Protein identifications with >95% confidence are reported. KO:KO = *Lman1* null:*Lman1* null, WT:KO = wild-type:*Lman1* null.

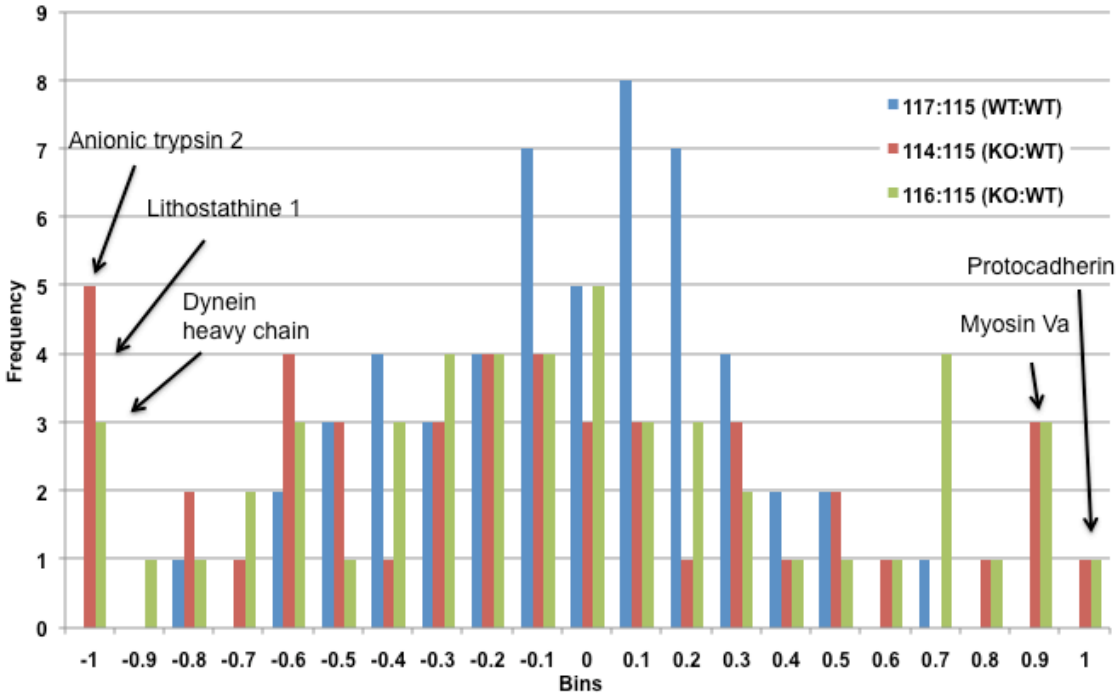


**Figure 3-13: iTRAQ ratio distribution for MS/MS analysis of zymogen granules**

Zymogen granules (ZG) from wild-type mice were labeled with 115 and 117; ZG from Lman1 null mice were labeled with 114 and 116. The 117:115 and 116:114 ratios represent duplicate comparisons for the wild-type and null samples, respectively. Cross comparisons (null:wild-type) are shown as the 114:115 and 116:115 ratios. The numbers on the x-axis represent a log transformation of raw iTRAQ ratios. Proteins that appear to be decreased in the null relative to wild-type are at the lower end of the histogram, whereas proteins that appear to be increased in the null relative to wild-type are at the upper end of the histogram. The iTRAQ ratios are shown for all protein identifications with >95% confidence. A few of the corresponding protein identifications are shown on the graph (see Table 3-9).

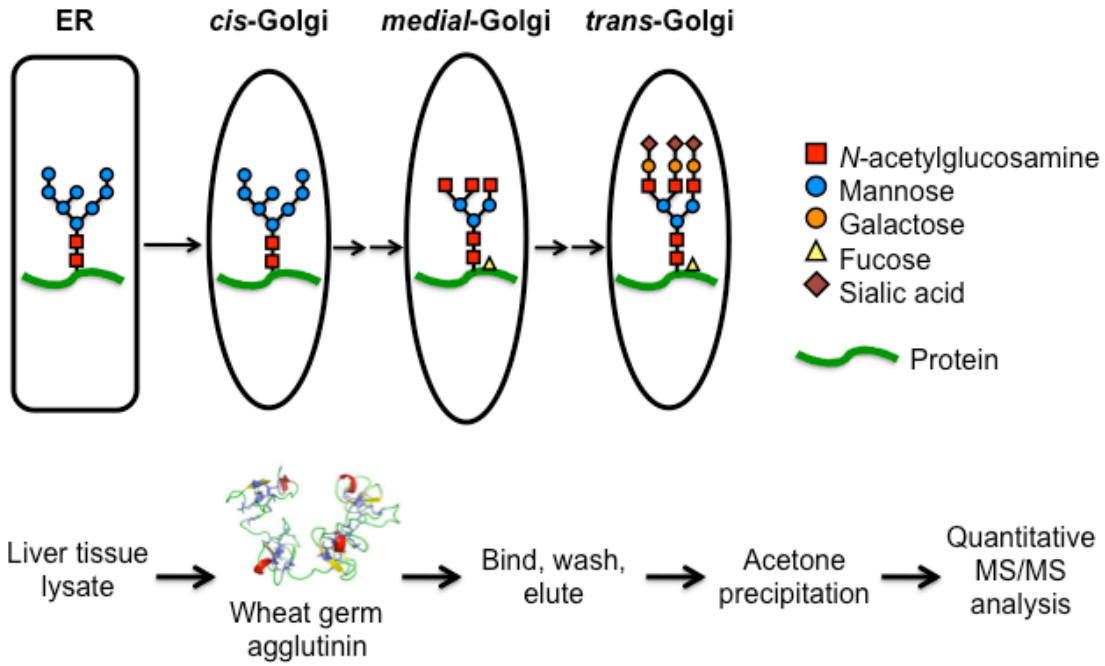


**Figure 3-14: Schematic of glycosylation patterns in the Golgi**

Newly synthesized glycoproteins exit the ER with a $(Man)_8(GlcNAc)_2$ glycan structure. In the Golgi, a series of processing steps (highly simplified in this drawing) trim or add new sugar residues to further elaborate the oligosaccharide chain. Glycans containing terminal *N*-acetylglucosamine or sialic acid are generated in the *medial*- and *trans*-Golgi. These terminal structures are recognized and bound by the lectin, wheat germ agglutinin (WGA). Therefore, WGA-affinity purification represents a method to select for proteins that have trafficked through the *medial*- and *trans*-Golgi.

**Table 3-10: iTRAQ-labeling of WGA-affinity purified liver tissue lysates**

| Mouse sample | Genotype | Analytical replicates | iTRAQ labels |
|---|---|---|---|
| 1 | *Lman1*+/+ A | | 113 |
| 2 | *Lman1*-/- A | | 114 |
| 3 | *Lman1*+/+ B | | 115 |
| 4 | *Lman1*-/- B | | 116 |
| | | | |
| 1 | | *Lman1*+/+ A | 117 |
| 2 | | *Lman1*-/- A | 118 |
| 3 | | *Lman1*+/+ B | 119 |
| 4 | | *Lman1*-/- B | 121 |

WGA-affinity purified liver tissue lysates from two individual wild-type mice were labeled with 113 and 115 (biologic replicates); WGA-affinity purified liver tissue lysates from *Lman1* null mice were labeled with 114 and 116 (biologic replicates). The remaining iTRAQ tags, 117, 118, 119 and 121, were added to duplicates of these four individual samples to serve as analytic replicates.

**Table 3-11: MS/MS analysis of WGA-affinity purified mouse liver tissue lysates.**

| Protein Identification (Name) | Accession | Peptide Count | 114: 113 | 115: 113 | 116: 113 | 117: 113 | 118: 113 | 119: 113 | 121: 113 |
|---|---|---|---|---|---|---|---|---|---|
| Carbamoyl-phosphate synthase, mitochondrial | IPI00111908.8 | 48 | 1.08 | 0.99 | 1.09 | 1.08 | 1.07 | 1.05 | 1.13 |
| Argininosuccinate synthase | IPI00134746.5 | 15 | 1.25 | 1.15 | 1.03 | 1.05 | 1.27 | 1.18 | 1.05 |
| ATP synthase subunit beta, mitochondrial | IPI00468481.2 | 14 | 1.11 | 1.04 | 1.04 | 1.07 | 1.13 | 1.07 | 1.06 |
| ATP synthase subunit alpha, mitochondrial | IPI00130280.1 | 13 | 1.04 | 1.01 | 0.99 | 1.05 | 1.05 | 1.10 | 1.02 |
| Hypoxia up-regulated protein 1 precursor | IPI00123342.4 | 12 | 1.03 | 1.04 | 0.93 | 1.01 | 1.10 | 1.05 | 0.90 |
| Endoplasmin precursor | IPI00129526.1 | 12 | 1.36 | 1.21 | 1.08 | 1.04 | 1.34 | 1.26 | 1.03 |
| 3-ketoacyl-CoA thiolase, mitochondrial | IPI00653158.1 | 12 | 1.10 | 1.15 | 0.96 | 1.02 | 1.12 | 1.22 | 0.98 |
| Aldehyde dehydrogenase, mitochondrial | IPI00111218.1 | 11 | 1.00 | 0.94 | 1.09 | 1.12 | 0.93 | 0.98 | 1.15 |
| Glutamate dehydrogenase 1, mitochondrial | IPI00114209.1 | 10 | 1.05 | 1.03 | 1.11 | 1.14 | 1.05 | 1.06 | 1.17 |
| ADP/ATP translocase 2 | IPI00127841.3 | 10 | 1.18 | 1.18 | 1.00 | 1.07 | 1.14 | 1.22 | 1.01 |
| UDP-glucuronosyltransferase 2 | IPI00112322.2 | 9 | 1.27 | 1.24 | 0.90 | 1.10 | 1.31 | 1.34 | 0.88 |
| Pyruvate carboxylase | IPI00114710.2 | 9 | 0.83 | 0.82 | 0.81 | 0.79 | 0.87 | 0.80 | 0.83 |
| 78 kDa glucose-regulated protein | IPI00319992.1 | 9 | 1.15 | 0.99 | 1.11 | 1.06 | 1.18 | 1.06 | 1.16 |
| 49 kDa protein | IPI00761732.1 | 9 | 1.04 | 1.00 | 1.07 | 1.07 | 1.05 | 1.07 | 1.13 |
| Murinoglobulin-1 | IPI00123223.2 | 8 | 1.40 | 1.00 | 1.32 | 1.05 | 1.38 | 0.94 | 1.25 |
| Macrophage mannose receptor 1 | IPI00126186.1 | 8 | 1.04 | 0.99 | 0.92 | 1.03 | 1.05 | 1.03 | 0.91 |
| Low-density lipoprotein receptor-related protein 1 | IPI00119063.2 | 7 | 0.86 | 0.84 | 0.80 | 0.97 | 0.83 | 0.84 | 0.84 |
| Dipeptidyl peptidase 4 | IPI00125813.1 | 7 | 1.04 | 0.98 | 0.88 | 1.05 | 1.05 | 1.06 | 0.87 |
| Betaine--homocysteine S-methyltransferase 1 | IPI00130950.1 | 7 | 0.78 | 0.80 | 1.07 | 1.17 | 0.75 | 0.79 | 1.12 |
| UDP glucuronosyltransferase 1 | IPI00480429.2 | 7 | 1.10 | 1.10 | 0.98 | 1.00 | 1.16 | 1.16 | 0.98 |
| UDP glucuronosyltransferase 2 | IPI00153143.1 | 6 | 1.01 | 1.01 | 0.91 | 1.07 | 1.06 | 1.03 | 0.89 |
| Heat shock cognate 71 kDa protein | IPI00480560.1 | 6 | 1.04 | 0.94 | 1.02 | 1.05 | 1.01 | 0.96 | 1.07 |
| Peroxiredoxin-1 | IPI00648615.3 | 6 | 0.75 | 0.72 | 1.04 | 1.14 | 0.76 | 0.76 | 1.13 |
| Bile salt export pump | IPI00828714.1 | 6 | 1.01 | 1.14 | 0.94 | 0.99 | 1.00 | 1.14 | 0.94 |
| Stress-70 protein, mitochondrial | IPI00133903.1 | 5 | 0.99 | 0.97 | 0.92 | 0.92 | 0.97 | 0.94 | 0.97 |
| Cytochrome P450 3A11 | IPI00134504.1 | 5 | 1.03 | 1.09 | 1.03 | 0.99 | 1.07 | 1.14 | 1.03 |
| Dihydroxyacetone kinase | IPI00310669.3 | 5 | 1.14 | 1.16 | 1.09 | 1.09 | 1.17 | 1.29 | 1.07 |
| Hydroxymethylglutaryl-CoA synthase, mitochondrial | IPI00420718.4 | 5 | 1.02 | 0.94 | 0.90 | 1.00 | 0.97 | 0.88 | 0.95 |
| Glycine N-methyltransferase | IPI00467066.7 | 5 | 0.81 | 0.77 | 1.25 | 1.02 | 0.82 | 0.71 | 1.27 |
| Fructose-bisphosphate aldolase B | IPI00127206.6 | 4 | 0.87 | 0.85 | 1.20 | 1.12 | 0.85 | 0.93 | 1.32 |
| Legumain | IPI00130627.1 | 4 | 0.90 | 0.96 | 0.93 | 0.99 | 0.92 | 1.07 | 0.91 |
| Thioredoxin domain-containing protein 4 | IPI00134058.3 | 4 | 0.78 | 0.92 | 0.74 | 0.98 | 0.82 | 1.00 | 0.86 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| UDP glucuronosyltransferase 2 | IPI00169666.3 | 4 | 0.96 | 1.05 | 0.92 | 1.09 | 1.08 | 1.08 | 0.93 |
| Isoform Long of Galectin-9 | IPI00230649.1 | 4 | 1.00 | 0.98 | 0.80 | 0.97 | 1.00 | 0.91 | 0.92 |
| Dolichyl-diphosphooligosaccharide | IPI00309035.2 | 4 | 0.96 | 0.88 | 0.92 | 0.99 | 0.96 | 0.94 | 0.95 |
| Sodium/potassium-transporting ATPase subunit alpha-1 | IPI00311682.5 | 4 | 1.11 | 1.14 | 1.08 | 1.01 | 1.11 | 1.16 | 1.03 |
| Solute carrier family 2 | IPI00311809.3 | 4 | 1.20 | 1.19 | 1.00 | 0.95 | 1.17 | 1.22 | 0.99 |
| Catalase | IPI00312058.5 | 4 | 0.92 | 0.93 | 1.18 | 1.08 | 0.99 | 1.04 | 1.21 |
| Flavin containing monooxygenase 5 | IPI00352124.2 | 4 | 1.11 | 1.07 | 1.17 | 1.06 | 1.08 | 1.09 | 1.29 |
| Tubulin alpha-1B chain | IPI00403810.2 | 4 | 1.15 | 1.10 | 1.00 | 1.05 | 1.12 | 1.05 | 1.08 |
| ATP-binding cassette sub-family G | IPI00468691.3 | 4 | 1.01 | 0.99 | 0.88 | 0.92 | 1.06 | 0.98 | 0.98 |
| Glutathione S-transferase P 1 | IPI00555023.2 | 4 | 0.77 | 0.72 | 1.23 | 1.22 | 0.77 | 0.78 | 1.36 |
| ATP synthase gamma chain, mitochondrial | IPI00750074.1 | 4 | 0.99 | 0.87 | 0.97 | 0.91 | 0.91 | 0.88 | 0.97 |
| Intergral membrane protein 1 | IPI00109108.2 | 3 | 0.92 | 0.93 | 0.85 | 0.95 | 0.89 | 0.98 | 0.89 |
| Cytochrome P450 2E1 | IPI00110556.1 | 3 | 0.99 | 1.01 | 1.03 | 1.03 | 1.06 | 1.07 | 1.00 |
| Abhydrolase domain-containing protein 14B | IPI00111876.2 | 3 | 0.97 | 0.99 | 1.06 | 0.92 | 0.96 | 0.93 | 1.03 |
| Arginase-1 | IPI00117914.3 | 3 | 0.96 | 0.88 | 1.17 | 1.08 | 1.03 | 0.90 | 1.18 |
| Cytochrome P450, family 2 | IPI00122634.1 | 3 | 0.96 | 1.03 | 0.98 | 1.01 | 1.08 | 0.99 | 0.97 |
| C-1-tetrahydrofolate synthase, cytoplasmic | IPI00122862.4 | 3 | 1.14 | 1.08 | 1.02 | 1.09 | 1.08 | 1.09 | 1.03 |
| NADH dehydrogenase | IPI00125929.2 | 3 | 1.04 | 0.95 | 0.92 | 0.97 | 1.05 | 0.95 | 0.90 |
| Cytochrome P450 1A2 | IPI00128287.1 | 3 | 0.93 | 0.93 | 0.82 | 0.99 | 0.92 | 0.96 | 0.88 |
| TFG protein | IPI00130013.1 | 3 | 1.29 | 1.19 | 1.11 | 0.98 | 1.26 | 1.35 | 1.05 |
| Prolyl 4-hydroxylase, beta | IPI00133522.1 | 3 | 1.00 | 0.81 | 1.17 | 1.15 | 0.95 | 0.91 | 1.40 |
| Calcium-binding mitochondrial carrier | IPI00135651.1 | 3 | 1.06 | 1.10 | 0.98 | 1.04 | 0.98 | 1.15 | 1.00 |
| 10-formyltetrahydrofolate dehydrogenase | IPI00153317.3 | 3 | 0.93 | 0.86 | 1.10 | 1.10 | 0.92 | 0.92 | 1.27 |
| Acetyl-CoA acetyltransferase, mitochondrial | IPI00154054.1 | 3 | 0.90 | 0.91 | 1.06 | 1.08 | 0.91 | 0.90 | 1.15 |
| Tubulin beta-2C chain | IPI00169463.1 | 3 | 1.14 | 0.93 | 0.87 | 0.89 | 1.03 | 0.99 | 0.91 |
| Long-chain-fatty-acid--CoA ligase 5 | IPI00170363.1 | 3 | 1.15 | 1.35 | 1.17 | 1.04 | 1.14 | 1.41 | 1.15 |
| Methyltransferase-like protein 7B | IPI00173179.3 | 3 | 1.12 | 1.11 | 0.96 | 1.01 | 1.09 | 1.16 | 0.98 |
| Alcohol dehydrogenase 1 | IPI00221400.5 | 3 | 0.95 | 1.02 | 1.14 | 1.16 | 0.97 | 1.01 | 1.18 |
| Uricase | IPI00223367.5 | 3 | 0.97 | 1.26 | 1.09 | 1.03 | 0.89 | 1.27 | 1.12 |
| Isoform 2 of Solute carrier organic anion transporter family | IPI00230319.1 | 3 | 1.02 | 1.12 | 0.81 | 0.95 | 1.06 | 1.14 | 0.73 |
| Selenium-binding protein 2 | IPI00323816.4 | 3 | 0.78 | 0.77 | 0.95 | 0.98 | 0.78 | 0.75 | 0.92 |
| Dolichyl-diphosphooligosaccharide | IPI00410819.3 | 3 | 1.05 | 0.96 | 1.01 | 1.13 | 1.10 | 1.00 | 1.00 |
| Phosphotyrosine protein phosphatase | IPI00474286.1 | 3 | 1.05 | 0.96 | 0.94 | 1.02 | 0.95 | 0.98 | 0.96 |
| Actin, gamma, cytoplasmic | IPI00652436.3 | 3 | 0.90 | 1.00 | 1.02 | 0.90 | 0.94 | 0.94 | 0.99 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Trypsinogen 10 | IPI00719897.1 | 3 | 0.80 | 0.93 | 0.82 | 0.74 | 0.98 | 0.68 | 1.43 |
| Putative integral membrane transport protein UST1R | IPI00754925.1 | 3 | 1.19 | 1.17 | 0.91 | 0.84 | 1.17 | 1.23 | 0.90 |
| Cytochrome P450 2D10 | IPI00762851.1 | 3 | 0.92 | 0.99 | 1.11 | 1.00 | 0.97 | 1.01 | 0.98 |
| Ubiquilin-2 | IPI00828670.2 | 3 | 1.26 | 1.16 | 0.94 | 1.09 | 1.37 | 1.18 | 1.00 |
| Isoform 1 of Dehydrogenase/reductase SDR family | IPI00110224.1 | 2 | 1.04 | 1.13 | 1.16 | 0.98 | 1.22 | 1.12 | 1.06 |
| Fatty acid synthase | IPI00113223.2 | 2 | 0.98 | 1.27 | 1.07 | 1.01 | 1.03 | 1.12 | 0.99 |
| Tubulin alpha-4A chain | IPI00117350.1 | 2 | 0.99 | 1.03 | 1.04 | 0.97 | 1.15 | 1.04 | 0.99 |
| Epididymis-specific alpha-mannosidase | IPI00117842.2 | 2 | 0.79 | 0.83 | 0.74 | 0.86 | 0.74 | 0.83 | 0.81 |
| NADH dehydrogenase 1 alpha subcomplex, mitochondrial | IPI00120212.1 | 2 | 1.28 | 1.06 | 1.11 | 0.91 | 1.19 | 1.22 | 1.14 |
| Fatty acid-binding protein, liver | IPI00120451.1 | 2 | 0.79 | 0.89 | 1.24 | 1.20 | 0.82 | 1.01 | 1.37 |
| Isoform 1 of NADH-cytochrome b5 reductase 3 | IPI00121079.3 | 2 | 0.94 | 0.92 | 0.98 | 1.12 | 0.99 | 0.95 | 0.93 |
| Glycogen [starch] synthase, liver | IPI00121962.3 | 2 | 1.24 | 1.32 | 0.90 | 1.00 | 1.37 | 1.06 | 0.93 |
| Mitochondrial antiviral-signaling protein | IPI00122075.1 | 2 | 1.28 | 1.16 | 1.13 | 1.06 | 1.22 | 1.17 | 1.14 |
| Hemopexin | IPI00128484.1 | 2 | 0.87 | 0.89 | 0.76 | 0.86 | 0.88 | 0.87 | 0.82 |
| Serine protease inhibitor A3K | IPI00131830.1 | 2 | 1.44 | 1.21 | 1.32 | 0.92 | 1.50 | 1.18 | 1.28 |
| Integrin beta-1 | IPI00132474.3 | 2 | 1.05 | 1.04 | 0.94 | 1.03 | 1.15 | 1.06 | 0.95 |
| Cytochrome P450 4A14 | IPI00133877.1 | 2 | 0.86 | 1.16 | 0.95 | 0.77 | 0.83 | 1.00 | 0.94 |
| Glutamyl aminopeptidase | IPI00134585.1 | 2 | 1.14 | 1.31 | 0.95 | 1.00 | 1.21 | 1.35 | 0.87 |
| Isoform 1 of ATP-binding cassette | IPI00170146.3 | 2 | 0.88 | 0.99 | 0.75 | 0.87 | 0.97 | 1.02 | 0.76 |
| Synaptic glycoprotein SC2 | IPI00262743.6 | 2 | 1.10 | 1.17 | 0.93 | 1.02 | 1.10 | 1.22 | 0.94 |
| Solute carrier family 25 | IPI00276926.2 | 2 | 0.95 | 0.97 | 0.91 | 0.95 | 0.95 | 1.12 | 0.87 |
| D-beta-hydroxybutyrate dehydrogenase, mitochondrial | IPI00330754.1 | 2 | 1.01 | 1.08 | 1.05 | 1.02 | 1.08 | 1.13 | 0.97 |
| Heat shock protein HSP 90-alpha | IPI00330804.4 | 2 | 0.98 | 0.90 | 1.03 | 0.84 | 0.98 | 0.89 | 0.98 |
| Glutathione S-transferase A3 | IPI00331241.8 | 2 | 0.86 | 0.89 | 1.18 | 1.19 | 1.02 | 0.97 | 1.31 |
| Microsomal glutathione S-transferase 1 | IPI00331322.3 | 2 | 1.06 | 1.19 | 1.07 | 1.30 | 1.07 | 1.15 | 1.13 |
| Protein-glutamine gamma-glutamyltransferase 2 | IPI00387333.4 | 2 | 0.81 | 0.80 | 0.89 | 0.92 | 0.85 | 0.82 | 0.91 |
| 2,4-dienoyl-CoA reductase, mitochondrial | IPI00387379.1 | 2 | 1.13 | 0.98 | 0.90 | 1.07 | 1.06 | 1.08 | 0.95 |
| Beta-glucuronidase | IPI00421209.2 | 2 | 1.09 | 1.01 | 1.02 | 1.05 | 1.17 | 1.16 | 0.95 |
| Probable N-acetyltransferase | IPI00459579.2 | 2 | 1.09 | 1.11 | 1.02 | 1.07 | 1.05 | 1.13 | 1.03 |
| Alpha-2-macroglobulin receptor-associated protein | IPI00469307.2 | 2 | 0.86 | 0.90 | 0.76 | 0.86 | 0.97 | 0.90 | 1.02 |
| 3-hydroxyacyl Coenzyme A dehydrogenase | IPI00554834.2 | 2 | 1.00 | 1.07 | 1.24 | 0.97 | 1.03 | 1.09 | 1.27 |
| Glutathione S-transferase, theta 1 | IPI00554933.2 | 2 | 1.14 | 1.09 | 1.07 | 1.08 | 1.18 | 1.24 | 1.09 |
| Hexose-6-phosphate dehydrogenase | IPI00624765.3 | 2 | 1.52 | 1.54 | 1.06 | 0.97 | 1.61 | 1.48 | 0.99 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Isoform 1 of Elongation factor Tu, mitochondrial | IPI00625588.1 | 2 | 0.83 | 0.95 | 0.96 | 0.90 | 0.92 | 0.92 | 0.91 |
| Solute carrier organic anion transporter family | IPI00625835.1 | 2 | 1.22 | 1.54 | 0.84 | 1.11 | 1.40 | 1.35 | 0.92 |
| Retinal dehydrogenase 1 | IPI00626662.3 | 2 | 0.97 | 0.82 | 1.09 | 1.14 | 0.89 | 0.95 | 1.28 |
| Glutamine synthetase | IPI00626790.2 | 2 | 0.65 | 0.61 | 0.86 | 0.83 | 0.72 | 0.71 | 0.90 |
| Glutathione S-transferase Mu 1 | IPI00649450.1 | 2 | 0.85 | 0.95 | 1.24 | 1.28 | 0.82 | 1.00 | 1.53 |
| Serine carboxypeptidase 1 | IPI00653972.1 | 2 | 0.93 | 1.23 | 0.95 | 1.03 | 1.09 | 1.24 | 0.92 |
| Similar to Es31 protein isoform 9 | IPI00664955.1 | 2 | 0.84 | 0.81 | 0.99 | 1.01 | 0.76 | 0.86 | 0.92 |
| Ribophorin II | IPI00754992.1 | 2 | 0.92 | 1.07 | 0.89 | 0.94 | 0.97 | 0.89 | 0.77 |
| Peroxiredoxin-6 | IPI00758024.1 | 2 | 0.81 | 0.78 | 1.15 | 1.09 | 0.82 | 0.77 | 1.11 |
| Acads Protein | IPI00776154.1 | 2 | 1.06 | 1.10 | 1.02 | 0.92 | 0.96 | 1.06 | 1.13 |
| Lysosomal alpha-mannosidase | IPI00785267.1 | 2 | 0.73 | 0.93 | 0.75 | 0.82 | 0.85 | 0.89 | 0.82 |
| UDP-glucuronosyltransferase 1-9 | IPI00785460.2 | 2 | 1.12 | 1.29 | 1.01 | 1.01 | 1.15 | 1.36 | 1.05 |
| Cytochrome P450 2A5 | IPI00798613.1 | 2 | 0.86 | 0.89 | 0.83 | 0.87 | 0.74 | 0.89 | 0.86 |
| Hydroxyacyl-Coenzyme A dehydrogenase type II | IPI00830581.1 | 2 | 1.21 | 1.13 | 1.09 | 1.13 | 1.12 | 1.14 | 1.14 |
| Electron transfer flavoprotein subunit beta | IPI00848492.1 | 2 | 0.93 | 0.97 | 1.24 | 1.07 | 0.89 | 0.94 | 1.41 |
| Hypothetical protein LOC219189 | IPI00850133.1 | 2 | 1.01 | 1.10 | 0.96 | 0.98 | 1.06 | 0.99 | 1.02 |

Protein identifications with >95% confidence and at least two unique peptides are reported. Columns shaded in gray represent ratios comparing *Lman1* null to wild-type samples (KO:WT).

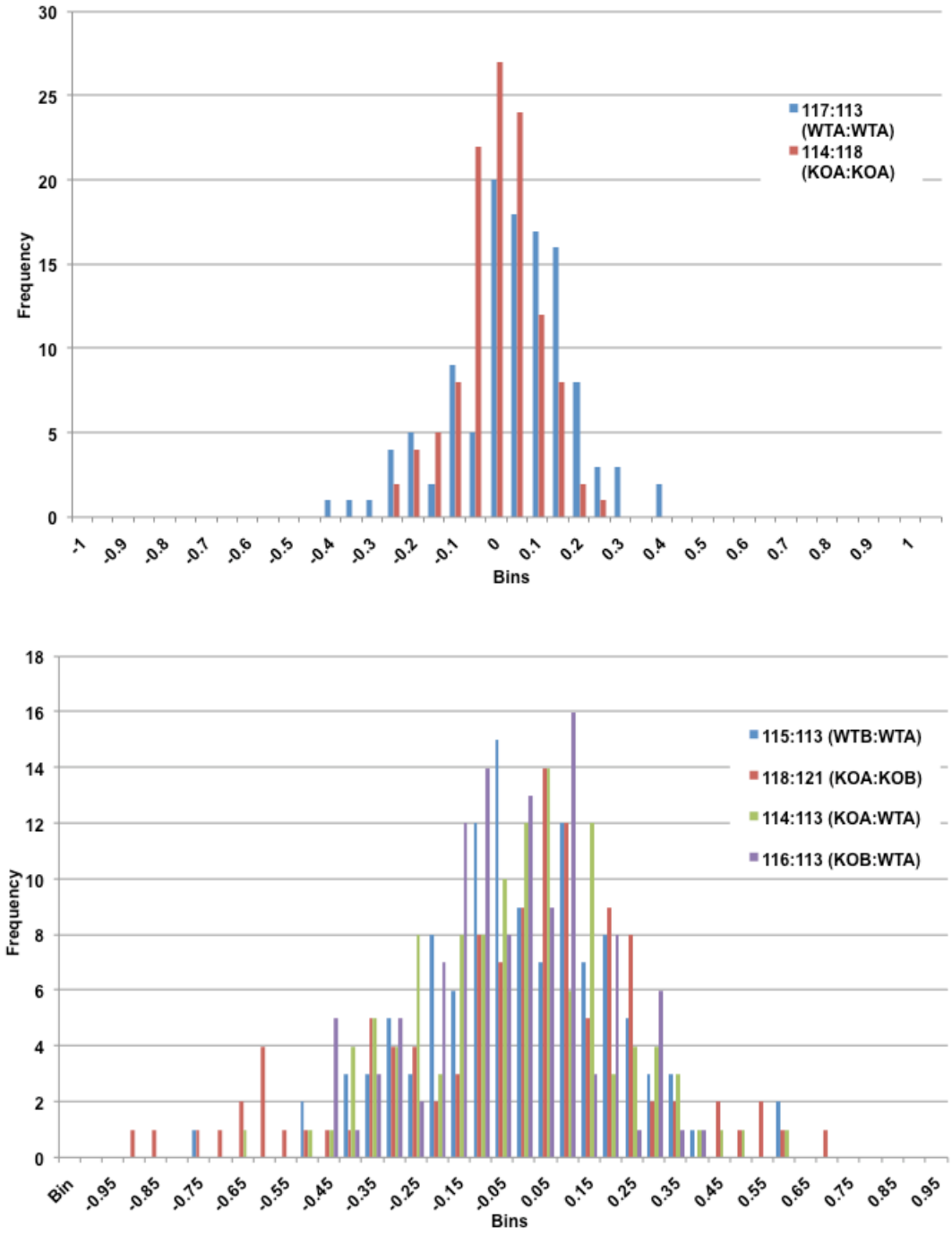


**Figure 3-15: iTRAQ ratio distribution for liver glycoprotein analysis**

iTRAQ ratio distributions (log transformed) are shown for control samples only (top) and all samples (bottom). Ratios are shown for all protein identifications with >95% confidence and at least 2 unique peptides (see Table 3-11).

## References


1. Nichols WC, Seligsohn U, Zivelin A, Terry VH et al. Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII. Cell. 1998;93:61-70.

2. Zhang B, Cunningham MA, Nichols WC, Bernat JA et al. Bleeding due to disruption of a cargo-specific ER-to-Golgi transport complex. Nat Genet. 2003;34:220-5.

3. Zhang B, McGee B, Yamaoka JS, Guglielmone H et al. Combined deficiency of factor V and factor VIII is due to mutations in either LMAN1 or MCFD2. Blood. 2006;107:1903-7.

4. Moussalli M, Pipe SW, Hauri HP, Nichols WC et al. Mannose-dependent endoplasmic reticulum (ER)-Golgi intermediate compartment-53-mediated ER to Golgi trafficking of coagulation factors V and VIII. J Biol Chem. 1999;274:32539-42.

5. Zhang B, Kaufman RJ, Ginsburg D. LMAN1 and MCFD2 form a cargo receptor complex and interact with coagulation factor VIII in the early secretory pathway. J Biol Chem. 2005 Jul 8;280(27):25881-6.

6. Zhang B, Spreafico M, Zheng C, Yang A et al. Genotype-phenotype correlation in combined deficiency of factor V and factor VIII. Blood. 2008;111:5592-600.

7. Nyfeler B, Michnick SW, Hauri HP. Capturing protein interactions in the secretory pathway of living cells. Proc Natl Acad Sci U S A. 2005;102:6350-5.

8. Nyfeler B, Zhang B, Ginsburg D, Kaufman RJ, Hauri HP. Cargo selectivity of the ERGIC-53/MCFD2 transport receptor complex. Traffic. 2006;7:1473-81.

9. Nyfeler B, Reiterer V, Wendeler MW, Stefan E et al. Identification of ERGIC-53 as an intracellular transport receptor of alpha1-antitrypsin. J Cell Biol. 2008;180:705-12.

10. Ong SE, Mann M. Mass spectrometry-based proteomics turns quantitative. Nat Chem Biol. 2005;1:252-62.

11. Ross PL, Huang YN, Marchese JN, Williamson B et al. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics. 2004;3:1154-69.

12. Stryke D, Kawamoto M, Huang CC, Johns SJ et al. BayGenomics: a resource of insertional mutations in mouse embryonic stem cells. Nucleic Acids Res. 2003;31:278-81.

13. Chen X, Walker AK, Strahler JR, Simon ES et al. Organellar proteomics: analysis of pancreatic zymogen granule membranes. Mol Cell Proteomics. 2006;5:306-12.

14. Linke T, Doraiswamy S, Harrison EH. Rat plasma proteomics: effects of abundant protein depletion on proteomic analysis. J Chromatogr B Analyt Technol Biomed Life Sci. 2007;849:273-81.

15. Hortin GL, Sviridov D. The dynamic range problem in the analysis of the plasma proteome. J Proteomics. 2009 Jul 17. [Epub ahead of print].

16. Pieper R, Su Q, Gatlin CL, Huang ST et al. Multi-component immunoaffinity subtraction chromatography: an innovative step towards a comprehensive survey of the human plasma proteome. Proteomics. 2003;3:422-32.

17. Wang H, Clouthier SG, Galchev V, Misek DE et al. Intact-protein-based high-resolution three-dimensional quantitative analysis system for proteome profiling of biological fluids. Mol Cell Proteomics. 2005;4:618-25.

18. Rindler MJ, Xu CF, Gumper I, Smith NN, Neubert TA. Proteomic analysis of pancreatic zymogen granules: identification of new granule proteins. J Proteome Res. 2007;6:2978-92.

19. Tartakoff AM, Vassalli P. Lectin-binding sites as markers of Golgi subcompartments: proximal-to-distal maturation of oligosaccharides. J Cell Biol. 1983;97:1243-8.

# CHAPTER IV:

# CONCLUSION AND FUTURE DIRECTIONS

Protein trafficking in the intracellular secretory pathway is carefully regulated to ensure the proper sorting and delivery of a diverse range of cargo proteins to many different cellular environments. As they progress through the early secretory pathway, newly synthesized proteins are exported from the endoplasmic reticulum (ER) in COPII vesicles and transported to the Golgi apparatus for further processing and sorting. The work described in this dissertation examines the role of specific components of COPII transport vesicles, including the major cargo-binding subunit of the COPII coat, SEC24, and a more specialized transport receptor, the LMAN1-MCFD2 complex. The existence of multiple isoforms of SEC24 (SEC24A-D) suggests that different types of COPII-coated vesicles, each with a particular set of cargo proteins, may function in a tissue-and cell-type specific manner. Specialized cargo receptors such as LMAN1-MCFD2 may further expand the cargo-binding repertoire of COPII vesicles by linking soluble cargo proteins to SEC24.

## Essential role for SEC24D in early embryonic development

In the first study, we demonstrated a critical role for SEC24D in mammalian development. *Sec24d*-deficient mice were generated from two independent gene trap ES-cell clones, designated Line 1 and Line 2. In both gene trap lines, we observed a complete absence of *Sec24d* null mice when the offspring of F1 intercrosses were

genotyped at two to three weeks of age. Analysis of embryos at the blastocyst stage showed a complete deficiency of *Sec24d-/-* embryos in Line 1, indicating a requirement for SEC24D in very early development. However, in Line 2, *Sec24d-/-* blastocysts were present in expected numbers. This discrepancy is possibly due to a low level of normal splicing around the gene trap vector or residual function of the SEC24D/β-geo fusion protein in Line 2, as discussed below. A statistically significant excess of *Sec24d* heterozygous mice was observed in Line 1 (discussed in more detail in Chapter II). Although an explanation for this phenomenon is not yet clear, the failure of Line 2 to recapitulate this finding suggests that it may be an artifact unique to Line 1, independent of *Sec24d*.

**Determining the time point of *Sec24d* null embryonic lethality**

In Line 1, *Sec24d* homozygous null embryos were absent as early as the blastocyst stage (embryonic day 3.5), indicating a critical role for SEC24D in very early stages of mammalian development. Although we observed one 8-cell stage embryo that was *Sec24d-/-* on repeat genotyping, this result might not be reliable due to the very low copy number of the DNA template used for PCR genotyping. It remains plausible that SEC24D is absolutely required for survival at the level of a single cell; if *Sec24d* null embryos are present at the 8-cell stage, their survival may be permitted by maternal *Sec24d* transcripts and protein, which are still present during the earliest rounds of cell division [1].

As noted, there was discrepancy between Line 1 and Line 2 with respect to the time point at which *Sec24d* null embryos disappear. While no *Sec24d-/-* blastocysts were observed in Line 1, they were present in expected numbers in Line 2. In a preliminary

analysis of one litter of Line 2 embryos at E10.5, we observed 3 *Sec24d*+/+ and 5 *Sec24d*+/- embryos, but no *Sec24d*-/- embryos. However, we would have to genotype at least 11 embryos and observe no *Sec24d*-/- embryos to reach a p-value < 0.05. Timed matings to harvest additional litters of Line 2 embryos at E10.5 are currently underway. As discussed in Chapter II, two possible mechanisms could explain the persistence of *Sec24d* null embryos to the blastocyst stage in Line 2. Although the gene trap vectors used to generate these ES-cell clones is designed to trap and inactivate the gene by splicing to the upstream exon [2], pre-mRNA splicing may occur around the gene trap vector, resulting in a hypomorphic allele that produces a low level of wild-type transcript [3, 4]. Splicing around the gene trap vector could produce enough SEC24D to promote survival to the blastocyst stage, but not enough to sustain later stages of development. Alternatively, the SEC24D/β-geo fusion protein could represent a partial loss-of-function allele that retains some normal SEC24D function. Because the gene trap vector insertion site is located much farther downstream in Line 2 compared to Line 1 (intron 20, as opposed to intron 8 in Line 1), the predicted fusion transcript in Line 2 contains a greater proportion of *Sec24d* coding sequence. In Line 1, the gene trap allele creates a fusion protein containing only the N-terminal 347 amino acids of SEC24D (the truncation occurs prior to the SEC23/SEC24 zinc finger domain, leaving only the N-terminal hypervariable region intact), whereas in Line 2, the fusion protein contains the first 892 out of 1032 amino acids (truncated just before the gelsolin repeat domain). To test the first hypothesis, RNA could be isolated from Line 2 blastocysts and used to perform RT-PCR with primers complementary to sequences located in the exons immediately 5' (exon 20) and 3' (exon 21) to the gene trap vector insertion site. To determine the lower

limit of sensitivity for detecting a small amount of wild-type transcript, known amounts of wild-type RNA can be serially diluted into the RNA isolated from *Sec24d-/-* blastocysts. These studies could be complemented by a single nucleotide polymorphism (SNP)-based analysis of *Sec24d* expression in heterozygous mice. Because the gene trap allele originated in ES-cells derived from the 129P2 strain of mice, and our mice have been maintained by backcrossing to the C57BL6/J strain, SNPs between these two strains located in the coding region of *Sec24d* could allow us to distinguish between transcripts produced from the wild-type and gene trap alleles via a quantitative primer extension assay. The second partial loss-of-function hypothesis could be explored by generating constructs corresponding to the truncated SEC24D proteins encoded by the gene trap fusion transcripts in Line 1 and Line 2, and testing whether they have any residual activity in a functional assay for SEC24, such as an in vitro budding assay, or a cell culture-based assay to see whether the fusion proteins can be recruited to ER exit sites in the presence of constitutively active SAR1.

**Normal appearance of Sec24d heterozygous mice**

Sec24d heterozygous mice from both lines were viable, fertile, and did not exhibit any obvious phenotypic abnormalities. Histological analysis of a wide range of tissue samples also failed to identify any differences between *Sec24d+/+* and *Sec24d+/-* mice. Because of the critical role of SEC24 in the early secretory pathway, we decided to focus our analysis on highly secretory tissues, such as the liver and pancreas. Disruption of COPII export is expected to lead to a build-up of cargo proteins, which may result in dilation of the endoplasmic reticulum that can be visualized at the ultrastructural level. For example, in the human disorder, CLSD, which results from a missense mutation in

*SEC23A,* ER dilation is observed in cultured skin fibroblasts from both homozygous affected individuals and phenotypically normal heterozygous carriers [5]. Our electron microscopic analysis of liver and pancreatic tissue, as well as mouse embryonic fibroblasts derived from *Sec24d*+/- embryos failed to demonstrate consistent ER dilation; however, it would be worthwhile to repeat this analysis in a blinded fashion with a more objective scoring method to determine whether subtle defects are present. Similarly, in preliminary experiments using siRNAs and shRNAs to knockdown expression of SEC24D in HeLa cells (data not shown), we did not observe any alterations in the pattern or distribution of staining for markers of the ER, ERGIC and Golgi compartments. However, even stable expression of *SEC24D* shRNAs did not result in a complete knockdown of *SEC24D* (~60-80% knockdown), and it remains possible that a more substantial reduction in the level of SEC24D is required to elicit a cellular phenotype in this setting. Although *Sec24d* heterozygous mice do not exhibit any gross abnormalities under normal circumstances, further analysis might identify a less obvious defect, perhaps in a tissue not yet studied in detail. It would be interesting to determine if the basal levels of ER stress markers are elevated in *Sec24d* heterozygous mice, or whether additional challenges (i.e. high fat diet, tunicamycin injection) might elicit an identifiable phenotype.

**Generation of mouse models to study the role of SEC24D beyond the blastocyst stage**

The early embryonic lethality of *Sec24d* null embryos precludes the study of SEC24D function later in development or in the adult mouse. Several approaches could allow us to generate genetically engineered mice that would permit studies of the *in vivo* function of SEC24D later in development or adulthood. Generation of a conditional

'knockout' allele of *Sec24d* by conventional gene targeting and homologous recombination of a *floxed* allele would allow *cre*-mediated excision and deletion of *Sec24d* at precise developmental time points or in specific tissue lineages. Mice such as these would allow us to study the effect of loss of SEC24D at later stages of development or in defined tissues and cell-types. Engineering of a conditional null allele of *Sec24d* is currently underway as part of the Knockout Mouse Project (www.knockoutmouse.org), which aims to provide researchers with targeted ES-cell lines for the purpose of generating knockout mouse models [6]. In the meantime, we have begun efforts to create a mouse model carrying a bacterial artificial chromosome (BAC) transgene that rescues *Sec24d* embryonic lethality. BAC transgenes are advantageous because they often recapitulate the endogenous levels, timing, and pattern of gene expression better than minigene-based transgenes [7]. We identified three BAC clones from a C57BL6/J library that span the entire *Sec24d* locus and contain considerable amounts of genomic sequence 5' and 3' to *Sec24d* (RP23-355K12, RP24-271N21 and RP23-381M6) [8]. These clones were acquired (bacpac.chori.org) and BAC DNA was purified and submitted to the University of Michigan Transgenic Core. BAC clones RP23-355K12 and RP24-271N21 were injected into fertilized oocytes and BAC-positive transgenic founders were obtained for both BAC clones. Matings between *Sec24d* heterozygous mice and BAC-positive *Sec24d* heterozygous mice are currently in progress to determine whether either *Sec24d*–containing BAC is able to rescue the embryonic lethality of *Sec24d-/-* mice. Successful rescue will also provide unequivocal confirmation that the observed embryonic lethality in our *Sec24d* null mice from Line 1 is due to the absence of SEC24D.

**Uncovering SEC24 isoform-specific functions**

The cause of early embryonic lethality in *Sec24d* null mice is unknown. An absence of SEC24D may lead to a general block in the export of cargo proteins from the ER, or it may lead to a more cargo-specific defect in protein sorting or secretion. Our results suggest that the other isoforms of SEC24 are unable to compensate for loss of SEC24D in this context. However, it remains to be determined whether the observed embryonic lethal phenotype is a manifestation of SEC24D-specific function or related to differential expression of the various isoforms. Although we detected expression of all four isoforms in total RNA isolated from E10.5 embryos, the developmental timing and expression pattern of the SEC24 proteins prior to that stage is not known. Our initial studies of adult mouse tissues provide evidence that the expression patterns of the four SEC24 isoforms are at least partially overlapping; however, higher-resolution studies are required to determine if there is diversification of SEC24 isoform expression at the cell-type specific level within a given tissue. Further analysis by western blot and immunohistochemistry of paraffin-embedded tissue sections using the anti-peptide antibodies recently generated by our laboratory should provide a more detailed description of isoform-specific expression patterns.

To address the question of whether there is functional redundancy among the four isoforms, we plan to use BAC recombineering to determine whether different isoforms of SEC24, when expressed under control of the *Sec24d* promoter, can rescue the embryonic lethality of *Sec24d* null mice. As discussed above, experiments to determine whether a *Sec24d* BAC transgene can rescue the *Sec24d-/-* mice from Line 1 and recapitulate the endogenous *Sec24d* expression pattern are in progress. Once we have confirmed that at

least one of the two BACs achieves this result, BAC recombineering can be used to engineer *Sec24d* BACs containing cDNA minigenes encoding each of the other three isoforms. Given the higher level of sequence homology between SEC24C and SEC24D, it makes sense to first attempt this with a S*ec24c* minigene. If expression of SEC24C in this manner were sufficient to rescue the embryonic lethality of *Sec24d* null mice, this would provide strong evidence for functional overlap between SEC24C and SEC24D.

**Identifying novel cargo proteins for the LMAN1-MCFD2 transport receptor**

In the second study, we carried out an initial quantitative mass spectrometry-based proteomic analysis of a mouse model of the human bleeding disorder, combined deficiency of factor V and factor VIII (F5F8D). Although our laboratory has generated *Mcfd2* null mice that are now available for similar analyses, all of these studies were performed on samples isolated from *Lman1* null mice. We examined plasma, pancreatic zymogen granules, and lectin affinity-purified liver glycoproteins from *Lman1*+/+ and *Lman1*-/- mice in a quantitative manner using the iTRAQ method of labeling tryptic peptides with isobaric tags [9].

First, we performed three separate analyses on plasma samples from wild-type and *Lman1* null mice that were pooled from four mice of each genotype (*Lman1*+/+ and *Lman1*-/-). Because of the complexity of the plasma proteome and the fact that a small number of highly abundant proteins account for 99.9% of the total plasma protein mass, a key step in our sample preparation was the depletion of the three most abundant plasma proteins (albumin, IgG and transferrin) with the use of a monoclonal antibody-based immunodepletion column. However, given the fact that this column only depletes the top three most abundant proteins, many predominant plasma proteins were still present in

very high concentrations, likely masking our ability to detect the full range of lower abundance proteins. Currently, more resources exist for proteomic analysis of human plasma, and immunodepletion columns that deplete the top twenty most abundant plasma proteins are available. Whereas albumin is present at a concentration of approximately 40mg/ml, factor V (FV) and factor VIII (FVIII) have plasma concentrations of 10μg/ml and 100ng/ml, respectively. Because FV and FVIII are known to have reduced levels in plasma from *Lman1*-/- mice relative to *Lman1*+/+ controls (Bin Zhang, manuscript in preparation), they could potentially serve as internal controls for iTRAQ quantitation. Despite the dynamic range problem presented by the high order-of-magnitude differences in plasma protein abundance [10], we did identify FV and FVIII in our analysis; however, each was identified only once, and in separate analyses. Second, there were inconsistencies between the iTRAQ ratios for highly confident protein identifications from one MS/MS analysis to the next (such as the example of ApoAI, discussed in Chapter III). α-1-antitrypsin is another highly abundant plasma protein that has been proposed to be a cargo protein for LMAN1. Although we did identify α-1-antitrypsin in our analyses, there was no significant difference in the iTRAQ ratios comparing *Lman1*-/- to *Lman1*+/+ mice. Finally, some of the most significant 'hits' in our studies corresponded to intracellular (non-secreted) proteins. While the failure to make consistent protein identifications is disappointing, it is not entirely unexpected. Highly complex protein mixtures such as plasma remain a formidable challenge for mass spectrometry-based proteomics [11].

The difficulties we encountered in our analysis of mouse plasma led us to focus on more highly purified protein fractions in an effort to reduce sample complexity. The

analysis of pancreatic zymogen granules (ZG) was appealing for several reasons: (1) *Lman1* and *Mcfd2* are both expressed in the pancreas, (2) the pancreas is highly active in protein secretion, (3) our collaborators had already demonstrated success using iTRAQ labeling and MS/MS analysis to identify pancreatic ZG membrane proteins [12], and (4) the major ZG content proteins have been well-documented [13]. In pilot studies, we showed that a sufficient amount of protein could be obtained from ZG isolated from the pancreas of a single mouse. In a 4-plex iTRAQ experiment, this made it possible to compare the ZG proteomes of biological replicates (two individual mice of each genotype). Our preliminary analysis of ZG purified from *Lman1*+/+ and *Lman1*-/- mice yielded 16 highly confident protein identification (at least two unique peptides and greater than 95% confidence). Of these 16 proteins, 14 were known zymogen granule content proteins. While this result was encouraging, the number of protein identifications was surprisingly small, likely due to the confounding factor of non-specific peptide cleavage by endogenous zymogen granule proteases that were not effectively inhibited by treatment with 6M urea and 0.5% SDS. Future analyses of ZG content proteins should be carried out under stronger denaturing conditions to properly inhibit endogenous protease activity.

Based on known patterns of glycosylation in the early secretory pathway, we reasoned that enrichment for certain classes of glycoproteins might provide a method to separate out proteins that have trafficked through the Golgi apparatus from those that are unable to exit the ER. The liver is a highly active secretory tissue that expresses *Lman1* and *Mcfd2*, and it is a site of synthesis for FV and FVIII. In addition, large amounts of liver tissue can easily be obtained for protein purification. Whole-cell liver lysates were

prepared from *Lman1*+/+ and *Lman1*-/- mice and subjected to lectin-affinity chromatography, using wheat-germ agglutinin, a lectin that specifically binds complex glycans that are added in the *medial* and *trans*-Golgi. The newly available 8-plex iTRAQ kit allowed us to label lectin-affinity purified liver glycoproteins from four individual mice (two *Lman1*+/+ and two *Lman1*-/-) in duplicate. In our initial MS/MS analysis of these samples, there were no consistent differences in the iTRAQ ratios comparing wild-type and *Lman1* null mice.

Although we cannot reliably interpret any individual data points from our proteomic experiments, our histogram plots suggest that there is true biological variation in protein abundance between *Lman1*+/+ and *Lman1*-/- mice. In the future, any proteins that appear to be consistently up- or down-regulated in multiple experiments would warrant further validation as possible novel cargo proteins for the LMAN1-MCFD2 complex.

**New approaches for proteomics-based analysis of defects in protein sorting and secretion**

There is growing interest in utilizing proteomics to address important questions in cell biology [14]. Efforts are underway to characterize the proteomes of a wide range of cellular organelles [15], and some have even focused on components of the early secretory pathway [16]. A proteomics approach allows an unbiased global comparison of differences in protein populations, and remains a promising method by which to identify novel cargo proteins for the LMAN1-MCFD2 pathway.

New methods are also being developed for quantitative proteomic analysis. The stable isotope labeling by amino acids in cell culture (SILAC) method has been in use for a number of years to differentially label cells grown in culture [19], and this method has

recently been employed to metabolically label whole animals with stable non-radioactive isotopes [20]. By feeding mice a special diet containing either the natural $^{12}C_6$-lysine or lysine substituted with a non-radioactive stable isotope ($^{13}C_6$-lysine) it is possible to differentially label samples for quantitative MS/MS analysis. We are in the process of generating a set of $^{13}C_6$-lysine C57BL6/J mice in our laboratory that will serve as the wild-type controls for quantitative MS/MS comparisons of samples from *Lman1* null and *Mcfd2* null mice. In the future, we plan to extend our quantitative proteomic analyses to other mouse models generated in our laboratory, such as those deficient in components of COPII vesicles.

**Final remarks: the connection between protein trafficking in the early secretory pathway and human disease**

In summary, this dissertation has focused on two important aspects of protein transport in the early secretory pathway. Future investigations will yield important insight into the contributions of the various isoforms of SEC24, and specialized cargo adaptors like LMAN1-MCFD2, to the process of selective protein transport. The identification of additional cargo proteins whose transport is dependent upon the LMAN1-MCFD2 pathway could have important implications for patients with F5F8D. In addition, with the recent discovery that mutations in specific components of COPII vesicles underlie several human disorders, our findings could have an even larger impact on the study of human diseases related to defects in protein transport.

**References**


1. Minami N, Suzuki T, Tsukamoto S. Zygotic gene activation and maternal factors in mammals. J Reprod Dev. 2007;53:707-15.

2. Stryke D, Kawamoto M, Huang CC, Johns SJ et al. BayGenomics: a resource of insertional mutations in mouse embryonic stem cells. Nucleic Acids Res. 2003;31:278-81.

3. Stanford WL, Cohn JB, Cordes SP. Gene-trap mutagenesis: past, present and beyond. Nat Rev Genet. 2001;2:756-68.

4. McClive P, Pall G, Newton K, Lee M et al. Gene trap integrations expressed in the developing heart: insertion site affects splicing of the PT1-ATG vector. Dev Dyn. 1998;212:267-76.

5. Boyadjiev SA, Fromme JC, Ben J, Chong SS et al. Cranio-lenticulo-sutural dysplasia is caused by a SEC23A mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking. Nat Genet. 2006;38:1192-7.

6. The International Mouse Knockout Consortium. A mouse for all reasons. Cell. 2007;128:9-13.

7. Van Keuren ML, Gavrilina GB, Filipiak WE, Zeidler MG et al. Generating transgenic mice from bacterial artificial chromosomes: transgenesis efficiency, integration and expression outcomes. Transgenic Res. 2009 Apr 26. [Epub ahead of print].

8. Osoegawa K, Tateno M, Woon PY, Frengen E et al. Bacterial artificial chromosome libraries for mouse sequencing and functional analysis. Genome Res. 2000;10:116-28.

9. Ross PL, Huang YN, Marchese JN, Williamson B et al. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics. 2004;3:1154-69.

10. Hortin GL, Sviridov D. The dynamic range problem in the analysis of the plasma proteome. J Proteomics. 2009 Jul 17. [Epub ahead of print].

11. Bell AW, Deutsch EW, Au CE, Kearney RE et al. A HUPO test sample study reveals common problems in mass spectrometry-based proteomics. Nat Methods. 2009 May 17. [Epub ahead of print].

12. Chen X, Walker AK, Strahler JR, Simon ES et al. Organellar proteomics: analysis of pancreatic zymogen granule membranes. Mol Cell Proteomics. 2006;5:306-12.

13. Rindler MJ, Xu CF, Gumper I, Smith NN, Neubert TA. Proteomic analysis of pancreatic zymogen granules: identification of new granule proteins. J Proteome Res. 2007;6:2978-92.

14. Bell AW, Nilsson T, Kearney RE, Bergeron JJ. The protein microscope: incorporating mass spectrometry into cell biology. Nat Methods. 2007;4:783-4.

15. Au CE, Bell AW, Gilchrist A, Hiding J et al. Organellar proteomics to create the cell map. Curr Opin Cell Biol. 2007;19:376-85.

16. Gilchrist A, Au CE, Hiding J, Bell AW et al. Quantitative proteomics analysis of the secretory pathway. Cell. 2006 Dec 15;127(6):1265-81.

17. Mallick P, Schirle M, Chen SS, Flory MR et al. Computational prediction of proteotypic peptides for quantitative proteomics. Nat Biotechnol. 2007;25:125-31.

18. Bergeron JJ, Hallett M. Peptides you can count on. Nat Biotechnol. 2007;25:61-2.

19. Ong SE, Mann M. Stable isotope labeling by amino acids in cell culture for quantitative proteomics. Methods Mol Biol. 2007;359:37-52.

20. Krüger M, Moser M, Ussar S, Thievessen I et al. SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. Cell. 2008;134:353-64.