Characterization of the function of the LMAN1 ER cargo receptor *in vivo*

by

Lesley Ann Everett

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Human Genetics)
in The University of Michigan
2014

Doctoral Committee:

Professor David Ginsburg, Chair
Assistant Professor Anthony Antonellis
Professor Jeff Innis
Professor Miriam Meisler
Professor Lois Weisman

DEDICATION

This thesis is dedicated to Jennifer Duncan-Taylor, whose infectious enthusiasm for genetics and biology inspired me to learn more.

ACKNOWLEDGEMENTS

Thank you to my Dissertation Committee (David Ginsburg, Miriam Meisler, Tony Antonellis, Jeff Innis, and Lois Weisman) for your time, your thoughtful and encouraging support, your intellectual input, and for the many helpful discussion we have shared. A special thanks to David Ginsburg, my mentor, whose guidance and mentorship proved invaluable to my dissertation research and to my personal scientific development.

Thank you to the University of Michigan Medical Scientist Training Program (Ron Koenig, Ellen Elkin, Hilkka Ketola, and Laurie Koivupalo) for your tremendous support and guidance, and for all of the work you do to constantly grow and develop our MSTP program. A special thanks to Ron and Ellen for the opportunity to serve on the MSTP Program Activities Committee and to help recruit new fellows to our program. Thank you to the Department of Human Genetics faculty and staff for your administrative and scientific support, particularly to Karen Grahl, JoAnn Sekiguchi, Didi Robins, and Sue Kellog. Thank you to Bev Yashar and Tom Gelehrter for the opportunity to serve as a teaching assistant for the human genetics component of the medical school Patients and Populations sequence, a highly enjoyable peer-teaching experience. In addition, thank you to Dr. Gelehrter for the many insightful and thought-provoking discussions we have enjoyed.

Thank you to the National Institutes of Health, the National Science Foundation, and the University of Michigan Horace B. Rackham Graduate School for funding portions of my training. Thanks to the University of Michigan Basic Science Core Facilities (Transgenic Animals, Histology, Microscopy and Image-Analysis, Proteomics, PCAR, etc.) for making it possible to do and learn incredible science at this institution.

Finally, my most grateful thanks to the friends, family members, MSTP students, and lab mates who have encouraged and inspired me to pursue an MD-PhD. Mom, Andrea, Rami, Cristina, Kirstin, John, Sarah, Michele, and Norbert – thank you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF SUPPLEMENTARY TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTARY FIGURES

LIST OF ABBREVIATIONS

A1AT	Alpha-1 antitrypsin
B6;129	C57BL/6J, 129S1/SvImJ mixed genetic background
CBC	Complete blood count
COPII	Coat protein complex II
CRD	Carbohydrate recognition domain
CTSC	Cathepsin C
CTSZ	Cathepins Z
ER	Endoplasmic reticulum
ERGIC	ER-Golgi Intermediate compartment
ERGIC-53	ER-Golgi Intermediate compartment 53 kDa protein (LMAN1)
F5	Gene encoding coagulation factor V
F8	Gene encoding coagulation factor VIII
F5F8D	Combined deficiency of coagulation factors V and VIII
FV	Coagulation factor V (protein)
FVIII	Coagulation factor VIII (protein)
HA	Hemagglutinin protein tag
IVC	Inferior vena cava
KOMP	Knockout mouse project
LMAN1	Lectin, mannose-binding 1 (ERGIC-53)
LSEC	Liver sinusoidal endothelial cells
MCFD2	Multiple coagulation factor deficiency 2
PT	Prothrombin time
PTT	Partial thromboplastin time
SILAC	Stable isotope labeling by amino acids in cell culture
TEM	Transmission electron microscopy
VWD	von Willebrand disease
VWF	von Willebrand Factor
Y2H	Yeast two-hybrid

ABSTRACT

Combined deficiency of coagulation factors V and VIII (F5F8D) is a human autosomal recessive bleeding disorder caused by defects in LMAN1 or MCFD2. These proteins form a complex, which facilitates secretion of coagulation factors V (FV) and VIII (FVIII), serving as the only known cargo-receptor in the mammalian secretory pathway. F5F8D patients exhibit FV and FVIII levels that are ~10-15% of normal. This dissertation focuses on three major questions regarding F5F8D pathophysiology. First, to answer a long-standing question in the hemostasis field, what is the cellular source of FVIII biosynthesis? Second, what is the relationship between *Lman1* expression level and the corresponding secretion of FV and FVIII? Third, can additional LMAN1-dependent cargos be identified?

Although hepatocytes and megakaryocytes are the known biosynthetic sources of murine FV, the primary cellular source of FVIII biosynthesis is controversial, with contradictory evidence supporting an endothelial or hepatocyte origin. Conditional knockout mice with *Lman1* deleted in the endothelium or the hepatocytes were generated in order to determine the relative contribution of these two tissues to the plasma FVIII pool. The endothelium-knockout mice exhibit normal FV levels, but significantly reduced FVIII levels. Conversely, hepatocyte-specific knockout mice exhibit normal FVIII levels, but low FV levels. These results indicate that endothelial cells are the primary site of FVIII biosynthesis, and that hepatocytes do not significantly contribute to the plasma FVIII pool. qRT-PCR analysis of endothelial cell-specific mRNA confirmed the endothelial-specific expression of *F8*, and demonstrated extensive heterogeneity of *F8* expression in different vascular beds, with the highest expression in the endothelial cells of liver and kidney.

A series of LMAN1-deficient mice with variable *Lman1* expression levels (0% to ~7% of normal) was generated. LMAN1-null mice exhibit FV and FVIII levels that are ~50% of normal (versus ~10-15% in human patients), while mice expressing ~7% of

wild-type *Lman1* levels show an intermediate reduction in FV and FVIII (~70% of wild-type levels). A similar result was demonstrated for alpha-1 antitrypsin secretion, another LMAN1-dependent cargo. Given that nearly all reported *LMAN1* F5F8D patients carry two null *LMAN1* alleles, our results suggest that hypomorphic *LMAN1* mutations in humans may result in milder phenotypes that are missed clinically. LMAN1-null mice also demonstrate a partially penetrant, strain-specific, perinatal lethality. This result suggests a broader role for murine LMAN1 in protein secretion in addition to FV and FVIII. LMAN1-deficient mice also exhibit a mild thrombocytopenia that has not been observed in humans, suggesting a potential role for LMAN1 in platelet development.

In preliminary studies, proteomic profiling of mouse tissues by mass-spectrometry and yeast-two-hybrid assays were applied in an effort to identify novel LMAN1/MCFD2 cargos and interacting proteins. Comparison of plasma samples from $Lman1^{+/+}$ and $Lman1^{-/-}$ mice by mass-spectrometry revealed SERPINA3K, a serine protease inhibitor, as a potential novel LMAN1 cargo that was in lower abundance in the plasma of LMAN1-deficient mice than wild-type mice. Complement component C3, a serine protease, was also under-represented in the plasma of $Lman1^{-/-}$ mice. It is possible that SERPINA3K and complement component C3 cleavage fragments represent newly identified protease/protease inhibitor complexes, and that the efficient secretion of one of these components to the plasma is LMAN1-dependent.

Taken together, the work outlined here provides new insight into the function of LMAN1/MCFD2 in the mammalian secretory pathway and the phenotype of LMAN1-deficiency *in vivo*.

CHAPTER 1 :
INTRODUCTION

COMBINED DEFICIENCY OF COAGULATION FACTORS V AND VIII

Defects in the synthesis, secretion, and function of coagulation factors V (FV) and VIII (FVIII) are important causes of inherited bleeding disorders. FV and FVIII are large, secreted glycoproteins (280-330 kDa) with critical roles in hemostasis and thrombosis. FV and FVIII function as essential cofactors for the proteolytic activation of prothrombin and factor X in the coagulation cascade, respectively. Genetic deficiency of FVIII results in hemophilia A, whereas inherited FV deficiency results in parahemophilia. A gain-of-function mutation (FV Leiden) or increased FVIII activity are both major risk factors for venous thrombosis.[1,2] Finally, combined deficiency of factors V and VIII (F5F8D, OMIM 227300) is an autosomal recessive bleeding disorder characterized by reduction of both FV and FVIII to ~10-15% of normal levels.[3,4] F5F8D was first identified and described in two brothers by Oeri *et al.* in 1954.[5] The molecular etiology of F5F8D is distinct from that of parahemophilia or hemophilia A, resulting from impaired post-translational secretion of FV and FVIII, rather than from a defect within the *F5* or *F8* genes.

Epidemiology

Data on the epidemiology of F5F8D and the distribution of this disease around the world are available in the Rare Bleeding Disorders Database (http://www.rbdd.org) and from the World Federation of Hemophilia (WFH) global survey (http://www1.wfh.org/publications/files/pdf-1488.pdf).

F5F8D is an autosomal recessive disorder; males and females are affected equally. F5F8D is extremely rare in the general population (1:1,000,000), although it is observed with a higher frequency in populations with consanguineous marriages.[6-8] As compared to the general population, the prevalence of F5F8D is much higher among Middle Eastern Jewish individuals and non-Jewish Iranians (~1:100,000).[9] Most reported cases of F5F8D are from Iran and Italy, with smaller numbers of case reports originating in Pakistan, Iraq, Lebanon, and Algeria.[6] Additional cases have been reported from European countries, India, Asian countries such as Japan, China, and Thailand, and also from South Africa, the United States, and South America.[6] Overall, the global distribution of F5F8D patients is broad with variable frequencies, with the majority of new F5F8D cases in Middle-Eastern countries and India, an intermediate frequency in South America and Asia, and a small proportion of cases in North America, Europe, and Africa.[6] As of 2004, at least 140 patients in 81 families had been diagnosed with F5F8D.[10] According to the 2011 World Federation of Hemophilia Global Survey, 62 countries reported a current total of 352 F5F8D patients.[11] However, given the mild bleeding manifestations in most F5F8D patients, it is likely that F5F8D is under-diagnosed, or that patients with F5F8D may be misdiagnosed with mild hemophilia A or parahemophilia. Accurate data about the global distribution of F5F8D is also lacking because many countries do not maintain reliable national registries, and there are probably higher rates of detection and accurate F5F8D diagnosis in countries with higher F5F8D prevalence. There are 117 *LMAN1* variants reported in the University of Washington Exome Variant Server database.[12] Among European Americans represented in the database, the frequency of these variants is ~1/8,500, while the frequency is ~1/4,400 among African Americans. Of the 117 *LMAN1* variants in the database, 24 are synonymous coding variants, 31 are missense variants, 1 is a 3' UTR variant, and 1 is a frameshift mutation.

Clinical Presentation

F5F8D results in mild-to-moderate bleeding symptoms associated with prolonged prothrombin time (PT) and partial thromboplastin time (PTT) due to deficiency of FV and FVIII, respectively.[7-9,13] The FV and FVIII levels (both activity and antigen) in most

F5F8D patients are ~10-15% of normal.[3,4] Easy bruising, epistaxis, and excessive bleeding during or after trauma are some of the most common symptoms, as well as menorrhagia and postpartum hemorrhage in affected women.[9,14] However, more severe bleeding symptoms may include hemarrthroses, muscular hematomas, umbilical cord bleeding, and rarely gastrointestinal and central nervous system bleeding.[6] F5F8D patients have normal VWF levels, as is the case for hemophilia A patients.[9] Except for the isolated combined deficiency of FV and FVIII, there are currently no other known clinical phenotypes associated with F5F8D. The primary goal of treatment for bleeding episodes in F5F8D patients is replacement of FV and FVIII, but specific management is determined on a case-by-case basis by the nature of the bleeding and the FV and FVIII levels of the patient.[6] Prophylactic treatment is not usually necessary, but may be used in cases of severe recurrent bleeding.

PATHOPHYSIOLOGY OF F5F8D

LMAN1/MCFD2 cargo receptor

F5F8D is caused by mutations in *LMAN1* (also known as *ERGIC-53*) or *MCFD2*, which result in a defect in the early secretory pathway and the impaired secretion of FV and FVIII to the plasma.[3,4] LMAN1 is encoded by a ~29 kb gene (13 exons) located on chromosome 18q21, and MCFD2 is encoded by a ~19 kb gene (4 exons) located on chromosome 2p21. LMAN1 and MCFD2 bind together in a 1:1 ratio to form the only known mammalian secretory cargo adaptor/receptor complex. The LMAN1/MCFD2 complex is responsible for efficient secretion of FV and FVIII to the plasma. F5F8D is therefore an example of locus heterogeneity, in which identical phenotypes result from mutations at two or more distinct genetic loci. Although the clinical features of F5F8D are similar to that of FV deficiency (parahemophilia, chromosome 1) and FVIII deficiency (hemophilia A, chromosome X), the inheritance of F5F8D is distinct from the coinheritance of both of these disorders. Whereas parahemophilia and hemophilia A result from defects in the *F5* and *F8* genes, F5F8D results from the impaired post-

translational secretion of FV and FVIII to the plasma mediated by the shared LMAN1/MCFD2 secretory cargo receptor.

LMAN1 is a homo-hexameric, 53-kDa type I transmembrane receptor protein (510 amino acids) that spans the ER membrane. The protein structure consists of a 480 amino acid ER luminal domain, a transmembrane domain, and a short 12 amino acid cytoplasmic tail domain (Figure 1-1A). The luminal domain includes an N-terminal carbohydrate recognition domain (CRD, ~200 amino acids) and a membrane-proximal α-helical coiled domain (the stalk domain). The CRD is homologous to other lectin domains known to bind mannose-rich glycoproteins. The LMAN1 C-terminal cytosolic tail domain contains both a di-phenylalanine ER-exit motif (FF) and a di-lysine ER-retrieval signal (KK). The di-phenylalanine motif allows for interactions with components of COPII vesicle coats, promoting efficient ER export of LMAN1 and its bound cargos.[15] The di-lysine motif serves as an ER retrieval signal for LMAN1,[16] facilitating the constitutive recycling of LMAN1 between the endoplasmic reticulum (ER) and the Endoplasmic Reticulum Golgi Intermediate Compartment (ERGIC) in the early secretory pathway (Figure 1-2).[17]

LMAN1 was discovered in 1988,[18] and has been used extensively as a molecular marker for the ERGIC since that time. However, its function as a secretory cargo receptor remained unknown for many years. LMAN1 is a member of a family of homologous leguminous type (L-type) mannose-binding lectins,[19-21] along with VIP36,[22] VIPL,[23,24] and ERGL.[25] The L-type lectins share similar carbohydrate recognition domains, but they have different subcellular localization patterns, distinct glycoprotein-binding specificities, and different pH optima for glycoprotein binding.[26,27] The intracellular localization pattern of LMAN1 (cycling between the ER and the ERGIC) and the homology of LMAN1 to the family of mannose-binding lectins were two of the first indications that LMAN1 may function as a secretory cargo receptor in mammalian cells. The surprising identification of *LMAN1* as the disease-causing gene in F5F8D provided the first direct evidence for LMAN1 as a mammalian ER-to-Golgi cargo receptor for FV and FVIII.[4] In contrast, evidence suggests that the VIP36 homolog does not play a role in anterograde transport of secretory cargos (i.e. not an ER cargo receptor), but rather that it

binds glycoproteins in the Golgi[19,28] (including A1AT) and recycles them back to the ER as part of a post-ER quality control mechanism.[27,29,30] The lectin family member VIPL (VIP36-like) is more similar in sequence to VIP36 than to LMAN1, and is a non-cycling resident protein of the ER.[24]

MCFD2 is a 16-kDa soluble luminal protein that contains two calmodulin-like EF-hand motifs (Figure 1-1B). MCFD2 contains a signal sequence mediating translocation into the ER, but does not contain a C-terminal KDEL retrieval signal or any other identifiable sorting signal to direct its intracellular localization or recycling to the ER. Rather, calcium-dependent interactions of the MCFD2 EF-hand domains with LMAN1 in the ER lumen cause MCFD2 to be retained in the early secretory pathway, co-localizing with LMAN1 to the ERGIC.[3]

LMAN1 and MCFD2 bind via calcium-dependent interactions in a 1:1 stoichiometric ratio to form a hexameric complex at the ER membrane (Figure 1-1C). This complex serves as a specific cargo receptor for the selective packaging of FV and FVIII into COPII-coated vesicles for subsequent ER-to-Golgi transport.[3] LMAN1/MCFD2 then recycles from the ERGIC back to the ER to facilitate another round of secretory cargo transport. LMAN1/MCFD2 complex formation is essential for cargo receptor function. MCFD2 missense mutations, particularly within the EF-hand domains, disrupt this interaction and cause F5F8D by abolishing the complex formation.[31-33] Although both LMAN1 and MCFD2 are required for efficient secretion of FV and FVIII, LMAN1 binds some cargo glycoproteins in an MCFD2-independent manner.[34,35] In the absence of MCFD2, LMAN1 is still localized properly to the ER and ERGIC.[35] In contrast, in the absence of LMAN1, MCFD2 is secreted (MCFD2 is normally retained within the ER and the ERGIC due to its physical interactions with LMAN1).[35]

Genetic analysis of F5F8D patients

The underlying molecular pathophysiology of F5F8D was not understood until 1998, when Nichols *et al.* used homozygosity mapping and positional cloning to discover that homozygous or compound heterozygous mutations in *LMAN1* account for ~70% of

F5F8D cases.[4] In 2003, Zhang *et al.* identified *MCFD2* as a second locus associated with F5F8D (~30% of cases),[3] and it is now clear that homozygous or compound heterozygous mutations in either *LMAN1* or *MCFD2* are responsible for every reported case of F5F8D.[36] No F5F8D patients have been reported with mutations in both *LMAN1* and *MCFD2*. As of 2013, the International Society on Thrombosis and Haemostasis and the Rare Bleeding Disorders Database (RBDD) of known mutations causing rare bleeding disorders documents 34 distinct *LMAN1* mutations and 22 distinct *MCFD2* mutations in F5F8D patients. Three *LMAN1* mutations are missense mutations, while the majority of mutations are predicted to be null (7 nonsense mutations, 7 splicing defect mutations, and 17 deletion/insertions). Two of the missense mutations are also likely to be functionally null alleles. The first is a substitution of a threonine for the initiator methionine (codon 1), which would interfere with normal translation initiation of LMAN1 and would be predicted to result in no functional protein product. The second missense mutation is an arginine to cysteine substitution near the transmembrane domain. This arginine residue is known to play an important role in the formation of inter-molecular disulfide bonds, which are important for LMAN1 oligomerization and function.[37] The functional significance of this arginine to cysteine substitution was demonstrated *in vitro* and resulted in failed accumulation of the mutant LMAN1 in the cell, indicating that this mutant LMAN1 protein is likely unstable and nonfunctional.[31] The abundance of null *LMAN1* mutations in F5F8D patients is suggestive of a threshold effect of LMAN1 function, but this has never been confirmed experimentally and no data exists about the relationship between *LMAN1* expression levels and the level of LMAN1-dependent cargo secretion.

Among the reported *MCFD2* mutations, there are 11 missense mutations, one nonsense mutation, four splicing defect mutations, and six deletion/insertions.[38] All missense mutations are localized to the EF-hand domains, which mediate interactions with LMAN1 as well as FV and FVIII.[32] Missense mutations are sufficient to impair the interaction of MCFD2 with LMAN1, which results in F5F8D.[31,32]

Genotype-phenotype correlations

Establishing genotype-phenotype correlations in rare bleeding disorders such as F5F8D is often difficult, primarily because the bleeding phenotypes are highly variable, even among patients with the same genotype.[38] Modifier genes and environmental factors likely contribute to the variable severity of bleeding phenotypes among F5F8D patients, although no specific modifier genes for F5F8D clinical variability are currently known.

F5F8D patients with *MCFD2* mutations are reported to have slightly lower levels of FV and FVIII than patients with *LMAN1* mutations,[39] based on a review of available published data for 46 patients with *MCFD2* mutations and 96 patients with *LMAN1* mutations. This suggests that MCFD2 may play a more direct role in transporting FV and FVIII than does LMAN1, consistent with the finding that MCFD2 can bind to FVIII independent of LMAN1.[40] Decreases in plasma FV and FVIII levels are correlated in F5F8D patients, and platelet FV levels are reduced to the same extent as plasma FV.[36] These data are consistent with the loss of function of the common ER-to-Golgi receptor complex for FV and FVIII.[41]

THE CELLULAR SOURCE OF FVIII BIOSYNTHESIS IS CONTROVERSIAL

FV and FVIII structure and function

FV and FVIII are secreted glycoproteins that share a similar domain structure (A1-A2-B-A3-C1-C2), with 40% amino acid sequence homology between their corresponding A and C domains (Figure 1-3A).[42] The heavily glycosylated B-domain of FV and FVIII is required for efficient secretion by LMAN1 (Figure 1-3B).[43,44] After secretion to the plasma, FV and FVIII circulate as inactive precursors that are activated by thrombin in order to serve as non-enzymatic cofactors for FXa and FIXa, respectively. Numerous post-translational modifications of FV and FVIII occur, including cleavage of the signal peptide, disulfide bond formation, glycosylation, and sulfation of specific tyrosine residues.[45]

FV is found in the plasma and in platelet α granules as a 330 kDa single chain polypeptide.[46] Human plasma FV concentrations are ~ 7 ug/mL. It was previously reported that the human plasma and platelet FV pools are both synthesized by hepatocytes;[47] megakaryocytes endocytose and modify FV from the plasma in order to form the platelet FV pool.[48] In contrast, the murine plasma and platelet FV pools are biosynthetically distinct, with plasma FV synthesized exclusively by hepatocytes, and platelet FV synthesized by megakaryocytes.[49]

Plasma FVIII is present at much lower concentration than FV (~200 ng/mL), and the full length protein is processed to a heterodimer comprised of an 80 kDa light chain and a 200 kDa heavy chain.[50] The FVIII light chain binds non-covalently to the amino-terminus of von Willebrand Factor (VWF), and this tight association between FVIII and VWF increases the plasma half-life of FVIII.[51]

Conflicting evidence for the cellular source of FVIII synthesis

Though VWF is synthesized exclusively in endothelial cells and megakaryocytes, the primary cellular source of FVIII biosynthesis has been controversial. Delayed rise of FVIII in von Willebrand Disease (VWD) patients infused with VWF demonstrates that VWF and FVIII do not require synthesis in the same cell for complex formation. Liver, spleen, and lung transplants have corrected hemophilia A in patients and dog models,[52-60] and dozens of conflicting reports have implicated hepatocytes, endothelial cells, lymphoid tissues, and renal or pulmonary epithelial cells as the source of FVIII.[61-64] Cumulatively, the majority of evidence points to either an endothelial and/or hepatocyte origin of FVIII synthesis. Since LMAN1 is part of a ubiquitously expressed ER cargo receptor that is responsible for the efficient secretion of both FV and FVIII to the plasma, manipulation of LMAN1 presents a convenient molecular approach to potentially identify the cellular source of FVIII *in vivo* (see Chapter 2). The identification of the biosynthetic source of FVIII may have important translational implications for the future development of *F8* gene therapy for the treatment of hemophilia A.

LMAN1 BIOLOGY AND ER-TO-GOLGI PROTEIN TRANSPORT

Overview of the secretory pathway

The protein secretory pathway (Figure 1-2) is composed of a series of specialized compartments. Proteins destined for secretion from the cell or for residence in secretory compartments (lysosomes, endosomes, plasma membrane, etc.) traffic through this pathway, undergoing quality-control checks and modifications. The first critical step in this process is transport from the ER to the Golgi. Correctly folded proteins in the ER lumen are packaged into COPII-coated vesicles for anterograde transport towards the Golgi. The vesicles uncoat and fuse with one another, forming the ERGIC, and resident ER proteins like LMAN1 recycle back to the ER in COPI-coated retrograde vesicles to participate in another cycle of protein transport.

Nearly 30% of all mammalian proteins travel through the secretory pathway. The mechanism(s) by which cargo proteins are recruited into COPII vesicles are not clear. Some transmembrane cargos interact directly with the COPII coat component SEC24 via specific sorting signals.[65] However, for soluble cargos in the ER lumen, two competing models have been proposed.[66,67] Some secretory cargos are thought to be secreted via a passive bulk-flow model (no signal or receptor required); the export of amylase and chymotrypsinogen from the ER of pancreatic exocrine cells is reported to follow this model.[68,69] Under the bulk-flow model, soluble proteins that do not contain a specific ER retention signal are gathered into transport vesicles without a concentration step, but this model does not explain why known ER resident proteins have been shown to be depleted from secretory transport vesicles.[70,71] In contrast, evidence now suggests that the majority of soluble secretory protein cargos require a selective cargo-receptor for efficient recruitment into budding COPII vesicles, as demonstrated in Figure 1-4.[72-78]

ER cargo receptors

Approximately 8,000 proteins are trafficked through the mammalian secretory pathway, some of which are ubiquitously expressed, and some of which have highly specialized

functions in specialized secretory tissues. The intracellular transport machinery must be able to efficiently and accurately sort and transport this highly variable population of proteins to their correct intra- or extra-cellular destinations. Recent developments suggest that this sorting and transport specificity is conferred, in part, by signaling elements encoded within secretory proteins, and by the selective recruitment of these secretory cargos into COPII vesicles by cargo receptors at the ER membrane.[66]

Much of the characterization of ER cargo receptors and their role in selective recruitment of cargos into COPII-coated vesicles has been performed in yeast. A few clearly defined yeast cargo receptors are known to cycle between the ER and Golgi compartments, including the p24 proteins and Erv29p.[79-81] The p24 protein Emp24p is required for efficient ER-export of specific secretory cargos, including the GPI-anchored protein Gas1p.[79,82] More recent reports indicate that Erv29p acts as an export receptor for another selective subset of soluble secretory proteins in yeast (glycol-pro-α factor, carboxypeptidase Y, and proteinase A).[80,83] The specific signal embedded in the glycol-pro-α factor cargo that is recognized by Erv29p (the ER receptor) was recently identified, consisting of an I-L-V motif in the pro-region of glycol-pro-α factor.[84] The identification of additional signal-receptor pairs will be critical for understanding the general mechanism of receptor-mediated cargo transport. As single-celled organisms, yeast secrete a relatively small number of proteins outside of the cell. In contrast, mammalian cells have highly complex secretory proteomes and secrete thousands of proteins outside of the cell. It is therefore very surprising and intriguing that LMAN1/MCFD2 is currently the only known mammalian cargo-receptor in the early secretory pathway.

LMAN1 is highly evolutionarily conserved, with close homologs in multiple vertebrate and invertebrate species, including *C. elegans* and *D. melanogaster*. Similarly, MCFD2 has multiple vertebrate orthologs with > 60% amino acid identity in species such as mouse, rat, and zebrafish. Possible MCFD2 paralogs in invertebrate species may also exist.[3] Importantly, the evolutionary divergence of invertebrates from vertebrates occurred prior to the establishment of a blood clotting system, suggesting a more general role for LMAN1/MCFD2 in protein secretion in addition to FV and FVIII. This

evolutionary conservation of LMAN1 and MCFD2, in combination with the ubiquitous expression patterns of these genes, presents an exciting opportunity to explore other roles for LMAN1.

Additional LMAN1 cargos

In addition to FV and FVIII, which are the best characterized LMAN1 secretory cargos, several other proteins have been reported to rely on LMAN1 for secretion. Utilizing a fluorescent protein-based protein fragment complementation assay to identify protein interactions in the lumen of the secretory pathway, Nyfeler *et al.* reported LMAN1 as an intracellular transport receptor for alpha-1 antitrypsin (A1AT).[85] A1AT is a protease inhibitor belonging to the serpin superfamily, and it is of clinical relevance because A1AT deficiency leads to liver and lung disease.[86] However, no known A1AT secretion defect has been observed in F5F8D patients. In addition to FV, FVIII, and A1AT, two lysosomal cathepsin proteases (cathepsin C (CTSC)[87] and cathepsin Z (CTSZ)[88]) have also been reported to rely on LMAN1 for ER-to-Golgi transport. Most recently, Chen *et al.* performed a cDNA library screen from HepG2 cells via a fluorescent protein fragment complementation assay and identified Mac-2 binding protein (Mac-2BP) as a novel LMAN1-transported glycoprotein cargo.[89]

Other proposed roles for LMAN1

In addition to its known role in FV and FVIII secretion and F5F8D, LMAN1 has been implicated as a mediator of several other human diseases. By analyzing mutation frequencies of genes of the cellular glycosylation machinery in microsatellite unstable tumors, Roeckel *et al.* determined that *LMAN1* shows a high mutation frequency in microsatellite unstable colorectal cancer cell lines, carcinomas, and adenomas.[90] The authors proposed that LMAN1 mutational inactivation may contribute to tumorigenesis in tumors with microsatellite instability. In addition, LMAN1 has been reported to play a role (along with ERp44) in facilitating the efficient polymerization and secretion of immunoglobulin M (IgM) polymers in antibody-secreting cells, but no defects in IgM secretion have been reported in F5F8D patients.[91,92] Furthermore, it was previously

shown that N-glycosylation of human nicastrin is required for it to interact with calnexin and ERGIC-53.[93] Nicastrin is a component of the gamma-secretase complex, which is responsible for cleavage of amyloid precursor protein, Notch, and several other type 1 transmembrane proteins. Most recently, Klaus *et al.* reported an entirely novel role for the LMAN1 cargo receptor, demonstrating that it is required for the production of infectious particles from several highly pathogenic RNA virus families (arenavirus, coronavirus, and filovirus).[94]

ANIMAL MODELS OF LMAN1 DEFICIENCY

Our lab previously generated an LMAN1-deficient mouse line[95] carrying a genetically engineered gene-trap allele (*Lman1*^{gt1}) designed to disrupt the *Lman1* gene. Mice homozygous for this allele reproduce the human F5F8D phenotype, exhibiting low FV and FVIII plasma levels that are ~50% of wild-type levels. The reduction in FV and FVIII levels in this mouse model are therefore less severe than the reduction of FV and FVIII levels in human F5F8D patients (~10-15% of normal levels). These data suggest that the *Lman1*^{gt1} allele does not result in complete disruption of LMAN1 function, or alternatively, that F5F8D manifests differently in mice and humans. In addition, *Lman1*$^{gt1/gt1}$ mice demonstrate a partially penetrant, perinatal lethality that is specific to some genetic strain backgrounds.[95] This unexpected lethality and the broad evolutionary conservation of both *Lman1* and *Mcfd2* suggest the hypothesis that additional critical LMAN1 secretory cargos remain to be discovered in addition to FV and FVIII.

Mutations in the *drosophila LMAN1* homolog, *rhea*, lead to a late embryonic recessive lethal phenotype.[96,97] Mutant embryos develop somatic wing and muscle developmental defects (complete detachment of the somatic musculature from the epidermis). In contrast, unpublished work performed in our lab showed that RNAi inactivation of *ile-1*, the *C. elegans* homolog of *LMAN1*, results in no detectable phenotype (B. Zhang, R. Ellis, and D. Ginsburg).

IMPORTANT REMAINING QUESTIONS FOR LMAN1 BIOLOGY

The growing body of literature characterizing the LMAN1/MCFD2 cargo receptor and the genetic etiology of F5F8D serves as an instructive example of a rare heritable human disease that can provide fundamental insights into basic biologic processes. Use of *in vitro* and *in vivo* models of LMAN1 and MCFD2 deficiency will enhance our understanding of FV and FVIII biosynthesis, and provide critical insight into the basic cellular process of ER-Golgi transport. Identification of other cargo proteins relying on LMAN1/MCFD2 for efficient secretion should provide further understanding of this important pathway, and subsequently it may be possible to define the common structural features that mediate these interactions. Such work may also inform the identification and characterization of other novel ER cargo receptors. Further study of selective ER-to-Golgi cargo transport may have important implications for the therapeutic modulation of FV and FVIII levels for the treatment of bleeding disorders or the prevention of thrombosis. In addition, investigation of LMAN1/MCFD2 receptor mediated secretion should provide useful information relevant to the future development of safe, efficient, and efficacious gene therapies for hemophilia A.

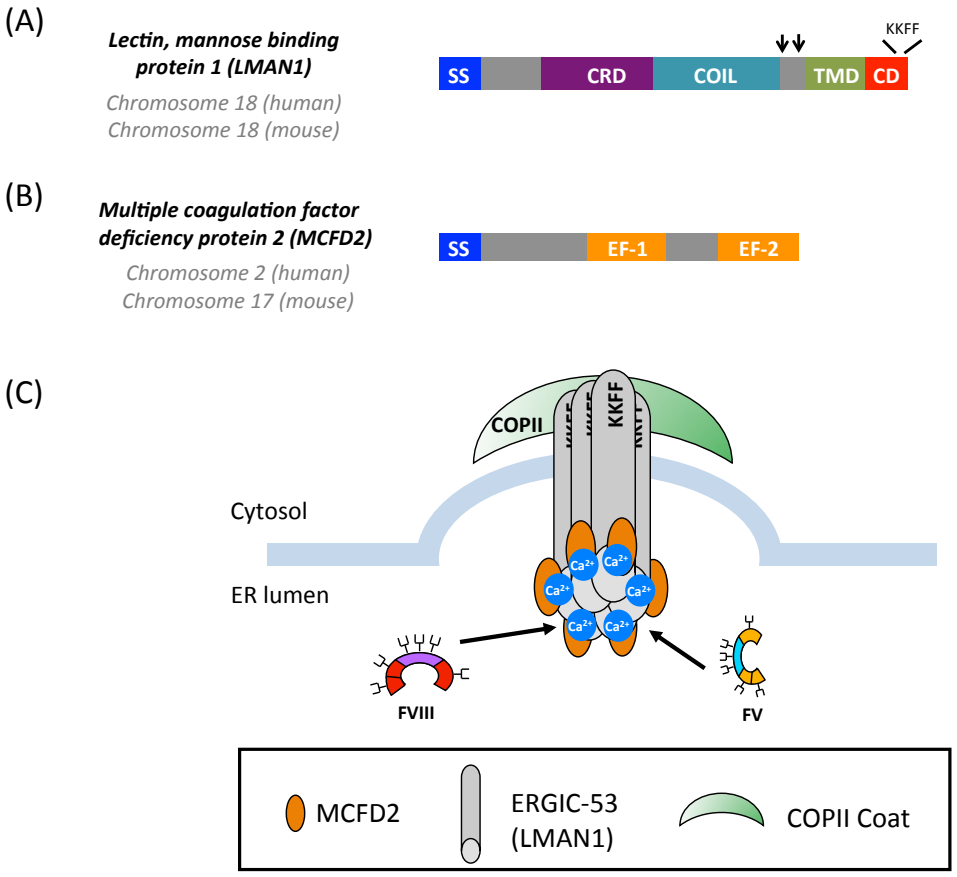
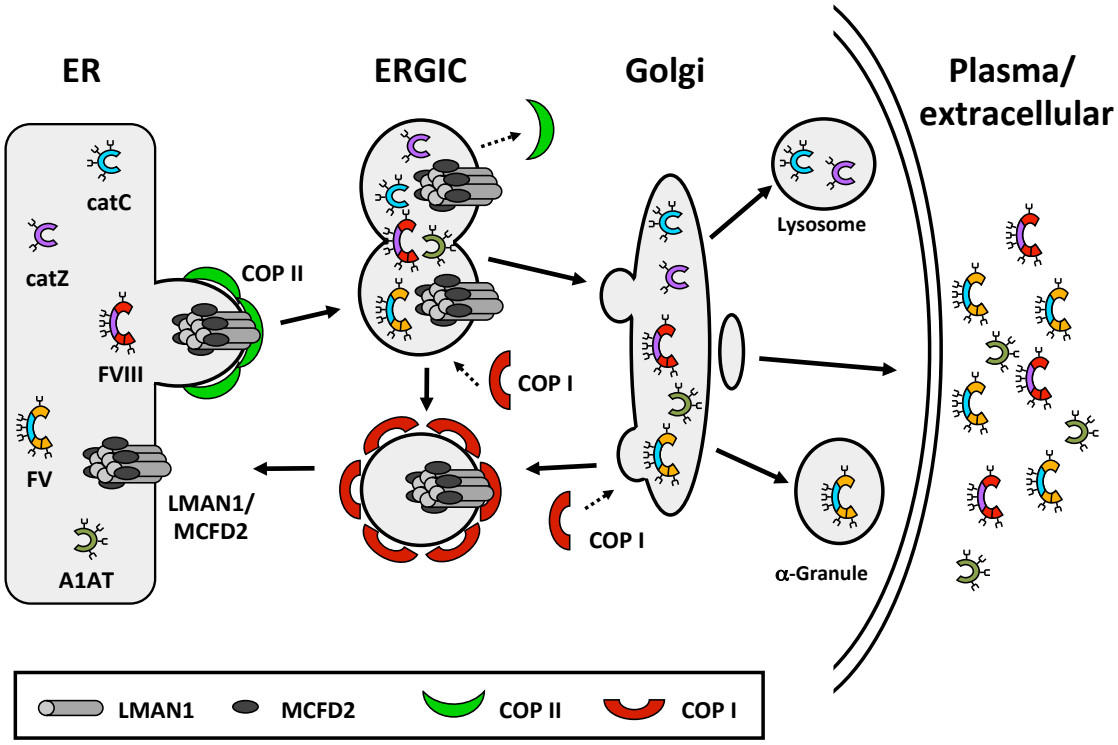


Figure 1-1: Structure of the LMAN1/MCFD2 receptor complex

(A) LMAN1 is a 53-kDa type-I transmembrane protein. Features include a signal sequence (SS), a carbohydrate recognition domain (CRD), a coiled-coil domain (COIL), a transmembrane domain (TMD), and a C-terminal cytosolic domain (CD) containing a di-phenylalanine (FF) ER exit motif and a di-lysine (KK) ER-retrieval signal. The KKFF motif enables cycling of LMAN1 between the ER and the ER-Golgi intermediate compartment (ERGIC). The 12 amino acid cytosolic domain facilitates interactions with the COPII coat. Two luminal cysteines near the TMD (black arrows) form inter-molecular disulfide bonds, which facilitate the oligomerization of LMAN1 along with structural determinants in the LMAN1 luminal and TM domains. **(B)** MCFD2 is a 145 amino acid 16-kDa soluble monomeric protein. Features include a signal sequence (SS) for ER translocation and two calcium-binding EF-hand domains (EF-1 and EF-2) at the C-terminus. Unlike LMAN1, MCFD2 does not contain an ER retrieval sequence. The recycling of MCFD2 between the ER and ERGIC is dependent upon its interactions with LMAN1. **(C)** LMAN1 and MCFD2 form a cargo receptor complex that is required for efficient transport of coagulation factor V (FV) and factor VIII (FVIII) from the ER to the Golgi. MCFD2 interacts with LMAN1 in the ER lumen with a 1:1 stoichiometry. A deficiency of either component, or a disruption of LMAN1/MCFD2 complex formation, results in inadequate transport and decreased secretion of FV and FVIII. Orange ovals represent MCFD2, grey cylinders represent LMAN1.



Adapted from Trends Biochem Sci. 2007 Aug;32(8):381-8.

Figure 1-2: Model for ER-to-Golgi cargo transport by the LMAN1/MCFD2 receptor complex

Hexameric LMAN1 is recruited into budding COPII vesicles at the ER membrane via a C-terminal diphenylalanine motif (FF). In combination with MCFD2, the LMAN1/MCFD2 cargo receptor complex recruits specific cargo proteins (FV, FVIII, A1AT, CTSC, CTSZ) into budding COPII vesicles for anterograde transport to the ERGIC. COPII coat components dissociate and the vesicles fuse at the ERGIC. LMAN1/MCFD2 are recycled back to the ER via a di-lysine motif (KK) that facilitates interactions with the retrograde COPI vesicle coats. Secretory cargos are transported onwards in the secretory pathway to their final intracellular (lysosomes, α-granules), membrane, or extracellular (plasma, extracellular matrix) destinations.

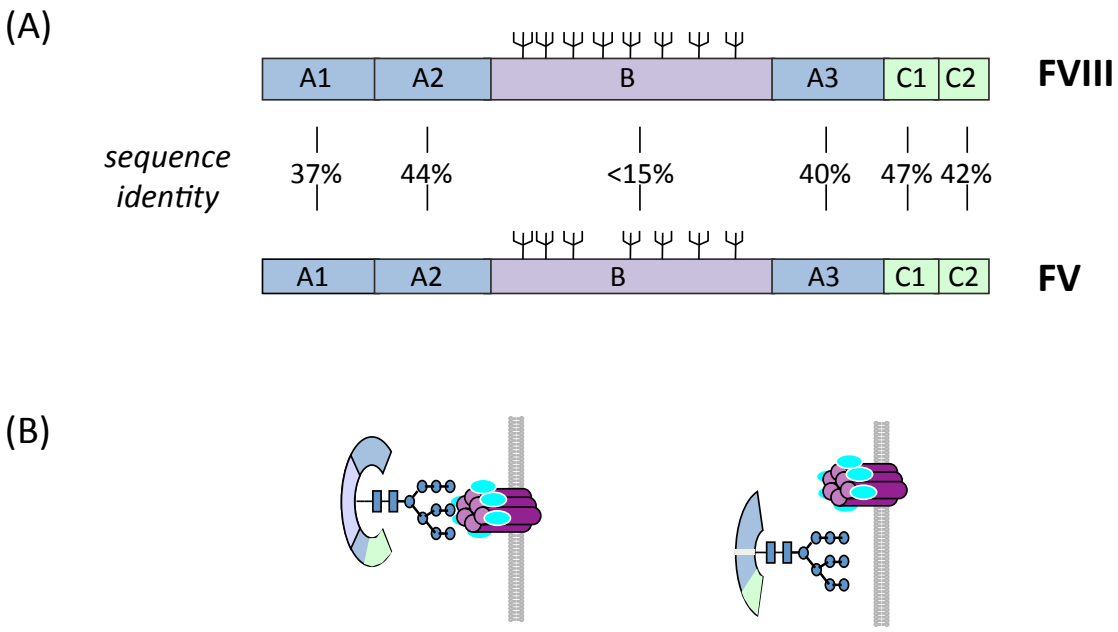
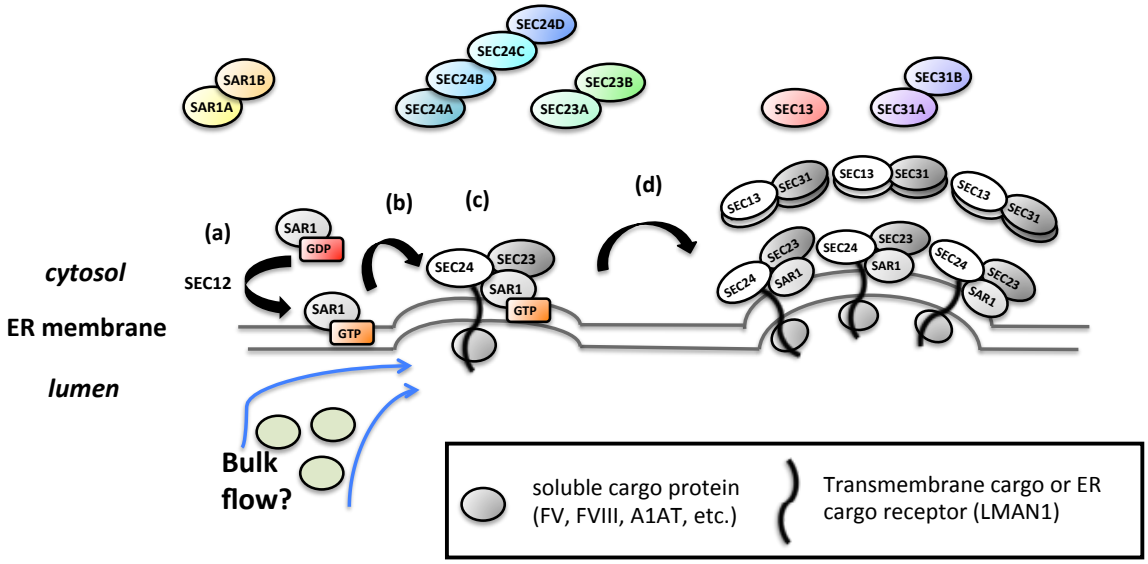


Figure 1-3: Domain structures and sequence identities of FV and FVIII

(A) FV and FVIII share a similar pro-cofactor structure containing six domains (A1-A2-B-A3-C1-C2), with ~40% homology in the A and C domains for FV and FVIII. The B domain is not homologous to other known structures and is heavily modified by N-linked glycosylation. The B domains of FV and FVIII share little sequence homology with each other. The A domains of FV and FVIII also share important sequence identify with the plasma copper-binding protein ceruloplasmin (not shown). Branched structures represent N-linked glycosylation sites. **(B)** The glycosylated B-domain of FVIII is required for its interaction with LMAN1.[42]



(a) SAR1 activation (b) SEC23/24 recruitment (c) Cargo binding (d) SEC13/31 polymerization

Figure 1-4: COPII vesicle composition and assembly

Proteins passing quality control checks in the ER are exported for anterograde transport by bulk-flow (no receptor or signal required) or by selective receptor-mediated packaging into COPII vesicles. COPII vesicle formation involves the hierarchical recruitment and oligomerization of several coat factors at specific ER Exit Sites. COPII-coated vesicle coats are composed of the small GTP-binding protein SAR1, the hetero-dimeric SEC23-SEC24 complex, and the hetero-tetrameric 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, some of which may 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.

CHAPTER 2 :
MURINE COAGULATION FACTOR VIII IS SYNTHESIZED IN ENDOTHELIAL CELLS

*Note: Chapter 2 was accepted for publication in **Blood** as a regular article in March 2014 (Authors: Lesley Everett, Audrey Cleuren, Rami Khoriaty, and David Ginsburg). Lesley Everett performed all experiments except for the RiboTag mouse studies, which were performed by Audrey Cleuren. Rami Khoriaty assisted with blood sample collection and analysis. Submission date: February 8, 2014.*

INTRODUCTION

Plasma coagulation factor VIII (FVIII) is tightly associated with von Willebrand Factor (VWF). Though it is well established that VWF is synthesized exclusively in endothelial cells and megakaryocytes, the cellular source of FVIII has been debated for decades. The delayed rise of FVIII in von Willebrand Disease patients infused with VWF demonstrates that VWF and FVIII do not require synthesis in the same cell for complex formation.[98] Dozens of conflicting reports have implicated numerous cell and tissue types as the source of FVIII, though the majority of evidence points to either an endothelial and/or hepatocyte origin of FVIII synthesis.[61-63]

Multiple curative organ transplantation studies in hemophilia A patients and experimental dog models have shown that the liver contributes significantly to FVIII biosynthesis,[52-55] though the failure of FVIII levels in hemophilia A liver transplant recipients to respond to desmopressin suggested FVIII localization to a different compartment than VWF.[99] As early as 1971, Webster *et al.* demonstrated extra-hepatic FVIII synthesis by showing that FVIII synthesis was maintained after replacing the livers of normal dog recipients with

hemophilic livers.[56] Similarly, extra-hepatic FVIII production was shown in humans after transplantation of a hemophilia A donor liver into a non-hemophilic recipient.[57] Spleen and lung transplantation has been shown to restore FVIII plasma levels in several animal models of hemophilia,[58-60,64] and FVIII production also has been documented by human purified microvascular endothelial cells from lung *in vitro*.[100]

Immunoradiometric assays localized FVIII antigen to guinea pig liver, spleen, lung, kidney, and to isolated hepatocytes,[61,101] and to human lymph nodes, lung, liver, and spleen,[102] whereas microscopy studies suggested synthesis in human hepatocytes[62] or human liver sinusoidal endothelial cells (LSECs).[103,104] *F8* mRNA transcripts have been detected in purified murine LSECs and hepatocytes at comparable levels, but not in Kupffer cells,[63] and by others in human spleen, lymph nodes, kidney, and in isolated hepatocytes, but not in white blood cells or cultured umbilical vein endothelial cells.[61] *F8* mRNA transcripts have also been shown to be synthesized in hemophilia A mouse livers by hepatocytes and LSECs that were derived from transplanted wild-type bone marrow progenitor cells.[105]

We took advantage of a genetically modified mouse model that is deficient for LMAN1 (Lectin, mannose-binding protein 1, also known as ERGIC-53). LMAN1 is part of a ubiquitously expressed cargo receptor that cycles between the endoplasmic reticulum (ER) and the ER-Golgi intermediate compartment (ERGIC) in the early secretory pathway. LMAN1 is required for the efficient secretion of FVIII as well as coagulation factor V (FV) to the plasma. Mutations in *LMAN1* cause an autosomal recessive bleeding disorder known as combined deficiency of coagulation factors V and VIII (F5F8D). Although FV and FVIII are synthesized at markedly different levels (human plasma FV ~7.0 ug/mL, and FVIII ~0.2 ug/mL) and potentially in different tissues, loss of the LMAN1 cargo receptor in F5F8D leads to parallel reduction of both of these factors to ~10-15% of normal levels.[3,4]

We previously reported a genetically engineered mouse carrying a gene-trap interruption of the *Lman1* gene.[95] Mice homozygous for this gene-trap allele ($Lman1^{gt1/gt1}$) exhibit

reduced FV and FVIII levels that are ~50% of normal, in contrast to the ~10-15% levels in human patients with F5F8D. As expected, this modest reduction in FV and FVIII results in a much milder phenotype than reported for hemophilia A mice carrying a *F8* gene-targeted allele leading to complete absence of FVIII.[106] We now describe a second, independent conditional *Lman1* allele, and analysis of mice with LMAN1 deficiency induced specifically in the hepatocyte or endothelial cell compartments. Deletion of *Lman1* only in hepatocytes results in reduced FV, but normal plasma FVIII, whereas endothelial *Lman1* deletion reduces FVIII but not FV plasma levels. In addition, analysis of purified, endothelial-specific polysomes by qRT-PCR confirms *F8* mRNA expression in endothelial cells (relative to total tissue lysates), with considerable heterogeneity in the level of *F8* mRNA expression across different vascular beds.

MATERIALS AND METHODS

Lman1 conditional knockout mice

Embryonic stem (ES) cell clone B06 (C57BL/6 genetic background) carrying a conditional *Lman1* gene-trap allele (*Lman1*^{cgt}, also known as *Lman1*$^{tm1a(KOMP)Wtsi}$) with a gene-trap insertion into *Lman1* intron 1 and *loxP* sites flanking exons 2-3 was obtained from the Knockout Mouse Project (KOMP; www.komp.org) (Figure 2-1A). Correct targeting of the conditional gene-trap construct was confirmed by long-range PCR at both the 5' and 3' ends of the targeted allele with Phusion Hot Start High-Fidelity DNA Polymerase (Finnzymes) (Supplementary Figure 2-1A) using the PCR primers listed in Supplementary Table 2-1. The ES clone was expanded and then injected into C57BL/6J blastocysts at the University of Michigan Transgenic Mouse Core. Germ-line transmission was achieved by mating male chimeric founders with C57BL/6J albino female mice (B6(Cg)-*Tyrc-2J*/J; stock no. 000058, The Jackson Laboratory). The resulting germ-line transmitted *Lman1*^{cgt} allele was continuously backcrossed and maintained on a C57BL/6J genetic background, as were all subsequent distinct alleles derived from *Lman1*^{cgt}. Primer trio LMAN1 C/D/E was used to distinguish between the

wild-type *Lman1*$^{+}$ allele and the targeted conditional *Lman1*^{cgt} allele (Figure 2-1B). Primer set LMAN1 A/B also distinguishes between the *Lman1*$^{+}$ allele and the *Lman1*^{cgt} allele (Figure 2-1C). Genotyping for all animals reported in this study was performed with mouse tail clip DNA using Go-Taq Green Master Mix (Promega) and the relevant primers listed in Supplementary Table 2-1, with the resulting PCR products resolved by 2% agarose gel electrophoresis. The expected PCR amplicon sizes for all *Lman1* alleles generated with primer sets C/D/E and A/B/C are listed in Table 2-1.

Transgenic mice (C57BL/6J background) carrying *Flpe* (flippase) recombinase driven by an actin promoter were obtained from The Jackson Laboratory (B6;SJL-Tg(ACTFLPe)9205Dym/J; stock no. 003800). Intercrosses of *Lman1*$^{cgt/+}$ mice and *Flpe* recombinase transgenic mice generated the *Lman1*^{fl} allele (also known as *Lman1*$^{tm1c(KOMP)Wsti}$) (Figure 2-1A). Transgenic *Tek-Cre*$^{+}$ *(Tie2-Cre)* recombinase mice (B6.Cg-Tg(Tek-cre)12Flv/J; stock no. 004128) and *Alb-Cre*$^{+}$ recombinase mice (B6.Cg-Tg(Alb-cre)21Mgn/J; stock no. 003574) were obtained from The Jackson Laboratory. Primer trio LMAN1 A/B/C was used to distinguish between the *Lman1*$^{+}$ allele and the *Lman1*^{fl} and *Lman1*$^{-}$ alleles derived from *Lman1*^{cgt} (Figure 2-1C) for the identification of conditional knockout *Lman1* mice, and *Alb-Cre* or *Tek-Cre* primers (listed in Supplementary Table 2-1) were used to determine the *Cre* genotype. The *Alb-Cre* Fwd/Rev primer pair generates a single 207 bp PCR amplicon. The *Tek-Cre* multiplex primer trio generates an internal control 200 bp PCR product from the endogenous *Tek* gene, and a 579 bp PCR product from the *Tek-Cre* transgene. *Tek-Cre* driven recombinase activity results in the deletion of *loxP* flanked target sequences in endothelial cells (of all tissues) and hematopoietic cells, but also to a lesser extent in the germline. Only male *Lman1*$^{fl/+}$, *Tek-Cre*$^{+}$ mice were crossed with female *Lman1*$^{fl/fl}$ mice to generate *Lman1*$^{fl/fl}$, *Tek-Cre*$^{+}$ offspring, since the *Tek-Cre* recombinase activity occurs at a lower frequency in the male germline than in the female germline.[107]

All animals were housed according to the guidelines of the University of Michigan Unit of Laboratory Animal Medicine. The University of Michigan's University Committee on

the Use and Care of Animals approved all animal protocols in this study under protocol number 08571.

Western blot antibodies

Antibodies used in Western blot experiments were: a monoclonal antibody against human β-actin (Santa Cruz Biotechnology), a monoclonal antibody against the RalA GTPase (Sigma-Aldrich), a goat-anti-mouse IgG HRP antibody, a polyclonal antibody against LMAN1 (Sigma-Aldrich), and a polyclonal rabbit antibody against the heavy chain of murine FV.[108] FV antigen levels in plasma and platelets were assessed by Western blot analysis using the FV antibody.

Measurement of FV and FVIII levels

Blood was collected into 4% sodium citrate solution (9:1) (Sigma-Aldrich) via inferior vena cava puncture essentially as previously described,[109] under pentobarbital anesthesia (Nembutal Sodium Solution) with a 1 mL syringe. Platelet poor plasma was isolated as previously described,[110] snap frozen, and stored at -80°C. Platelet rich plasma and platelet pellets were isolated and processed as follows: 2 mL of room temperature Buffered Saline Glucose Citrate (BSGC) (129 mM NaCl, 13.6 mM Na_3 citrate, 11.1 mM glucose, 1.6 mM KH_2PO_4, 8.6 mM NaH_2PO_4, pH 7.3) were placed in a 5 mL polypropylene tube, to which was added 1-1.5 mL of whole blood. BSGC was added to a final volume of 4 mL, and then gently mixed by inversion. The tubes were centrifuged at 180 x g for 10 minutes at room temperature. The supernatant (semi-platelet rich plasma) was removed and subjected to a platelet count on an Advia 120 Automated Hematology Analyzer (Siemens), and the number of platelets was normalized between samples before centrifuging in fresh tubes at 700 x g for 10 minutes. Platelet pellets were lysed with Triton-X 100 on ice (final concentration 0.2%) for 30 minutes and frozen as previously described.[95] Plasma and platelet samples from previously described *F5*$^{-/-}$ mice carrying either a platelet- or plasma- specific *F5* transgene,[49] respectively, were used as experimental controls.

Plasma FV and FVIII activity levels were determined in prothrombin time (PT) and partial thromboplastin time (PTT) based assays, respectively. Individual mouse plasma samples were mixed with human plasma deficient for the respective factor (George King) and analyzed on a KC-4 Coagulation Analyzer (Sigma-Aldrich). A pooled C57BL/6J plasma stock was diluted into human FV or FVIII deficient plasma (George King) to generate a reference standard curve.

RiboTag mice

RiboTag mice[111] (B6.129-*Rpl22*$^{tm1.1Psam}$/J) were a kind gift from Dr. Paul Amieux (University of Washington) and were also obtained from The Jackson Laboratory (stock no. 011029). Genotyping primers for the RiboTag mice are listed in Supplementary Table 2-1. RiboTag mice (*Rpl22*$^{HA/HA}$) in this study are homozygous for a modified *Rpl22* gene (ribosomal protein L22) that contains both a wild-type exon 4 flanked by *loxP* sites, as well as an alternate hemagglutinin (HA)-tagged exon 4 sequence.

RNA isolation

Mice were anesthetized followed by intracardiac perfusion with 12 mL of 100 ug/mL cyclohexamide (Sigma) in PBS, after which tissues were isolated and snap frozen in liquid nitrogen. Immunoprecipitation of endothelial cell-specific polysomes (and bound mRNA) from whole liver, kidney, brain, and heart lysates was performed essentially as previously described, with minor modifications.[111] Frozen tissues were pulverized and the resulting powders were transferred to tubes containing polysome buffer (50 mM Tris, pH 7.5, 100 mM KCl, 12 mM MgCl2, 1% Nonidet P-40, 1 mM DTT, 200 U/mL Promega RNasin, 1 mg/mL heparin, 100 ug/mL cycloheximide, Sigma protease inhibitor mixture). Cleared lysate was incubated with a purified mouse monoclonal antibody against the hemagglutinin tag (HA.11 clone 16B12, Covance, 3 ug per sample) for 1 hour with rotation at 4°C. Protein G magnetic beads (New England Biolabs) equilibrated in polysome buffer were added for an additional 30-minute rotated incubation at 4°C. The beads were washed in high salt buffer (50 mM Tris, pH 7.5, 300 mM KCl, 12 mM MgCl2, 1% Nonidet P-40, 1 mM DTT, 100 ug/mL cycloheximide), and ribosome-

associated mRNA was subsequently eluted from the beads by adding Qiagen buffer RLT supplemented with 2-mercapto-ethanol (1% v/v). Total RNA was isolated from the whole tissue lysates (input) and immunoprecipitated polyribosome samples with the RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol, and DNase treatment was performed on column (Qiagen). RNA samples were quantified on a spectrophotometer (Nanodrop) and the integrity of the RNA was assessed by gel electrophoresis. Total RNA from whole blood of C57BL/6J mice was extracted with Trizol LS and purified on RNeasy spin columns (Qiagen).

cDNA conversion and qPCR analysis

Total mRNA (100 ng) from input lysates and HA-tagged polysome samples was converted to cDNA with the Superscript First-Strand system (Invitrogen) and oligoDT primers. Gene-specific primers (Supplementary Table 2-1) were designed with Primer Express Software (Applied Biosystems) and qRT-PCR was executed with SYBR-Green RT-PCR Master Mix (Applied Biosystems) on a 7900HT Fast Real-Time PCR machine (Applied Biosystems). Data were analyzed using the comparative threshold cycle method[112] with *Rpl37* (Ribosomal protein L37), *Rpl38* (Ribosomal protein L38), and *Eif3e* (eukaryotic translation initiation factor 3, subunit E) as internal controls,[113] and the ΔCt value of the bead elution was related to the ΔCt value (Ct value of sequence of interest – Ct value of reference sequence) of its respective total tissue lysate to determine the fold-change. One microgram of mRNA from whole blood was used for subsequent cDNA conversion with oligoDT primers.

Statistical analysis

A chi-squared test was used to evaluate for statistical deviation from expected Mendelian ratios for the genotypes of offspring from matings between $Lman1^{fl/fl}$ mice and $Lman1^{fl/+}$, Cre^{+} mice. Analysis of the FV and FVIII activity data was performed using the Student's t-test comparing pooled levels from $Lman1^{fl/fl}$ and $Lman1^{fl/+}$ (control) mice relative to levels from the other experimental genotypes ($Lman1^{fl/fl}$, Cre^{+} and $Lman1^{fl/+}$, Cre^{+}). To evaluate the qPCR data with respect to enrichment or depletion in the endothelial cell-

fraction relative to the whole tissue lysate, a paired t-test was performed. A Student's t-test was used to analyze the qPCR data for whole blood relative to whole tissue lysates. A p-value < 0.05 was considered to be statistically significant for all analyses.

RESULTS

FVIII is synthesized in murine endothelial cells

Long-range PCR confirmed the correct genomic targeting of the *Lman1*^{cgt} KOMP targeting construct (Supplementary Figure 2-1B). *Lman1*$^{cgt/+}$ mice were crossed to *Flpe* recombinase transgenic mice to excise the gene-trap cassette, generating the *Lman1*^{fl} allele in which exons 2-3 are flanked by *loxP* sites, but no other gene-trap sequence is retained (Figure 2-1A). The *Lman1*^{fl} allele may be converted to the null *Lman1*$^{-}$ allele in a conditional manner by tissue-specific *Cre*-mediated excision of exons 2 and 3. Heterozygous *Lman1*$^{fl/+}$ mice were crossed with male transgenic *Tek-Cre*$^{+}$ (*Tie2-Cre*) mice that carry the *Cre* recombinase gene driven by the *Tek* promoter to generate *Lman1*$^{fl/+}$, *Tek-Cre*$^{+}$ mice with a single *Tek-Cre* allele. To delete *Lman1* in endothelial cells of all tissues, *Lman1*$^{fl/+}$, *Tek-Cre*$^{+}$ mice were subsequently crossed to *Lman1*$^{fl/fl}$ mice to generate *Lman1*$^{fl/fl}$, *Tek-Cre*$^{+}$ endothelium-specific knockout mice. The same general mating scheme was used to generate hepatocyte-specific *Lman1* knockout mice by utilizing an *Alb-Cre* recombinase. *Lman1*$^{fl/fl}$, *Tek-Cre*$^{+}$ mice (endothelial-knockout) and *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ (hepatocyte-knockout) mice are viable and were observed in the expected Mendelian ratios (Table 2-2). As demonstrated by PCR analysis of genomic DNA isolated from an *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ mouse tail snip and liver tissue, the conversion of the *Lman1*^{fl} allele to the *Lman1*$^{-}$ null allele by genetic excision of *Lman1* exons 2 and 3 was very efficient and specific to the liver (Figure 2-1C). Similarly, Western blot analysis of LMAN1 levels in multiple tissues harvested from wild-type C57BL/6J mice and *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ mice confirmed efficient loss of LMAN1 expression in the liver of *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ mice (Figure 2-1D). The equivalent levels of LMAN1 in other tissues of *Lman1*$^{fl/fl}$ mice compared to wild-type mice confirmed

normal LMAN1 expression from the *Lman1*^{fl} allele in the absence of *Cre* recombinase (Figure 2-1D). FV and FVIII activity levels for both *Lman1*$^{fl/+}$, *Tek-Cre*$^{+}$ and *Lman1*$^{fl/+}$, *Alb-Cre*$^{+}$ mice (heterozygous conditional knockout mice) were indistinguishable from control animals (Figure 2-2A and Figure 2-2B), consistent with the autosomal recessive inheritance of F5F8D, and with previous results in *Lman1*$^{gt1/+}$ mice.[95] *Lman1*$^{fl/fl}$ and *Lman1*$^{fl/+}$ mice were pooled together as controls for subsequent experiments.

As expected from the known synthesis of plasma FV in hepatocytes,[49] *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ mice demonstrate significant reduction of plasma FV activity relative to *Lman1*$^{fl/fl}$ and *Lman1*$^{fl/+}$ control mice (48.3% vs. 96.8%, $p < 1.5 \times 10^{-10}$) (Figure 2-2A), with consistent reductions in plasma FV antigen level also observed by Western blotting (Figure 2-3A). Plasma FV activity in *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ mice was indistinguishable from *Lman1*$^{gt1/gt1}$ mice ($p > 0.5$, Figure 2-2A), suggesting highly efficient excision in hepatocytes, as directed by the *Alb-Cre* transgene. In contrast, plasma FVIII levels in *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ mice were indistinguishable from controls (90.2% vs. 97.5%, $p > 0.12$) (Figure 2-2B). These data demonstrate that hepatocyte synthesis does not contribute significantly to the plasma FVIII pool in the mouse. The *Lman1*$^{fl/fl}$, *Tek-Cre*$^{+}$ mice exhibit normal plasma FV activity relative to *Lman1*$^{fl/fl}$ and *Lman1*$^{fl/+}$ control mice (99.8% vs. 96.8%, $p > 0.5$) (Figure 2-2A). In contrast, FVIII activity levels in *Lman1*$^{fl/fl}$, *Tek-Cre*$^{+}$ mice were markedly reduced (52.8% vs. 97.5%, $p < 1.7 \times 10^{-8}$) and were indistinguishable from the levels previously reported in *Lman1*$^{gt1/gt1}$ null mice (52.8% vs. 47.9%, $p > 0.30$) (Figure 2-2B). These results suggest that endothelial cells are the primary biosynthetic source of FVIII *in vivo*.

The murine plasma and platelet FV pools are biosynthetically distinct and are derived from synthesis in hepatocytes and megakaryocytes, respectively.[114] Consistent with these previous reports, platelet FV antigen, as assessed by Western blotting, appears roughly equivalent to control mice in *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ mice, with a modest reduction in *Lman1*$^{fl/fl}$, *Tek-Cre*$^{+}$ mice, (Figure 2-3B), consistent with the known expression of *Tek-Cre* in megakaryocytes.[115]

***F8* mRNA is localized to endothelial cells and not hepatocytes**

The previously reported RiboTag mouse[111] was used to isolate endothelial cell-specific mRNA for qPCR analysis from a number of different murine tissues. In the absence of *Cre*-recombinase, RiboTag mice ubiquitously express the wild-type RPL22 ribosomal protein in all cell types. However, in the presence of *Cre*-recombinase, RPL22 is fused to a hemagglutinin-tag, which can be used to selectively immunoprecipitate polyribosomes derived only from cells expressing *Cre*.[111] Mice homozygous for the RiboTag allele and hemizygous for the *Tek-Cre* transgene were generated, which should result in expression of the RPL22-HA fusion protein only in endothelial (and hematopoietic) cells.

qPCR analyses of endothelial cell mRNA isolated from unfractionated liver lysates from RiboTag, *Tek-Cre*$^{+}$ mice (n =4) (Figure 2-4A; Supplementary Table 2-2) demonstrated depletion (5-10 fold) of transcripts from several known hepatocyte-specific genes, including *Albumin (Alb)*, *prothrombin (F2)*, *coagulation factor V (F5)*, *coagulation factor VII (F7)*, and *coagulation factor X (F10)*. In contrast, a 15-22 fold enrichment was observed for gene transcripts that are known to be endothelial-specific, including *Cadherin 5 (Cdh5)*, *TEK tyrosine kinase (Tek)*, *Vascular cell adhesion molecule 1 (Vcam1)*, and *von Willebrand factor (Vwf)*.

Similarly, endothelial cell mRNA from kidney (Figure 2-4B; Supplementary Table 2-2), brain (Figure 2-4C; Supplementary Table 2-2), and heart (Figure 2-4D; Supplementary Table 2-2) prepared from the same four RiboTag, *Tek-Cre*$^{+}$ mice showed 10-50 fold depletion of corresponding tissue specific non-endothelial cell genes, and 4-20 fold enrichment of endothelial cell-specific genes (*Cdh5*, *Tek*, *Vcam1*, and *Vwf*).

qPCR analysis of *F8* mRNA levels revealed an unexpected variation in the level of *F8* expression across these vascular beds. *F8* mRNA transcripts were highly enriched (relative to total tissue lysate) in liver endothelial cells (16.5 fold, $p < 2.2 \times 10^{-4}$), comparable to that of known endothelial-specific genes. *F8* transcripts were also enriched in kidney endothelial cells (7.4 fold, $p < 1.6 \times 10^{-3}$). However, no *F8* mRNA

enrichment was observed in the endothelial cells of brain (0.25 fold, $p > 0.1$) or heart (0.67 fold, $p > 0.15$).

Circulating peripheral blood cells do not express *F8* mRNA

Since *Tek-Cre* is also expressed in hematopoietic cells,[115] qRT-PCR was performed on whole blood to exclude contaminating blood as the source of the enriched *F8* mRNA in RiboTag, *Tek-Cre*$^+$ tissues. *F8* mRNA in the blood was detected at < 2% of the levels in liver and kidney, with normalization to internal control genes *Rpl37*, *Rpl38*, and *Eif3e* (Table 2-3). Thus, peripheral blood cell mRNA can be excluded as the source of the *F8* mRNA signal in RiboTag, *Tek-Cre*$^+$ mouse tissues. The enrichment of mRNA transcripts for *Vwf* and *F5* in whole blood mRNA is consistent with the known expression patterns of these genes in murine megakaryocytes/platelets.[49,116,117]

DISCUSSION

The *in vivo* biosynthetic origin of FVIII has been controversial for decades.[52-57,59-63,100-105] We took advantage of the dependence of efficient FV and FVIII secretion on LMAN1 to examine the cellular biosynthetic source of each protein *in vivo*. While both the human plasma and platelet FV pools are produced by hepatocytes,[47] murine plasma FV is synthesized exclusively in hepatocytes, and murine platelet FV in megakaryocytes.[49,114] In the absence of LMAN1, FV levels are reduced in both plasma and platelets in humans and mice.[4,41,95]

Consistent with these previous data, we observed that mice with *Lman1* specifically deleted in the hepatocytes exhibit low plasma FV and normal platelet FV, while *Lman1* deletion in the hematopoietic compartment results in low platelet FV and normal plasma FV. Our finding of normal FVIII in mice with hepatocyte-specific knockout of *Lman1* excludes the hepatocyte as a major source of FVIII production *in vivo*. Conversely, the observation that endothelial deletion of *Lman1* results in reduced FVIII levels that are

comparable to the levels in LMAN1 null mice (Figure 2-2B) indicates that high level, tissue-specific *Lman1* excision was achieved with the *Tek-Cre* recombinase, and suggests that the endothelial cell is the primary source of FVIII biosynthesis *in vivo*.

The *Tek-Cre* used for specific targeting to the endothelium is known to be expressed both in the endothelial and hematopoietic compartments.[118,119] The absence of *F8* mRNA enrichment in peripheral blood of wild-type C57BL/6J mice excludes trapped peripheral blood cells as the source for the FVIII endothelial cell signal, as does the absence of *F8* mRNA enrichment in the brain or heart (Figure 4). The lack of *F8* mRNA in the blood is consistent with a recent clinical report of allogeneic bone marrow transplantation (BMT) in a child with severe aplastic anemia and hemophilia A. The child's FVIII levels and coagulation test results remained essentially unchanged from his pre-transplantation values when measured four months after BMT,[120] indicating that the bone marrow is not an important source of FVIII synthesis *in vivo*.

Examination of affinity-selected endothelial cell mRNA from actively translating ribosomes isolated from multiple tissues confirmed endothelial FVIII biosynthesis *in vivo*. However, the levels of *F8* mRNA observed in the endothelial cells of liver, kidney, brain, and heart tissues revealed surprising variation across these vascular beds, with the highest levels of expression in liver endothelial cells, intermediate levels in kidney, and no detectable enrichment of *F8* mRNA transcripts in brain or heart endothelial cells.

Taken together, these data demonstrate that endothelial cells from multiple (but not all) tissues and vascular beds contribute to the plasma FVIII pool in the mouse, with a large contribution from hepatic endothelial cells. These results account for the successful reversal of hemophilia A by liver transplantation, and presumably also by lung and spleen transplantation.[52-54,59,60] These results are also consistent with the maintenance of normal FVIII levels in patients with parenchymal liver disease in whom the endothelium may still be intact.[121-123] The variable levels of endothelial cell mRNA enrichment across tissues likely reflects differences in mRNA expression profiles and the relative ratio of endothelial to parenchymal cells in different vascular beds.[124-126] Though the functional

significance of this heterogeneity in FVIII endothelial production is unknown, it could reflect variation in local hemostatic demand across different vascular beds.

Our data are also consistent with several recent publications reporting the relative contributions of hepatocytes and endothelial cells to the plasma FVIII pool. Transplantation of unfractionated liver cells (a mixture of hepatocytes, LSECs, Kupffer cells, and hepatic stellate cells) into the peritoneal cavity of hemophilia A mice corrected the disease, but transplantation of a hepatocyte-enriched fraction alone had no effect.[127] Similarly, transplanted normal endothelial cells are able to repopulate the liver endothelium and correct the bleeding phenotype of hemophilia A mice.[128] In addition, FVIII activity levels were found to be 10- to 100-fold higher in human LSECs than in hepatocytes separated by flow cytometry cell sorting.[129] Nonetheless, other studies have detected FVIII in human hepatocytes,[52-55,130] and differences in the pathway for FVIII production between humans and mice cannot be excluded.

A growing body of evidence suggests extensive heterogeneity among endothelial cells *in vivo*.[131] For example, *Vwf* expression varies widely across different endothelial beds, with different tissue micro-environmental factors regulating its transcription.[132,133] Similarly, the expression of dozens to hundreds of genes may vary among different specialized vascular beds.[134] While we observed modest variation in the expression levels of *Vwf* and 3 other endothelial-specific genes in our studies of lung, kidney, heart, and brain endothelial cells, we observed large differences in *F8* gene expression amongst these four tissues. *F8* mRNA transcripts were most highly enriched in hepatic endothelial cells, with intermediate enrichment in kidney endothelial cells and little to no enrichment in the endothelial cells of heart or brain. Our results are consistent with recent transcriptional profiling of tissue-specific endothelial cell populations isolated by flow sorting, where *F8* mRNA transcripts were detected in endothelial cells from liver and glomeruli, but not in the testis, spleen, muscle, brain, lung, heart, or bone marrow.[124] This heterogeneity of endothelial cell gene expression has important implications, not only for FVIII biosynthesis, but also for regulation of other tissue-specific endothelial cell functions.[135,136] The finding that *F8* and *Vwf* are both highly expressed in endothelial

cells likely confers a regulatory mechanism for FVIII biosynthesis, consistent with previous reports that the coexpression of *F8* and *Vwf in vitro* result in increased stable accumulation of FVIII activity.[51] Finally, the establishment of endothelial cells as the primary site of FVIII biosynthesis also has important implications for the treatment of hemophilia A, particularly for gene-targeting approaches directed specifically to endothelial cells.[137-139]

Table 2-1: Expected PCR amplicon size (bp) for KOMP *Lman1* alleles				
Primer Set	*Lman1*$^{+}$	*Lman1*^{cgt}	*Lman1*^{fl}	*Lman1*$^{-}$
C/D/E	434	565	591	(no product)
A/B/C	444	508	508	635

Table 2-2: Generation of *Lman1* conditional knockout mice					
Crosses	Genotype Distribution at 3 weeks				p-value (χ^2)
	Lman1$^{fl/+}$, *Cre*$^{-}$	*Lman1*$^{fl/+,}$ *Cre*$^{+}$	*Lman1*$^{fl/fl}$, *Cre*$^{-}$	*Lman1*$^{fl/fl}$, *Cre*$^{+}$	
Expected %	25%	25%	25%	25%	
Lman1$^{fl/fl}$ x *Lman1*$^{fl/+}$, *Alb-Cre*$^{+}$	24% (17)	25% (18)	19% (14)	32% (23)	0.5 (n.s.)
Lman1$^{fl/fl}$ x *Lman1*$^{fl/+}$, *Tek-Cre*$^{+}$	19% (16)	28% (23)	26.5% (22)	26.5% (22)	0.7 (n.s.)
The chi-squared test (df = 3) was based upon an expected genotype ratio of 3:1, with *Lman1*$^{fl/fl}$/*Cre*$^{+}$ mice expected to represent 25% of the offspring from each mating, and all other genotypes cumulatively accounting for 75% of offspring.					

Table 2-3: Gene expression average fold-change in the whole blood mRNA relative to total liver and kidney lysates (input)

Whole blood vs. liver lysate				
Gene	Expression	Fold-change ± SEM	Range	p-value
Cdh5	Endothelial	0.19 ± 0.02	0.17 – 0.22	$< 3.5 \times 10^{-2}$
Tek	Endothelial	0.04 ± 0.01	0.034 – 0.046	$< 2.0 \times 10^{-3}$
Vcam1	Endothelial	0.001 ± 0.004	0.001 – 0.009	$< 6.0 \times 10^{-3}$
Vwf	Endothelial/Megakaryocyte	113.40 ± 7.09	106.6 – 120.7	$< 3.0 \times 10^{-6}$
F5	Hepatic/Megakaryocyte	1.31 ± 0.06	1.25 – 1.36	> 0.3
F8	?	0.01 ± 0.01	0.007 – 0.012	$< 1.0 \times 10^{-4}$
Whole blood vs. kidney lysate				
Gene	Expression	Fold-change ± SEM	Range	p-value
Cdh5	Endothelial	0.08 ± 0.01	0.074 – 0.095	$< 7.4 \times 10^{-5}$
Tek	Endothelial	0.002 ± 0.001	0.007 – 0.010	$< 2.1 \times 10^{-5}$
Vcam1	Endothelial	0.003 ± 0.004	0.001 – 0.009	$< 3.2 \times 10^{-3}$
Vwf	Endothelial/Megakaryocyte	248.92 ± 15.56	233.2 – 264.9	$< 3.3 \times 10^{-9}$
F5	Hepatic/Megakaryocyte	73.41 ± 3.15	70.3 – 76.6	$< 1.5 \times 10^{-5}$
F8	?	0.02 ± 0.01	0.011 – 0.029	$< 1.5 \times 10^{-4}$

To evaluate the potential contribution of blood contamination to transcript levels of endothelial-specific genes, the tissue lysates of either the kidney or liver samples were set as a reference and the ΔCt values of individual blood samples (n=4) were related to the mean ΔCt of the reference group.

Supplementary Table 2-1: Primer sequences	
Primer	5' → 3' Sequence
Genotyping Primers	
Lman1 primer A	GGCTTTCTTGACACCTTCAATTTAA
Lman1 primer B	CCAAGTGAAGGGAAGACCATCAAGC
Lman1 primer C	GACCCCTAGTGACGGGTTCTTGTC
Lman1 primer D	CACTGAGTCTCTGGCATCTCCCCT
Lman1 primer E	CCCACAAATTAAATGTCAAGGACA
Flpe transgene Fwd	GGTCCAACTGCAGCCCAAGCTTCC
Flpe transgene Rev	GTGGATCGATCCTACCCCTTGCG
Alb-Cre transgene Fwd	GAAAATGCTTCTGTCCGTTTGC
Alb-Cre transgene Rev	ATTGCTGTCACTTGGTCGTGGC
Tek Fwd	CCCTGTGCTCAGACAGAAATGAGA
Tek Rev	TAAGCCGGCTAAAGAGTCCA
Tek-Cre transgene Rev	CGCATAACCAGTGAAACAGCATTG
RiboTag Fwd	GGGAGGCTTGCTGGATATG
RiboTag Rev	TTTCCAGACACAGGCTAAGTACAC
qPCR Primers	
Rpl37 Fwd	CGGGACTGGTCGGATGAG
Rpl37 Rev	TCACGGAATCCATGTCTGAATC
Rpl38 Fwd	CGCCATGCCTCGGAAA
Rpl38 Rev	CCGCCGGGCTGTCAG
Eif3e Fwd	GGTTGGATGCCAAGATTGATTC
Eif3e Rev	GGGCGAGACTGCATTGTTG
Cdh5 Fwd	AGCGCAGCATCGGGTACT
Cdh5 Rev	TCGGAAGAATTGGCCTCTGT
Vcam1 Fwd	GGCAGAGTGTACAGCCTCTTTATG
Vcam1 Rev	GCTGACCCAGATGGTGGTTT
Vwf Fwd	AGCTGTCAGCCAGGTTTTTCTT
Vwf Rev	GCAGAGGGCAGGCACCTT
Tek Fwd	CTATCGGACTCCCTCCTCCAA
Tek Rev	TCAAATTTAGAGCTGTCTGGCTTTT
Alb Fwd	GGACCAGGAAGTGTGCAAGAA
Alb Rev	CAAGAACGTGCCCAGGAAGA
F2 Fwd	GGACGCTGAGAAGGGTATCG
F2 Rev	CCCCACACAGCAGCTCTTG
F5 Fwd	CCAGGTAGCTGGCATGCAA
F5 Rev	CCATTGGCATCTTACACTCTTTGT
F7 Fwd	CTGTCATGAGGACTACACGCTACA
F7 Rev	CACGGGTACTCAACTTTTGGTTT
F8 Fwd	CTTCACCTCCAGGGAAGGACTA
F8 Rev	TCCACTTGCAACCATTGTTTTG
F10 Fwd	GTGGCCGGGAATGCAA
F10 Rev	AACCCTTCATTGTCTTCGTTAATGA

Aqp2 Fwd	CCCTGGCTCCAGCAGTTG
Aqp2 Rev	CGATCCAGAAGACCCAGTGATC
Cdh16 Fwd	TGGAGGCCCAAGACACAGA
Cdh16 Rev	TGGATCACAACAGTGGCAGAA
Umod Fwd	CAGCCTGAAGACCTCCCTACA
Umod Rev	TCCCACCCAAGCTGATGTTC
Penk Fwd	TGCAGCTACCGCCTGGTT
Penk Rev	CAGTGTGCACGCCAGGAA
Tac1 Fwd	GATGCTGATTCCTCAGTTGAAAAA
Tac1 Rev	GATCTGGCCATGTCCATAAAGAG
Fabp3 Fwd	AGTCACTCGGTGTGGGCTTT
Fabp3 Rev	ATGATGGTAGTAGGCTTGGTCATG
Myh6 Fwd	CCCTCAATGACTTCACCACACA
Myh6 Rev	CTAGCCAACTCCCCGTTCTCT
Tnnt2 Fwd	GAAGTTCGACCTGCAGGAAAA
Tnnt2 Rev	TTCGCAGAACGTTGATTTCG
Genes: *Rpl37,* Ribosomal protein L37; *Rpl38*, Ribosomal protein L38; *Eif3e*, Eukaryotic translation initiation factor3, subunit E; *Cdh5*, Cadherin 5, type 2; *Vcam1*, Vascular cell adhesion molecule 1; *Vwf*, von Willebrand factor; *Tek*, TEK tyrosine kinase, endothelial; *Alb*, Albumin; *F2*, Prothrombin; *F5*, Coagulation factor V; *F7*, Coagulation factor VII; *F8*, Coagulation factor VIII; *F10*, Coagulation factor X; *Aqp2*, Aquaporin 2; *Cdh16,* Cadherin 16; *Umod*, Uromodulin; *Penk*, Preproenkephalin; *Tac1*, Tachykinin, precursor 1; *Fabp3*, Fatty acid binding protein 3, muscle and heart; *Myh6*, Myosin, heavy polypeptide 6, cardiac muscle, alpha; *Tnnt2*, Troponin T type 2 (cardiac).	

Supplementary Table 2-2: Gene expression average fold-change in the endothelial cell mRNA relative to total tissue lysate mRNA (input)

LIVER			
Gene	Expression	Fold-change ± SEM	p-value
Alb	Hepatic	0.15 ± 0.041	$< 5.74 \times 10^{-3}$
F2	Hepatic	0.20 ± 0.046	$< 6.00 \times 10^{-3}$
F5	Hepatic	0.19 ± 0.060	$< 1.24 \times 10^{-2}$
F7	Hepatic	0.15 ± 0.042	$< 4.50 \times 10^{-3}$
F10	Hepatic	0.19 ± 0.049	$< 6.75 \times 10^{-3}$
Cdh5	Endothelial	21.69 ± 6.02	$< 2.90 \times 10^{-3}$
Tek	Endothelial	20.46 ± 4.12	$< 8.70 \times 10^{-4}$
Vcam1	Endothelial	21.14 ± 2.22	$< 1.30 \times 10^{-4}$
Vwf	Endothelial	17.73 ± 5.27	$< 2.89 \times 10^{-3}$
F8	?	16.49 ± 2.09	$< 2.22 \times 10^{-4}$
KIDNEY			
Gene	Expression	Fold-change ± SEM	p-value
Aqp2	Kidney	0.066 ± 0.023	$< 4.10 \times 10^{-3}$
Cdh16	Kidney	0.15 ± 0.074	$< 1.79 \times 10^{-2}$
Umod	Kidney	0.068 ± 0.023	$< 1.76 \times 10^{-2}$
Cdh5	Endothelial	20.93 ± 4.40	$< 7.00 \times 10^{-4}$
Tek	Endothelial	22.54 ± 7.03	$< 3.06 \times 10^{-3}$
Vcam1	Endothelial	11.07 ± 2.78	$< 4.79 \times 10^{-3}$
Vwf	Endothelial	9.93 ± 2.74	$< 1.77 \times 10^{-2}$
F8	?	7.40 ± 1.11	$< 1.60 \times 10^{-3}$
BRAIN			
Gene	Expression	Fold-change ± SEM	p-value
Penk	Brain	0.043 ± 0.0034	$< 3.38 \times 10^{-5}$
Tac	Brain	0.035 ± 0.0051	$< 3.12 \times 10^{-4}$
Cdh5	Endothelial	11.77 ± 1.38	$< 2.50 \times 10^{-4}$
Tek	Endothelial	19.39 ± 1.87	$< 7.25 \times 10^{-5}$
Vcam1	Endothelial	4.66 ± 1.04	$< 7.95 \times 10^{-3}$
Vwf	Endothelial	18.41 ± 2.02	$< 1.10 \times 10^{-4}$
F8	?	0.25 ± 0.085	$> 1.10 \times 10^{-1}$
HEART			
Gene	Expression	Fold-change ± SEM	p-value
Fabp3	Heart	0.28 ± 0.012	$< 1.66 \times 10^{-3}$
Myh6	Heart	0.11 ± 0.056	$< 1.20 \times 10^{-2}$
Tnnt2	Heart	0.065 ± 0.017	$< 1.22 \times 10^{-3}$
Cdh5	Endothelial	4.56 ± 0.76	$< 3.20 \times 10^{-3}$
Tek	Endothelial	4.43 ± 0.92	$< 4.60 \times 10^{-3}$
Vcam1	Endothelial	4.79 ± 1.23	$< 6.60 \times 10^{-3}$
Vwf	Endothelial	4.40 ± 1.08	$< 7.30 \times 10^{-3}$
F8	?	0.67 ± 0.22	$> 1.75 \times 10^{-1}$

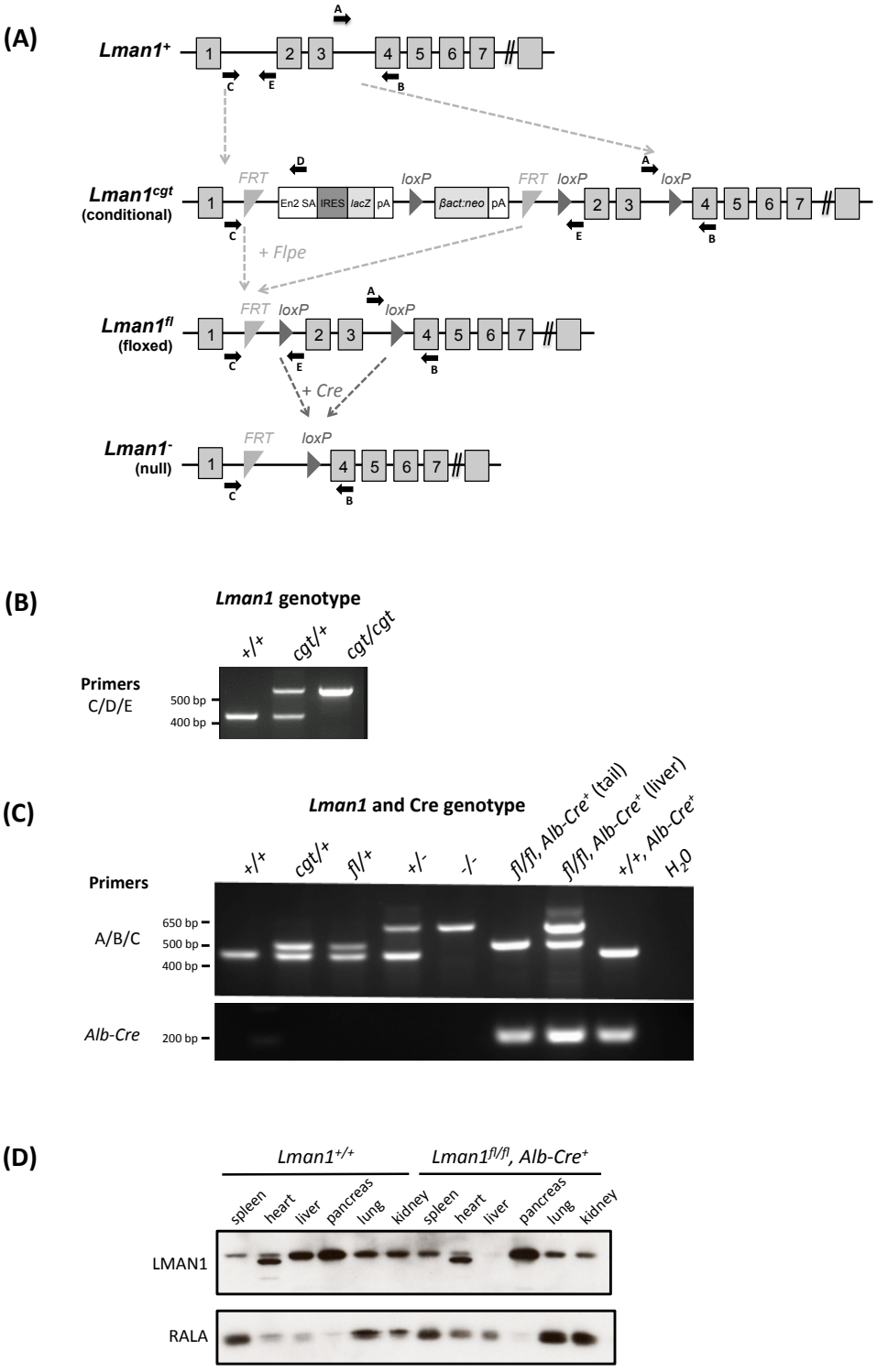


Figure 2-1: *Lman1* mutant alleles

(A) The *Lman1* conditional gene-trap allele (*Lman1*^{cgt}) contains a gene-trap insertion in intron 1 flanked by 2 *FRT* sites. Mice carrying this allele were crossed to β-actin FLP

transgenic mice. Mice heterozygous for the resulting *Lman1* floxed allele (*Lman1*^{fl}) were crossed to *Alb-Cre*$^{+}$ or *Tek-Cre*$^{+}$ transgenic mice to excise exons 2 and 3, generating the *Lman1* null allele (*Lman1*$^{-}$) in selected tissues. Gray blocks represent exons. A, B, C, D, and E represent genotyping primers. (Adapted from the Knockout Mouse Project. General conditional gene targeting scheme: https://www.komp.org/alleles.php#conditional-promoter; *Lman1* targeting: https://www.komp.org/geneinfo.php?geneid=66654); *cgt,* conditional gene trap; *fl*, floxed; En2 SA, splice acceptor of mouse *En2* exon 2; IRES, encephalomyocarditis virus (EMCV) internal ribosomal entry site; *lacZ*, *E.coli* β-galactosidase gene; pA, SV40 polyadenylation signal; β*act:neo*, human β-actin promoter-driven neomycin cassette. **(B)** A three-primer PCR assay (primers C, D, and E) distinguishes the *Lman1*$^{+}$ allele (434 bp) from *Lman1*^{cgt} (565 bp). **(C)** PCR genotyping assay used for tail-snip genomic DNA from a *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ mouse, with the addition of a liver biopsy genomic DNA sample from the *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ mouse to demonstrate excision of *Lman1* exons 2 and 3 in the liver. A three-primer PCR assay (primers A, B, and C) distinguishes the *Lman1*$^{+}$ (444 bp), *Lman1*^{cgt} (508 bp), *Lman1*^{fl} (508 bp), and *Lman1*$^{-}$ (635 bp) alleles. *Alb-Cre* primers were used to determine the *Cre* genotype of offspring from matings of *Lman1*$^{fl/+}$, *Alb-Cre*$^{+}$ mice with *Lman1*$^{fl/fl}$ mice. Comparison of the signal in DNA from liver tissue of an *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ mouse (lane 7) to tail DNA from the same animal (lane 6), indicates a high degree of specific excision in the liver. A similar PCR genotyping strategy of tail-snip genomic DNA was used to identify *Lman1*$^{fl/fl}$, *Tek-Cre*$^{+}$ mice, with primer trio A/B/C demonstrating the *Lman1* genotype and with *Tek-Cre* specific primers rather than *Alb-Cre* primers (not shown). **(D)** Western blot of spleen, heart, liver, pancreas, lung, and kidney tissues from a wild-type C57BL/6J mouse and a *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ (hepatocyte-knockout) mouse, demonstrating nearly-complete and tissue-specific *Lman1* knockout in the liver of *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ mice. The lower molecular weight band observed in the heart samples represents a non-specific band due to cross-reactivity with the LMAN1 antibody, since this same band is observed in the heart tissue of ubiquitous *Lman1*$^{-/-}$ null mice (not shown).

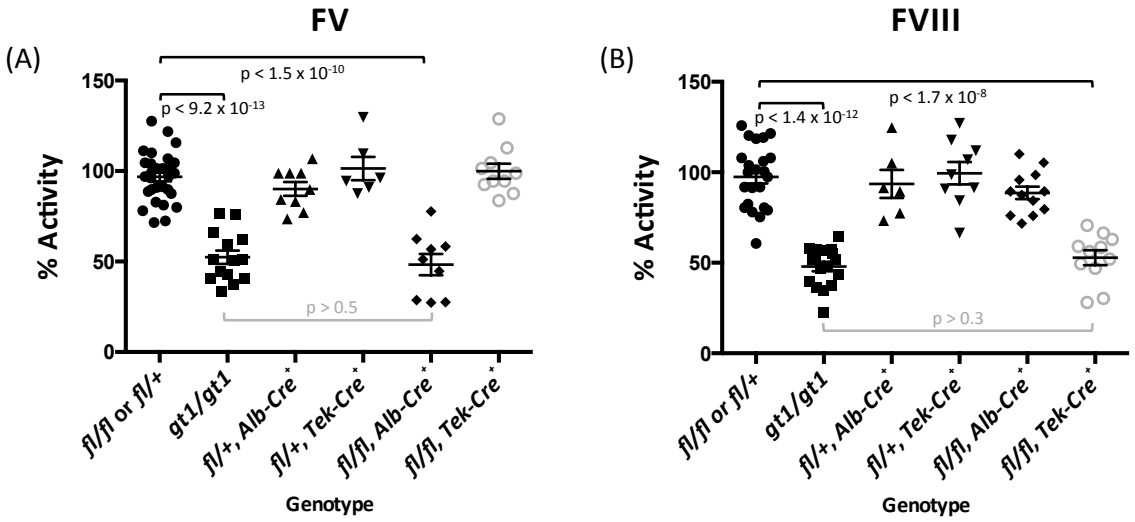


Figure 2-2: FV and FVIII activity levels for *Lman1* conditional knockout mice.

Lman1$^{fl/+}$ and *Lman1*$^{fl/fl}$ mice were pooled as controls. Data for *Lman1*$^{gt1/gt1}$ mice[95] (previously reported) is included for comparison. **(A)** Plasma FV activity measured by PT-based coagulation assay. *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ mice exhibit a statistically significant reduction in plasma FV activity, relative to control mice ($p < 1.5 \times 10^{-10}$), and are indistinguishable from *Lman1*$^{gt1/gt1}$ mice ($p > 0.5$). In contrast, FV activity levels for *Lman1*$^{fl/fl}$, *Tek-Cre*$^{+}$ mice are not reduced. **(B)** Plasma FVIII activity measured by PTT-based coagulation assay. *Lman1*$^{fl/fl}$, *Tek-Cre*$^{+}$ mice exhibit a statistically significant reduction in plasma FVIII activity, relative to control mice ($p < 1.7 \times 10^{-8}$), and are indistinguishable from *Lman1*$^{gt1/gt1}$ mice ($p > 0.3$). FVIII activity levels for *Lman1*$^{fl/fl}$, *Alb-Cre*$^{+}$ mice are not reduced. Each symbol represents an individual animal. Horizontal lines indicate mean and error bars indicate standard error of the mean for each genotype.

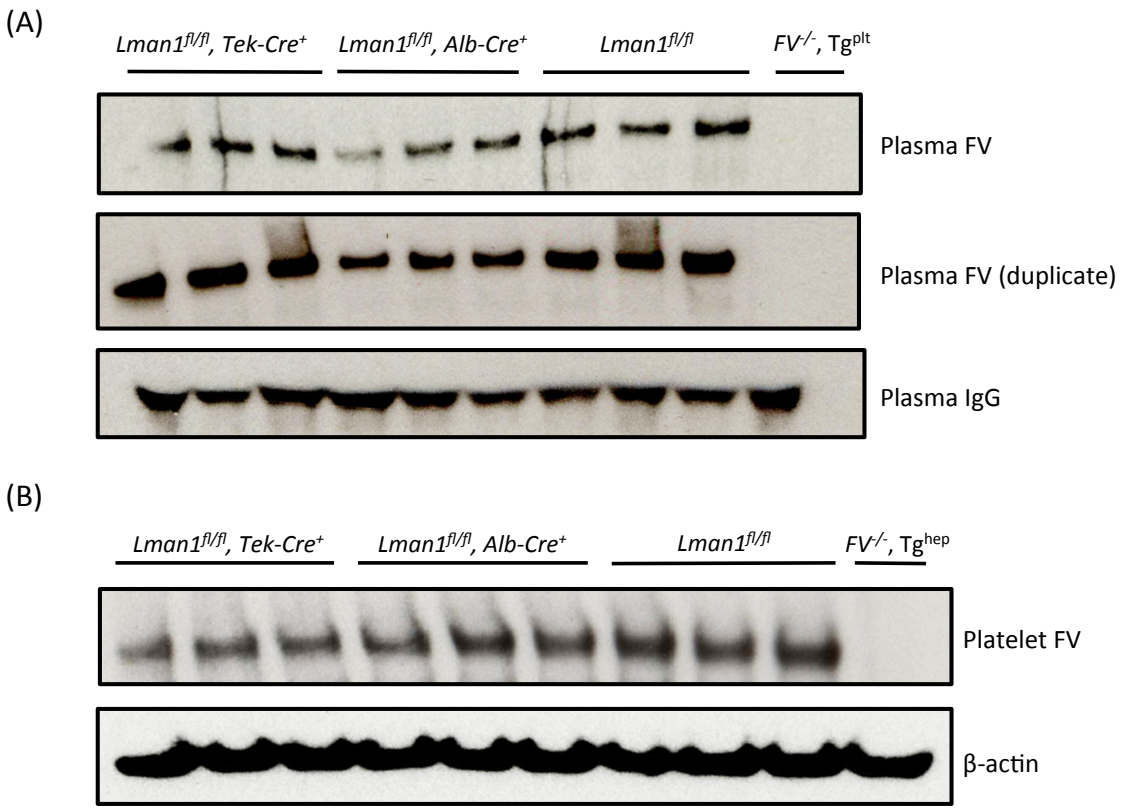


Figure 2-3: Relative FV antigen in plasma and platelets of *Lman1* conditional knockout mice.

(A) Platelet-poor plasma (equal volume) and **(B)** isolated platelets (equal number) from conditional *Lman1* knockout mice and littermate control mice were analyzed for relative FV antigen levels by Western blot. The same plasma samples were run twice as duplicates. Plasma IgG serves as a loading control for the plasma samples, and β-actin serves as a loading control for platelets. Three adult mice were analyzed per genotype. Plasma from $F5^{-/-}$ mice carrying a platelet-specific *F5* transgene[49] was included as a negative control for the plasma samples. Similarly, platelets isolated from a $F5^{-/-}$ mouse carrying a hepatocyte-specific *F5* transgene[49] were included as a negative control for the platelet samples.

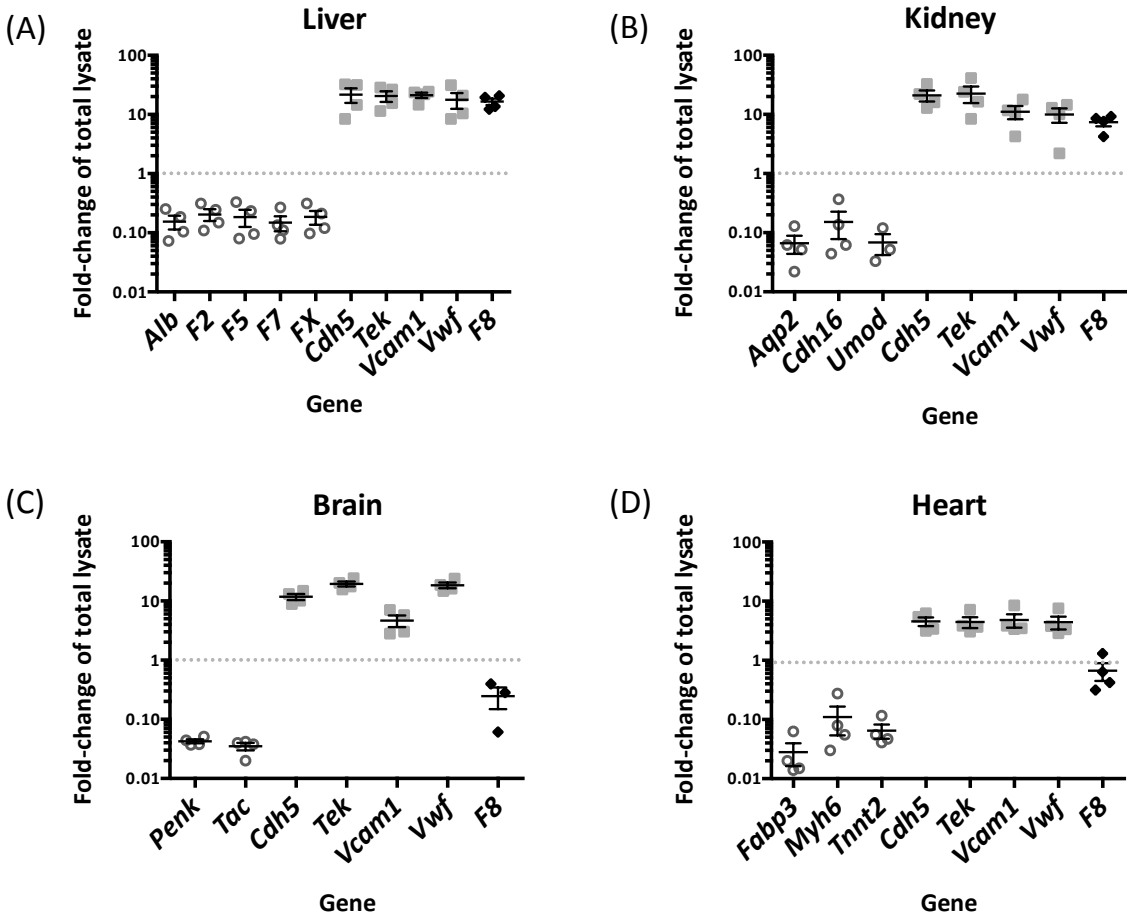
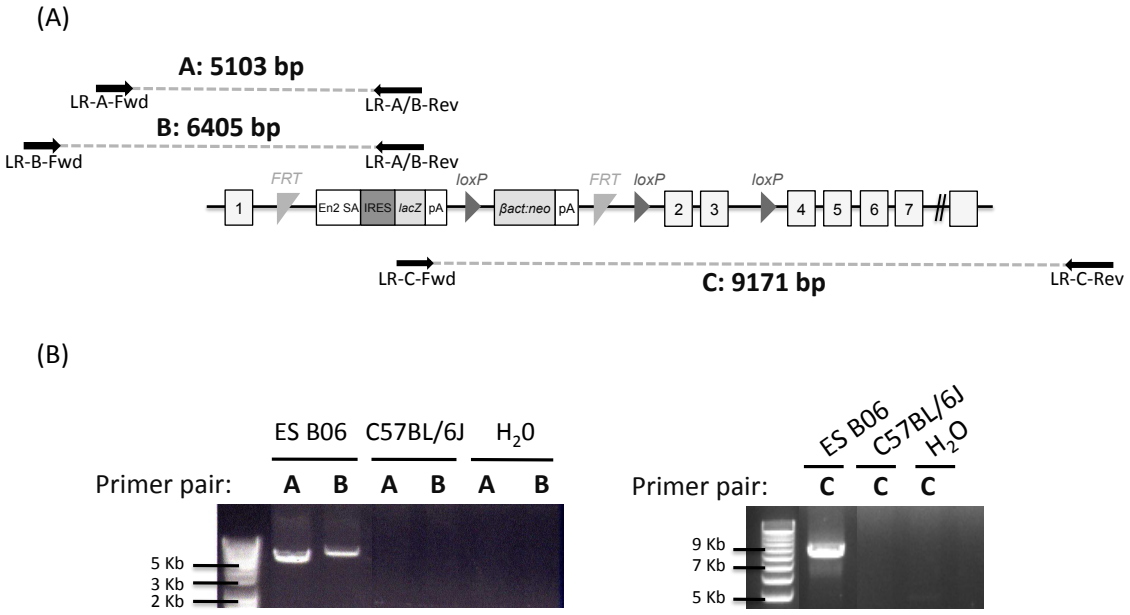


Figure 2-4: *F8* mRNA is expressed in selective endothelial cell beds.

Plots show fold-change (enrichment or depletion) of mRNA transcripts for a number of genes in the endothelial cell mRNA isolated from **(A)** liver, **(B)** kidney, **(C)** brain, and **(D)** heart tissues of four RiboTag, *Tek-Cre*$^{+}$ mice, relative to the whole tissue lysate input (set to 1.0). Values less than 1.0 indicate depletion in the endothelial cell fraction, relative to input. In contrast, values greater than 1.0 indicate enrichment. Each symbol represents an independent animal. Values for tissue-specific, non-endothelial cell genes are shown as open circles. Values for endothelial cell-specific genes are shown as light grey squares. Values for *F8* mRNA fold-change are shown as black diamonds. Horizontal lines indicate mean and error bars indicate standard error of the mean for each genotype.



Supplementary Figure 2-1: Confirmation of *Lman1*^{cgt} KOMP allele genomic targeting by long-range PCR

Long-range PCR was used to confirm correct genomic targeting of the *Lman1*^{cgt} allele in ES cell clone B06 on the 5' and 3' ends of the targeting construct. **(A)** Two independent forward primers were used to confirm the 5' targeting of the KOMP allele (LR-A-Fwd and LR-B-Fwd) in combination with a common reverse primer (LR-A/B-Rev). The expected PCR amplicon sizes for these two primer pairs are 5,103 bp and 6,405 bp, respectively. To confirm targeting on the 3' end of the construct, a forward primer in the gene-trap cassette (LR-C-Fwd) and a reverse primer downstream of the *Lman1* gene (LR-C-Rev) were used, and the resulting PCR amplicon is expected to be 9,171 bp. **(B)** All three long-range primer pairs (A, B, and C) generated a single specific band of the expected nucleotide length when amplified with genomic DNA from the ES B06 clone. As expected, no PCR product was generated when negative control wild-type C57BL/6J genomic DNA was used as the PCR template.

CHAPTER 3 :
CHARACTERIZATION OF LMAN1-DEFICIENT MICE

*Note: Chapter 3 is in preparation for submission to **Blood** as a regular article in March, 2014 (Authors: Lesley Everett, Rami Khoriaty, and David Ginsburg). Lesley Everett performed all experiments and prepared samples for imaging and histologic analyses. Rami Khoriaty assisted with collection and analysis of blood and bone marrow samples. Pathology analysis was performed by the University of Michigan Pathology Department and the ULAM Pathology Cores for Animal Research.*

INTRODUCTION

LMAN1 is a type-1 transmembrane receptor protein that cycles between the Endoplasmic Reticulum (ER) and the Endoplasmic Reticulum Golgi Intermediate Compartment (ERGIC) in the early secretory pathway. LMAN1 (also known as ERGIC-53) has an ER luminal carbohydrate recognition domain for the binding of high mannose glycans,[140,141] and it has a 12 amino-acid cytosolic domain for interactions with the COPII coat. LMAN1 binds in a 1:1 ratio with the soluble, 16 kDa protein MCFD2, which localizes to the ER and ERGIC due to its interactions with LMAN1. Together, these proteins form the only known mammalian cargo receptor complex in the early secretory pathway.

The LMAN1/MCFD2 receptor is required for efficient secretion of coagulation factor V (FV) and factor VIII (FVIII) to the plasma. Mutations in either *LMAN1* or *MCFD2* cause combined deficiency of coagulation factors V and VIII (F5F8D), an autosomal recessive bleeding disorder characterized by reduction of both FV and FVIII to ~10-15% of normal levels.[3,4] LMAN1 and MCFD2 are ubiquitously expressed and highly evolutionarily conserved. LMAN1 orthologs are present in multiple vertebrate and invertebrate species (including *C. elegans* and *D. melanogaster*), prior to the evolutionary establishment of a

coagulation system, suggesting that the LMAN1/MCFD2 cargo receptor plays a broader role in protein secretion. It is therefore surprising that only a small number of LMAN1-dependent secretory cargos are currently known. Other previously reported, but poorly characterized, LMAN1 secretory cargos include alpha-1 antitrypsin (A1AT),[85] cathepsin C (CTSC),[87] cathepsin Z (CTSZ),[88] and Mac-2 binding protein (M2BP).[89]

LMAN1 belongs to a family of homologous leguminous type (L-type) mannose-binding lectins that function in the trafficking of certain *N*-linked glycoproteins in the mammalian secretory pathway.[19-21] Other members of this lectin family include VIP36,[22] VIPL,[23,24] and ERGL.[25] The L-type lectins share similar carbohydrate recognition domains, but they have different subcellular localization patterns and distinct glycoprotein-binding specificities.[26] LMAN1 is transported from the ER to the ERGIC via COPII coated vesicles, and it is recycled back to the ER via COPI vesicles. The identification of *LMAN1* as the disease-causing gene in F5F8D provided the first direct evidence that LMAN1 acts as a mammalian ER-to-Golgi cargo receptor for FV and FVIII.[4] ER cargo receptors are transmembrane proteins that specifically recruit secretory proteins into budding COPII vesicles at the ER membrane. They contain ER exit signals that are recognized by the SEC24 component of COPII coats.[65,142] Only a small number of ER cargo receptors have been characterized in detail, most in yeast, and presumably many more mammalian ER cargo receptors with critical roles in selective, receptor-mediated protein secretion remain to be discovered.[67]

Analyses in human F5F8D patients with defects in *LMAN1* reveal that most patients carry two null *LMAN1* mutations predicted to result in complete absence of the LMAN1 protein, as opposed to missense mutations that could potentially confer a less damaging effect on protein function. This suggests that even low levels of residual LMAN1 function are sufficient to support normal FV and FVIII secretion levels.

We previously described *Lman1* mutant mice carrying a gene-trap insertion in intron 10 ($Lman1^{gt1}$).[95] $Lman1^{gt1/gt1}$ mice have ~50% of normal FV and FVIII activity level, relative to wild-type controls, and are therefore less severely affected than human F5F8D

patients. These mice also exhibit retention of A1AT in the hepatocyte ER, but normal steady state plasma A1AT levels and no significant induction of ER stress or unfolded protein response (UPR) genes.[95] The less severe reductions of FV and FVIII in *Lman1*$^{gt1/gt1}$ mice relative to that of F5F8D patients could be explained by low levels of residual LMAN1 from the gene-trap allele, or differences in FV and FVIII secretion between mice and humans. *Lman1*$^{gt1/gt1}$ mice also exhibit a partially penetrant, perinatal lethality that is restricted to certain genetic strain backgrounds.[95]

We now report analysis of a second, independent LMAN1-deficient mouse with no *Lman1* gene expression (*Lman1*$^{-/-}$), as well as hypomorphic LMAN1-deficient mice that retain ~7% of normal *Lman1* gene expression levels. The null *Lman1*$^{-/-}$ mice exhibit FV and FVIII plasma activity levels that are ~50% of wild-type levels (comparable to *Lman1*$^{gt1/gt1}$ mice). In contrast, hypomorphic LMAN1-deficient mice (*Lman1*$^{cgt/cgt}$) that retain a low level of wild-type LMAN1 expression exhibit intermediate reductions in FV and FVIII. We confirm the previously reported strain-specific, perinatal lethality of LMAN1-deficient mice, and we describe the previously unreported phenotype of moderate thrombocytopenia in *Lman1*$^{-/-}$ mice.

MATERIALS AND METHODS

Lman1 and *Mcfd2* mutant mice

Lman1^{gt1} mice were described previously[95] and are maintained on a C57BL/6J genetic background. The *Lman1*^{gt1} allele carries a gene-trap insertion in *Lman1* intron 10 (Figure 3-1A). PCR genotyping for the *Lman1*^{gt1} allele was performed as previously described.[95]

Lman1^{cgt} (Figure 3-1B) and *Lman1*^{fl} (see Chapter 2) mice were previously described (Everett *et al.*, manuscript in press; Chapter 2) and are maintained on a C57BL/6J genetic background. PCR genotyping for the *Lman1*^{cgt} and the *Lman1*^{fl} alleles was performed as previously described (Everett *et al.*, manuscript in press; Chapter 2). The *Lman1*^{fl} allele

may be converted to the null *Lman1*$^{-}$ allele (*Lman1*$^{tm1d(KOMP)Wtsi}$) (Figure 3-1C) by *Cre*-mediated excision of exons 2 and 3. Transgenic C57BL/6J mice carrying *Cre* recombinase driven by a ubiquitous *EIIA* promoter (B6.FVB-Tg(EIIa-cre)C5379Lmgd/J; stock no. 003724) were obtained from The Jackson Laboratory. *Lman1*$^{fl/+}$ mice were mated to *EIIA-Cre*$^{+}$ transgenic mice to generate *Lman1*$^{+/-}$ mice, which were subsequently backcrossed to C57BL/6J mice to achieve germline transmission of the *Lman1*$^{-}$ allele. The resulting *Lman1*$^{+/-}$ mice were intercrossed to generate *Lman1*$^{-/-}$ null mice. 129S1/SvImJ mice (stock no. 002448) were obtained from The Jackson Laboratory. *Lman1*$^{+/-}$ (C57BL/6J, 129S1/SvImJ) F1 mice (mixed genetic background designated B6;129) were generated and subsequently intercrossed to generate F2 mice. Primer trio A/B/C was used to distinguish between the *Lman1*$^{+}$, *Lman1*^{fl}, and *Lman1*$^{-}$ alleles (Figure 3-2).

An MCFD2 null mouse was previously generated by deleting exons 2 and 3 of the *Mcfd2* gene (Zhang *et al.*, manuscript in preparation). *Mcfd2*$^{-/-}$ mice exhibit FV and FVIII levels that are ~45-50% of wild-type levels, but unlike LMAN1-deficient mice, lethality of *Mcfd2*$^{-/-}$ mice was not observed. The *Mcfd2*$^{-}$ allele was maintained on a C57BL/6J genetic background and PCR genotyping was performed as previously described (Zhang *et al.*, manuscript in preparation).

Genotyping for all animals reported in this study was performed with mouse tail clip DNA using Go-Taq Green Master Mix (Promega) and the relevant primers listed in Table 3-1. The resulting PCR products were resolved by 2% agarose gel electrophoresis. For tissue collection, mice were anesthetized and cardiac perfusion with 10 mL of PBS was performed, after which tissues were isolated and snap frozen in liquid nitrogen for protein and RNA isolation. Timed matings were performed by intercrossing *Lman1*$^{+/-}$ mice. The following morning, designated embryonic day 0.5 (E0.5), matings were separated. Pregnant female mice were euthanized at E18.5 post-coitus. A tail biopsy was first obtained from each fetus for genotyping. Recovered fetuses were separated and placed individually in Z-fix (Aqueous Buffered Zinc Formalin, Anatech LTD) fixative solution in preparation for histologic examination. All animals were housed according to the

guidelines of the University of Michigan Unit of Laboratory Animal Medicine. The University of Michigan's University Committee on the Use and Care of Animals approved all animal protocols in this study under protocol number 08571.

mRNA isolation and gene expression analysis

Total mRNA was isolated from a panel of frozen tissues from adult wild-type mice, *Lman1*$^{gt1/gt1}$ mice, and *Lman1*$^{cgt/cgt}$ mice using the QIAshredder (Qiagen) and RNeasy Kits (Qiagen), as per manufacturer's instructions, including the optional DNaseI digestion step on column. For the RT-PCR experiment, cDNA synthesis and PCR were carried out in one reaction using the Superscript III One-step RT-PCR System with Platinum Taq (Invitrogen) and primers listed in Table 3-1. For the qRT-PCR analyses, total mRNA (1 ug) isolated from tissues was converted to cDNA with the Superscript First-Strand system (Invitrogen) using oligoDT primers. qPCR was executed with SYBR-Green RT-PCR Master Mix (Applied Biosystems) and primers listed in Table 3-1 on a 7900HT Fast Real-Time PCR machine (Applied Biosystems). Data were analyzed using the comparative threshold cycle method,[112] with the ΔCt value for *Lman1* related to the ΔCt value of *Gapdh*, a housekeeping gene used as an internal control.

The cDNA primers utilized for these assays were chosen based on the differences in the *Lman1*^{cgt} and *Lman1*^{gt1} alleles (Figure 3-3A). The *Lman1*^{cgt} allele has a gene-trap in *Lman1* intron 1, while the *Lman1*^{gt1} allele has a gene-trap in intron 10. Primer pairs F4-R2 and F6-R6 span intron 1 and should detect splicing around the *Lman1*^{cgt} gene-trap, while primer pair F5-R5 spans intron 11 and should detect transcriptional activity downstream of the *Lman1*^{gt1} gene-trap in intron 10 (in the case of splicing around the *Lman1*^{gt1} gene-trap). Primer pair F5-R5 was used, in part, because it is comparable to an RT-PCR primer pair used in the initial report of *Lman1*$^{gt1/gt1}$ mice.[95]

Tissue histology

An autopsy and histological survey was performed on adult *Lman1*$^{-/-}$ mice. Tissues were collected from anesthetized and perfused adult mice and placed directly in Z-Fix for

fixation prior to submission to the University of Michigan Cancer Center Histology Core for sectioning and H&E staining. Histologic analysis of adult tissues was performed by the University of Michigan Department of Pathology. Embryos harvested at E18.5 as well as five litters of newborn pups from *Lman1*$^{+/-}$ x *Lman1*$^{+/-}$ intercrosses were sacrificed within ~ 6 hours of birth for fixation (Z-Fix) and necropsy analysis, blinded to genotype, was provided by the veterinary pathologists of the University of Michigan Pathology Cores for Animal Research facility.

Blood and bone marrow analyses

Mice were anesthetized briefly with isoflurane and seventy microliters of blood were drawn from the retro-orbital venous sinus. Blood was diluted 70:130 in 5% Bovine Serum Albumin (BSA) in Phosphate Buffered Saline (PBS, pH 7.4). A complete blood count (CBC) was performed on the Advia120 Whole Blood Analyzer (Bayer) according to the manufacturer's instructions. Blood smears were performed as previously described[143] and were evaluated by a hematologist (RK) blinded to the mouse genotype.

Following pentobarbital-induced anesthesia, bone marrow (BM) was flushed from femurs and tibias using RPMI 1640 (Sigma-Aldrich) supplemented with 5% FBS. BM cells were collected by centrifugation in a Rotofix 32A cytospin (Hettich), stained with the HEMA 3 kit (Fisher), and examined under light microscopy. BM cytospins were evaluated by an investigator blinded to mouse genotype.

Femurs from *Lman1*$^{-/-}$ and wild-type control adult mice were harvested to assess bone marrow cellularity and architecture. The samples were processed, embedded, sectioned, and stained at the University of Michigan Pathology Cores for Animal Research facility. The femurs were fixed in 10% neutral buffered formalin prior to processing, and the bones were decalcified in Immunocal (Decal Chemical Corporation) for 24 hours.

Platelet isolation and transmission electron microscopy

Platelet-rich plasma and platelet pellets were isolated and processed as follows: 2 mL of room temperature Buffered Saline Glucose Citrate (BSGC) (129 mM NaCl, 13.6 mM Na_3 citrate, 11.1 mM glucose, 1.6 mM KH_2PO_4, 8.6 mM NaH_2PO_4, pH 7.3) were placed in a 5 mL polypropylene tube, to which was added 1-1.5 mL of whole blood. BSGC was added to a final volume of 4 mL, and gently mixed by inversion. The tubes were centrifuged at 180 x g for 10 minutes at room temperature without brake. The supernatant (semi-platelet rich plasma) was removed and centrifuged in fresh tubes at 700 x g for 10 minutes with brake. The resulting isolated platelet pellets were resuspended and fixed in 2.5% glutaraldehyde in 0.1 M Sorensen's buffer (0.1 M Na_2HPO_4, 0.1 M KH_2PO_4, pH 7.4) overnight at 4°C. Sample processing and transmission electron microscopy (TEM) was performed at the University of Michigan Microscopy & Image Analysis Laboratory. Platelets were rinsed in 0.1 M Sorensen's buffer and fixed with 1% osmium tetroxide in 0.1 M Sorensen's buffer, then rinsed in double distilled water, and subsequently stained *en bloc* with aqueous 3% uranyl acetate for 1 hour. Platelets were dehydrated in ascending concentrations of ethanol, rinsed twice in 100% ethanol, and embedded in epoxy resin. Samples were ultra-thin sectioned at 70 nm thickness and stained with uranyl acetate and lead citrate. TEM was performed using a Philips CM100 electron microscope at 60kV. Images were recorded digitally using a Hamamatsu ORCA-HR digital camera system operated with AMT software (Advanced Microscopy Techniques Corp., Danvers, MA).

Measurement of FV and FVIII levels

Blood collection via inferior vena cava (IVC) puncture and platelet-poor plasma isolation were performed as previously described (Everett *et al.*, manuscript in press; Chapter 2). Plasma FV and FVIII activity levels were determined in prothrombin time- (PT) and partial thromboplastin time- (PTT) based assays, respectively, as previously described (Everett *et al.*, manuscript in press; Chapter 2).

Western blots

Western blot analysis of mouse tissues and plasma samples was performed according to standard protocols. Briefly, frozen tissues were homogenized in ice-cold lysis buffer (100 mM Tris, pH 7.5, 1% NP-40, 10% glycerol, 25 mM β-glycerol phosphate, 130 mM NaCl, 10 mM NaF, 5 mM $MgCl_2$, 1 mM EDTA) supplemented with complete protease inhibitor (Roche). Cell lysates were cleared by centrifuging for 10 minutes at 13.2 x 10^3 rpm in an Eppendorf 5415D microcentrifuge. After normalizing protein concentrations by DC protein assay (Bio-Rad), 30 ug of protein from each tissue was fractionated by SDS-PAGE and transferred to nitrocellulose (Bio-Rad). Equal volumes of mouse plasma from individual animals were used for Western blot analysis of plasma proteins. Antibodies used in Western blot experiments were: a monoclonal antibody against GAPDH (EMD Millipore), a monoclonal antibody against RALA (Sigma-Aldrich), a polyclonal antibody against Alpha-1 antitrypsin (Genway Biotech), and a polyclonal antibody against LMAN1 (Sigma-Aldrich). HRP was detected using an enhanced chemiluminescence kit (Perkin-Elmer). ER accumulation of A1AT in liver tissues was evaluated by using an IRDye-conjugated secondary antibody (LI-COR Biosciences) and quantifying the band intensities using the Odyssey Infrared Imaging System. The Odyssey CLx Image Studio program (LI-COR Biosciences) was used for analysis.

Statistical analysis

A chi-squared test was used to evaluate statistical deviation from expected Mendelian ratios for the genotypes of offspring from the matings listed in Table 3-2. Analysis of the FV and FVIII activity assays was performed by one-way ANOVA for comparisons of more than two groups, or by student's t-test for comparison of two groups. Statistical analysis of CBC data for *Lman1* mutant mice compared to wild-type mice was performed using a student's t-test with a Bonferroni correction for multiple observations, with a p-value < 0.00625 considered statistically significant (8 CBC parameters evaluated). One-way ANOVA was also performed for an overall comparison of the mean platelet count for each genotype. A p-value ≤ 0.05 was considered to be statistically significant for all other analyses. Statistical analysis of the ratios of post-ER A1AT:ER-retained A1AT for *Lman1* mutant mice and wild-type control mice was performed using a student's t-test to

compare two specific genotypes, as well as with one-way ANOVA to compare the means of all genotypes.

RESULTS

LMAN1 deficiency results in perinatal lethality with incomplete penetrance

The genotype distribution of offspring from intercrosses of *Lman1*$^{gt1/+}$ mice on a C57BL/6J genetic background revealed a small but significant reduction in the number of observed homozygote *Lman1*$^{gt1/gt1}$ mice at 3 weeks of age compared to Mendelian prediction (17% observed vs. 25% expected, $p < 0.003$) (Table 3-2), consistent with the previous report (9-15% of pups were *Lman1*$^{gt1/gt1}$ on three genetic strain backgrounds).[95]

Heterozygous *Lman1*$^{fl/+}$ mice were crossed with transgenic *EIIA-Cre*$^{+}$ mice that carry the *Cre* recombinase gene driven by the ubiquitously expressed adenovirus *EIIA* promoter to generate *Lman1*$^{+/-}$ mice, which were maintained on a C57BL/6J genetic background. The resulting *Lman1*$^{+/-}$ mice were intercrossed to generate *Lman1*$^{-/-}$ mice. Similar to *Lman1*$^{gt1/gt1}$ mice, perinatal lethality with incomplete penetrance was observed for *Lman1*$^{-/-}$ mice ($p < 0.007$) (Table 3-2), thereby ruling out a passenger gene effect as the cause of the originally reported lethality. Genotyping of pups from this intercross that died spontaneously at P0 revealed an excess of *Lman1*$^{-/-}$ mice (11/22 pups) ($p < 0.007$). In contrast, *Lman1*$^{cgt/cgt}$ mice on a C57BL/6J genetic background are viable and were observed in the expected Mendelian ratios ($p > 0.19$) (Table 3-2).

To assess *Lman1*$^{-/-}$ lethality on a mixed genetic background, the *Lman1*$^{-}$ allele was outcrossed to the 129 SvImJ genetic background. Offspring from intercrosses of *Lman1*$^{+/-}$ (B6;129) F1 mice were observed in the expected Mendelian ratios, with no evidence for perinatal lethality among *Lman1*$^{-/-}$ mice on this genetic background ($p > 0.79$) (Table 3-2). These results are consistent with the original *Lman1*$^{gt1/gt1}$ report.[95]

No consistent abnormalities were observed on routine autopsy and histologic examination of adult *Lman1*$^{-/-}$ mice (Supplementary Figure 3-1) or on necropsy of *Lman1*$^{-/-}$ pups and wild-type littermates sacrificed at embryonic day 18.5 (E18.5) or several hours after birth.

Comparison of *Lman1* expression from different mutant alleles

LMAN1 protein levels in liver lysates prepared from *Lman1*$^{gt1/gt1}$ mice, *Lman1*$^{cgt/cgt}$ mice, and *Lman1*$^{-/-}$ mice were compared by Western blot analysis (Figure 3-4A), demonstrating that LMAN1 is expressed with variable levels in the different *Lman1* mutant mouse strains. Relative to wild-type mice, the *Lman1*$^{cgt/+}$ (lane 2), *Lman1*$^{+/-}$ (lane 4), and *Lman1*$^{gt1/+}$ (lane 6) heterozygote animals express moderately reduced levels of LMAN1. The *Lman1*$^{cgt/cgt}$ mice (lane 3) retain a low level of residual LMAN1 expression, indicating that the *Lman1*^{cgt} allele is not a null allele. In contrast, LMAN1 protein is not detectable by Western blot in either the *Lman1*$^{-/-}$ (lane 5) or *Lman1*$^{gt1/gt1}$ (lane 7) samples. Analysis of several *Lman1*$^{cgt/cgt}$ mouse tissues reveals detectable LMAN1 expression at varying levels (Figure 3-4B).

RT-PCR and qPCR on liver samples were performed to assess *Lman1* mRNA expression. *Lman1*$^{cgt/cgt}$ mice exhibit detectable wild-type *Lman1* mRNA by RT-PCR, consistent with the Western blot results (Figure 3-3B). An alternate splice product of the *Lman1*^{cgt} allele was also observed (Figure 3-5). In contrast, RT-PCR analysis detected no wild-type *Lman1* mRNA in *Lman1*$^{gt1/gt1}$ mice (Figure 3-3B).

qPCR analysis demonstrated that *Lman1*$^{cgt/cgt}$ mice retain ~5-7% of normal levels of *Lman1* mRNA expression in the liver and kidney (Figure 3-3 C-F). qPCR of liver mRNA from *Lman1*$^{gt1/gt1}$ mice detected no wild-type mRNA (Figure 3-3D), though with evidence for reduced stability of the β-geo fusion mRNA compared to wild-type *Lman1* mRNA (Figure 3-3C).

Lman1$^{-/-}$ mice are thrombocytopenic

Relative to wild-type mice, *Lman1*$^{-/-}$ mice exhibit normal CBC parameters (Figure 3-6 A-G), with the exception of a mild, isolated thrombocytopenia ($p < 0.004$ with Bonferroni correction) (Figure 3-6H). In contrast, *Lman1*$^{+/-}$ mice, *Lman1*$^{cgt/cgt}$ mice, and *Mcfd2*$^{-/-}$ mice all exhibit normal platelet counts indistinguishable from C57BL/6J wild-type control mice (Figure 3-7). No platelet abnormalities were detected in peripheral blood smears of *Lman1*$^{-/-}$ mice (Figure 3-8A). Longitudinal histologic sections of mouse femurs also revealed no abnormalities in bone marrow cellularity or architecture (Figure 3-8B; Supplementary Table 3-1). TEM of platelets isolated from *Lman1*$^{-/-}$ mice and wild-type littermate controls revealed no differences in platelet morphology, as assessed by three independent reviewers blinded to mouse genotype (Figure 3-9).

Relationship between *Lman1* expression and LMAN1-dependent FV and FVIII secretion

Lman1$^{-/-}$ null mice on a pure C57BL/6J genetic background or on a mixed B6;129 genetic background exhibit plasma FV (Figure 3-10A, Figure 3-10C) and FVIII (Figure 3-10B, Figure 3-10D) activity levels that are ~50% of wild-type levels, consistent with the previous report for *Lman1*$^{gt1/gt1}$ mice.[95] In contrast, *Lman1*$^{cgt/cgt}$ mice exhibit plasma FV (Figure 3-10A) and FVIII (Figure 3-10B) activity levels that are moderately reduced to ~67% ($p < 3.0 \times 10^{-4}$) and ~63% ($p < 9.2 \times 10^{-7}$) of wild-type levels, respectively. Plasma FV (Figure 3-11A) and FVIII (Figure 3-11B) activity levels are indistinguishable between *Lman1*$^{gt1/gt1}$ mice and *Lman1*$^{-/-}$ mice.

Consistent with previous reports of *Lman1*$^{gt1/+}$ mice[95] and with the autosomal recessive inheritance of F5F8D, heterozygous carriers of the *Lman1*^{cgt}, *Lman1*^{gt1}, and the *Lman1*$^{-}$ mutant alleles on either a pure C57BL/6J genetic background (Figure 3-12A and Figure 3-12B) or on a mixed B6;129 genetic background (Figure 3-12C and Figure 3-12D) demonstrate FV and FVIII activity levels indistinguishable from control mice.

FV and FVIII plasma activity levels of *Lman1*$^{-/-}$ mice were compared directly to that of *Mcfd2*$^{-/-}$ mice, the other murine model of F5F8D. There was a trend ($p > 0.08$) for FV

activity levels to be slightly lower in *Mcfd2*$^{-/-}$ mice as compared to *Lman1*$^{-/-}$ mice (Figure 3-13A), which is consistent with previous reports of FV and FVIII plasma levels in human F5F8D patients with *MCFD2* mutations as compared to *LMAN1* mutations.[39] However, FVIII activity was indistinguishable between *Lman1*$^{-/-}$ and *Mcfd2*$^{-/-}$ mice (Figure 3-13B).

ER accumulation of A1AT in the *Lman1* allelic series

Western blot analysis of *Lman1*$^{gt1/gt1}$ hepatic lysates confirmed the previous report of A1AT retention in the ER (Figure 3-14A).[95] A1AT also accumulates in the ER of *Lman1*$^{cgt/cgt}$ mice and *Lman1*$^{-/-}$ mice, relative to wild-type control mice (Figure 3-14A). Null *Lman1*$^{-/-}$ mice demonstrate the highest levels of A1AT ER retention, with intermediate levels in the *Lman1*$^{cgt/cgt}$ mice (Figure 3-14B).

DISCUSSION

We report the characterization of a series of *Lman1* mutant alleles (summarized in Table 3-3), with *Lman1* gene expression ranging from ~7% of normal levels (*Lman1*$^{cgt/cgt}$ mice) to no gene expression (*Lman1*$^{-/-}$). We exclude trace residual LMAN1 expression as the explanation for the more modest reduction in FV and FVIII in previously reported gene targeted mice[95] compared to humans with F5F8D, and observe an intermediate level of FV and FVIII reduction in *Lman1*$^{cgt/cgt}$ mice. We confirmed the previously reported partially penetrant, perinatal, strain-specific lethality of LMAN1-deficient mice.[95] In addition, we report that *Lman1*$^{-/-}$ mice are mildly thrombocytopenic, with no detectable histologic defect in the bone marrow or peripheral blood.

In the original report of the *Lman1*$^{gt1/gt1}$ F5F8D mice, partial perinatal lethality was observed on the C57BL/6J, 129/SvImJ, and DBA/2J genetic backgrounds, with no significant lethality observed on a mixed B6;129 or A/J genetic background.[95] We confirm this strain-specific partial lethality in an independent *Lman1* targeted allele,

excluding a passenger gene effect not due to LMAN1 deficiency as the cause of lethality in LMAN1 null mice.

In contrast to the partial lethality observed in *Lman1*$^{-/-}$ mice, mice homozygous for the hypomorphic *Lman1*^{cgt} allele are observed in expected Mendelian proportions. Taken together with the lack of a similar perinatal phenotype in human F5F8D patients, despite much greater reductions in FV and FVIII levels in humans, these data suggest the existence of another as yet unknown LMAN1-dependent cargo with a critical function in the neonatal period. Possible explanations for this difference in phenotype include the fact that humans (like the outbred B6;129 mice) are more genetically diverse than highly inbred C57BL/6J mice, and this genetic diversity is expected to be protective against the effects of a deleterious genetic mutation (relative to the homogeneous C57BL/6J genetic background). The different phenotypes in mice and humans could also result from biologic or physiologic differences between these species. For example, there may exist a secretory cargo that is essential in mice but not in humans, or that has a different function in mice than in humans. Alternatively, the same cargo may be present in both species, but its impaired secretion in humans may not be as limiting as in mice. Examination of *Lman1*$^{-/-}$ pups at E18.5 and at birth has revealed no obvious anatomic or histologic abnormalities to suggest the underlying cause of death in these animals. One possible exploratory approach to identify the cause of death in *Lman1*$^{-/-}$ neonates would be to generate a number of different tissue-specific *Lman1* knockout strains utilizing various tissue-specific Cre recombinases, which would reveal the tissues in which LMAN1 deficiency is sufficient to cause neonatal lethality.

The observation that reported F5F8D human patients with *LMAN1* mutations carry two null *Lman1* alleles suggested that even low levels of residual LMAN1 function are sufficient to support normal FV and FVIII secretion.[36] Our results in the mouse suggest an ascertainment bias, with F5F8D patients with milder LMAN1 mutations escaping clinical detection. Based on these data, LMAN1/MCFD2 genetic analyses could be considered in patients with only modest reductions in FV and FVIII levels (in the 30-70% range). Previous GWAS analyses of over 23,000 individuals for genetic loci associated

with plasma FVIII levels[144] failed to identify any signal at the *LMAN1* locus on chromosome 18, excluding common genetic variants in *LMAN1* as a key determinant of FVIII variation in the general population. However, we cannot rule out the possibility that *LMAN1* variation may moderately contribute to population variation in FV and FVIII levels. The observation that the low level of *Lman1* expression in *Lman1*$^{cgt/cgt}$ mice is associated with intermediate reductions in FV, FVIII, and A1AT secretion suggests that targeting of LMAN1 to reduce FV and FVIII as an anticoagulant strategy may only require partial LMAN1 reduction.

Though A1AT was identified *in vitro* as an LMAN1-dependent secretory cargo,[85] plasma levels of A1AT were indistinguishable between two human F5F8D patient siblings, their obligate carrier parents, and 2 normal controls.[95] Plasma A1AT levels are also indistinguishable between *Lman1*$^{gt1/gt1}$ mice and wild-type controls, though retained within the ER in *Lman1*$^{gt1/gt1}$ hepatocytes.[95] We confirm retention of A1AT in hepatocytes of *Lman1* null mice, with intermediate levels in the ER of *Lman1*$^{cgt/cgt}$ hepatocytes. Thus, a similar relationship between *Lman1* gene expression levels and LMAN1-dependent cargo secretion applies to A1AT secretion as was observed for FV and FVIII secretion.

Complete LMAN1 deficiency in the mouse results in plasma FV and FVIII levels reduced to ~50% of normal, compared to ~10-15% levels in human F5F8D patients, suggesting the existence of alternative compensatory or parallel mechanisms for the secretion of at least some LMAN1-dependent cargos, including FV and FVIII. Differences in these alternate secretion pathways between mice and humans (for example, the number of alternate pathways, the specificity of each pathway for certain cargos, or the rate of cargo transport in the pathways) may account for the difference in FV and FVIII secretion levels between mice and humans.

Finally, the observation of mild thrombocytopenia in *Lman1*$^{-/-}$ mice was unexpected, since no such abnormality has ever been noted in human patients. No other hematologic abnormalities were observed in the mouse or previously in human patients. A previous

GWAS analysis for platelet-associated traits in 66,867 subjects identified 68 genetic loci associated with platelet count or platelet volume, though no loci associated with platelet count are close to the *LMAN1* locus.[145] These data suggest that common variants at the *LMAN1* locus do not contribute significantly to platelet count variation in the general population. Given the mild reduction in platelet count in *Lman1*$^{-/-}$ mice, a similar subtle phenotype in human F5F8D patients could have been missed and would likely require analysis of a large cohort of F5F8D patients for detection. Morphologic analyses of platelets and megakaryocytes in *Lman1*$^{-/-}$ mice revealed no abnormalities, and the mechanism for this mild thrombocytopenia remains unknown. However, reduction in secretion of another, as yet unknown LMAN1-dependent cargo seems likely. In this case, the cargo could represent either a component of the platelet itself, or an extrinsic factor affecting platelet number that is secreted by another cell type.

Table 3-1: Primer sequences	
Primer	5' → 3' Sequence
Genotyping primers	
Lman1 primer A	GGCTTTCTTGACACCTTCAATTTAA
Lman1 primer B	CCAAGTGAAGGGAAGACCATCAAGC
Lman1 primer C	GACCCCTAGTGACGGGTTCTTGTC
Flpe transgene Fwd	GGTCCAACTGCAGCCCAAGCTTCC
Flpe transgene Rev	GTGGATCGATCCTACCCCTTGCG
EIIA-Cre transgene Fwd	CCGCTGGAGATGACGTAGTT
EIIA-Cre transgene Rev	CGCATAACCAGTGAAACAGCATTGC
RT-PCR and qPCR primers	
β-*actin* Fwd	CTAAGGCCAACCGTGAAAAG
β-*actin* Rev	GGGGTGTTGAAGGTCTCAAA
Gapdh Fwd	TGTGTCCGTCGTGGATCTGA
Gapdh Rev	ACCACCTTCTTGATGTCATCATACTT
F4	CCGTTTCGAGTACAAATACAGC
R2	TTGATTTGTCCATTGTTGCCTA
F5	ACGTGGTGAAGAGAGATATCGAC
R5	ACAAAGTGGATCGTGGATAGACA
F6	GCCCAGGCGGGGAATGCTATTCCAAG
R6	TCTCGAAGGCTGCTTTTGCTTTGGT

Table 3-2: Intercrosses of *Lman1* mutant mice

Crosses	Strain	Genotype distribution at 3 weeks: percent (number observed)		p-value (χ^2)
		+/+ and +/-	**-/-**	
Expected %		**75 %**	**25%**	
Lman1$^{gt1/+}$ x *Lman1*$^{gt1/+}$	C57BL/6J	83% (201)	17% (40)	< 0.003
Lman1$^{cgt/+}$ x *Lman1*$^{cgt/+}$	C57BL/6J	80% (107)	20% (27)	> 0.19
Lman1$^{+/-}$ x *Lman1*$^{+/-}$	C57BL/6J	83% (184)	17% (38)	< 0.007
Lman1$^{+/-}$ x *Lman1*$^{+/-}$	B6;129 mixed	76% (95)	24% (30)	> 0.79

The chi-squared test was based upon an expected genotype ratio of 3:1, with *Lman1*$^{-/-}$ mice expected to represent 25% of the offspring from each mating, and all other genotypes cumulatively accounting for 75% of offspring.

Table 3-3: Summary of homozygous *Lman1* mutant mice on C57BL/6J genetic background			
***Lman1* allele**	**% wild-type *Lman1* expression in homozygote mice (qRT-PCR)**	**Lethality of homozygotes?**	**FV and FVIII activity levels**
Lman1^{+}	100%	NA	100%
Lman1^{cgt}	~7%	No	~70%
Lman1^{-}	0%	Yes (~40%)	~50%
Lman1^{gt1}	<1%	Yes (~40%)	~50%

Supplementary Table 3-1: Histologic evaluation of mouse bone marrow			
Mouse	Genotype	Myeloid:Erythroid ratio	Megakaryocyte number (per longitudinal section)
77022	*Lman1*$^{-/-}$	1.85	322
77021	*Lman1*$^{-/-}$	1.77	341
77115	*Lman1*$^{-/-}$	1.70	268
77312	*Lman1*$^{+/+}$	1.94	322
77327	*Lman1*$^{+/+}$	1.63	354

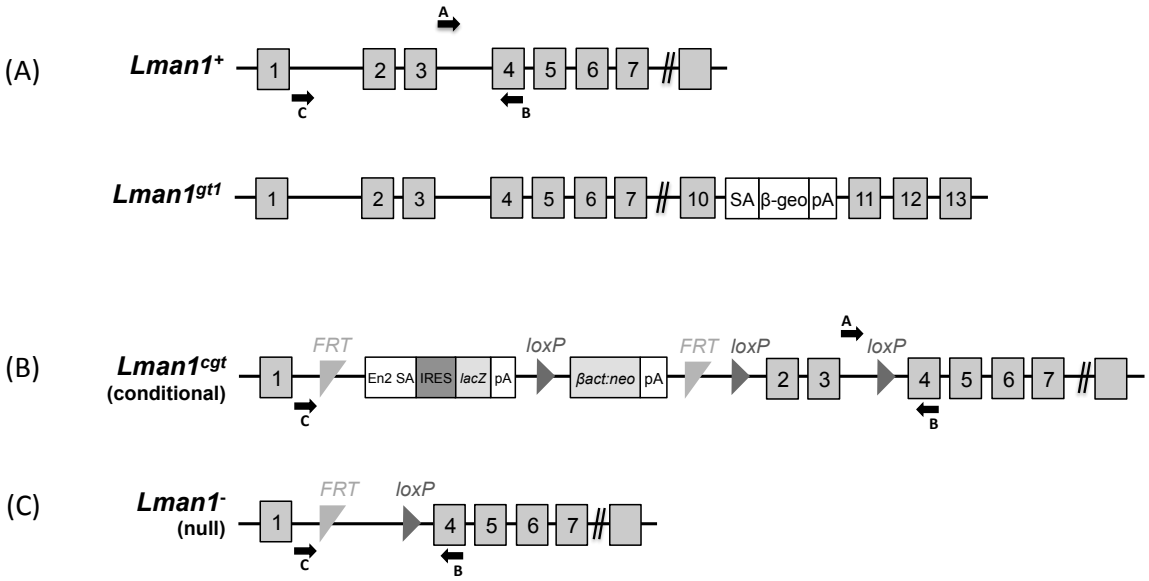


Figure 3-1: Schematic of *Lman1* mutant alleles

(A) The original, previously published *Lman1* gene-trap allele (*Lman1*^{gt1}) carries a gene-trap insertion in intron 10.[95] SA, splice acceptor cassette; β-Geo, β-galactosidase-neo fusion; pA, poly-adenylation sequence. **(B)** The *Lman1* conditional gene-trap allele (*Lman1*^{cgt}) contains a gene-trap insertion in intron 1 flanked by 2 *FRT* sites. Mice carrying this allele were crossed to β-actin *FLP* transgenic mice. Mice heterozygous for the resulting *Lman1* "floxed" allele (*Lman1*$^{fl/+}$) were crossed to *EIIA-Cre*$^{+}$ transgenic mice to excise exons 2 and 3, generating **(C)** the *Lman1* null allele (*Lman1*$^{-}$). Gray blocks represent exons. A, B, and C represent genotyping primers. (Adapted from the Knockout Mouse Project. *Lman1* targeting: https://www.komp.org/geneinfo.php?geneid=66654); *cgt,* conditional gene trap; *fl*, floxed; En2 SA, splice acceptor of mouse *En2* exon 2; IRES, encephalomyocarditis virus (EMCV) internal ribosomal entry site; *lacZ*, *E.coli* β-galactosidase gene; pA, SV40 polyadenylation signal; β*act:neo*, human β-actin promoter-driven neomycin cassette.

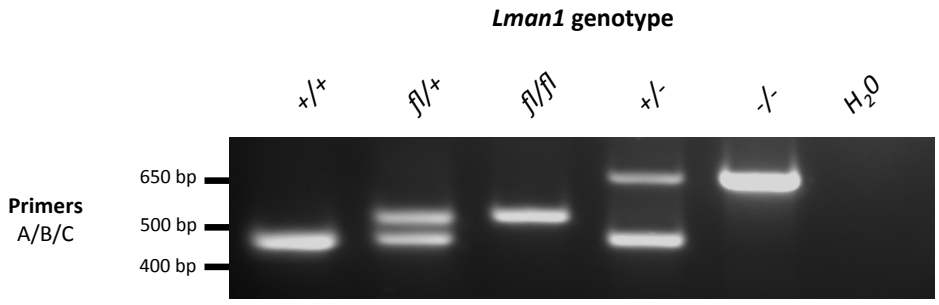


Figure 3-2: *Lman1* PCR genotyping strategy

Lman1$^{fl/+}$ mice were crossed to *EIIA-Cre*$^{+}$ transgenic mice to generate the *Lman1*$^{-}$ allele. A three-primer PCR assay (primers A, B, and C) distinguishes the *Lman1*$^{+}$ (444 bp), *Lman1*^{fl} (508 bp), and *Lman1*$^{-}$ (635 bp) alleles. Primer pair A/B also distinguishes between the *Lman1*$^{+}$ allele and the *Lman1*^{cgt} allele (508 bp) (not shown).

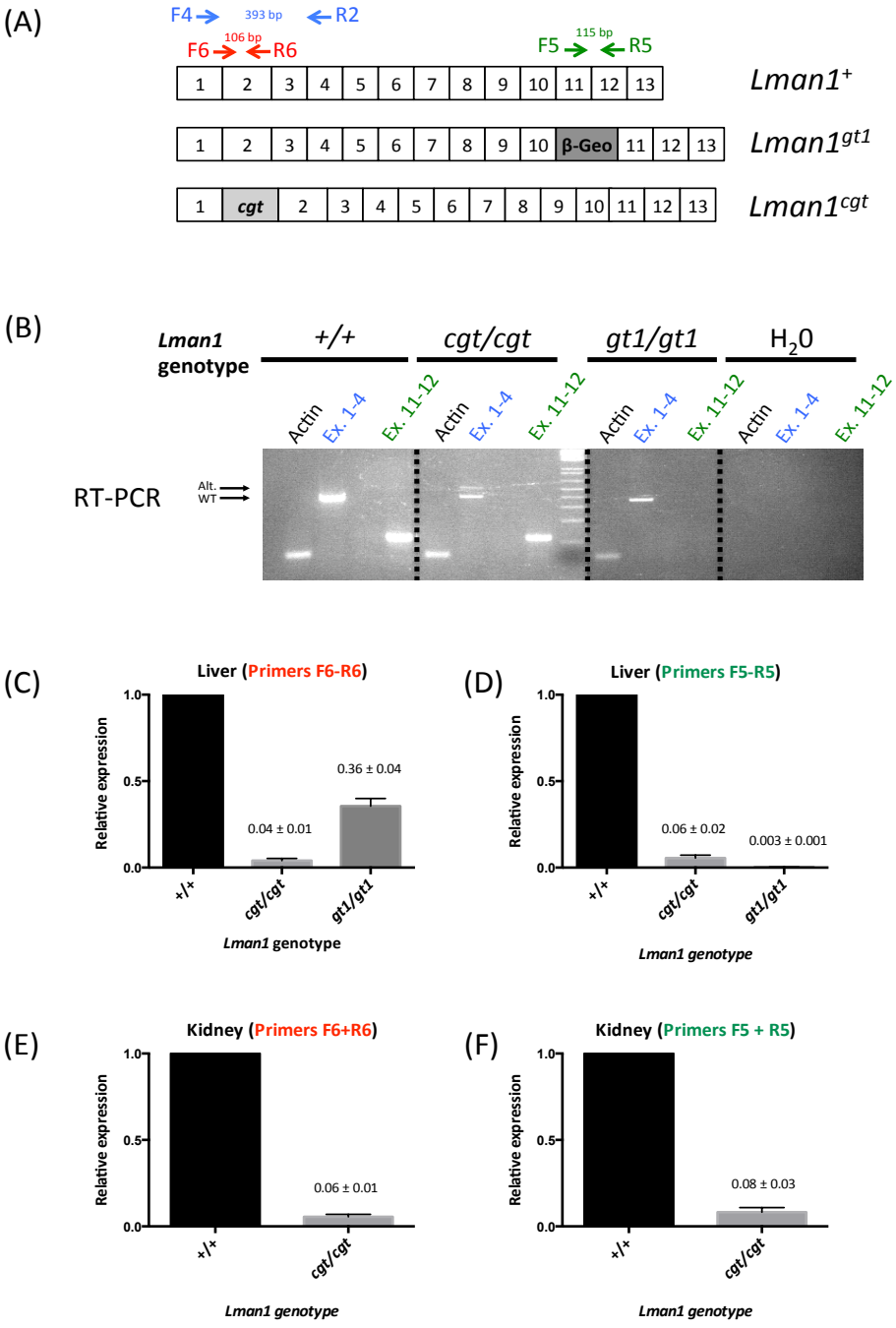


Figure 3-3: *Lman1* expression levels in *Lman1*$^{cgt/cgt}$ and *Lman1*$^{gt1/gt1}$ mice

(A) Schematic of *Lman1*^{gt1} and *Lman1*^{cgt} alleles. Primer pairs used for RT-PCR (blue and green) and qPCR (red and green) are indicated by arrows. Primer pair F4/R2 amplifies *Lman1* mRNA transcript between exons 1-4 (393 bp product), while primer pair F5/R5 amplifies *Lman1* mRNA transcript sequence between exons 11-12 (115 bp product). **(B)** RT-PCR analysis of *Lman1* expression. β-actin primers were used as a positive control. Primers F4/R2 were used to amplify sequence targets between *Lman1*

exons 1-4 (flanking the *Lman1*^{cgt} gene-trap), while primers F5/R5 were used to amplify target sequences downstream of the *Lman1*^{gt1} gene-trap. Two bands are detected in the Ex.1-4 lane for the *Lman1*$^{cgt/cgt}$ mouse, indicating that splicing around the *Lman1*^{cgt} gene-trap occurs. As indicated by the black arrows, the lower band is the expected 393 bp product (WT) while the upper band (Alt.) represents a 508 bp alternatively spliced product (see Figure 3-5). Similarly, downstream exon 11-12 sequences are also amplified in the *Lman1*$^{cgt/cgt}$ mouse, indicating that full-length, wild-type *Lman1* mRNA transcripts are present in this sample. In contrast, an expected Ex. 1-4 band is observed in the *Lman1*$^{gt1/gt1}$ mouse (since the *Lman1*^{gt1} gene-trap insertion occurs downstream), but no detectable band is observed in the Ex. 11-12 lane for this mouse, consistent with complete or nearly complete absence of wild-type *Lman1* mRNA transcripts in *Lman1*$^{gt1/gt1}$ mice. **(C and D)** qPCR of liver cDNA from *Lman1*$^{cgt/cgt}$ and *Lman1*$^{gt1/gt1}$ mice indicates that *Lman1*$^{cgt/cgt}$ mice retain ~5-6% of normal *Lman1* expression in the liver, whereas *Lman1*$^{gt1/gt1}$ mice retain < 1% of normal expression. **(E and F)** qPCR of kidney cDNA from *Lman1*$^{cgt/cgt}$ mice indicates that *Lman1*$^{cgt/cgt}$ mice retain ~7% of normal *Lman1* expression in the kidney. Three mice per genotype were used for qPCR analysis. Horizontal bars represent means, error bars represent standard error of the mean.

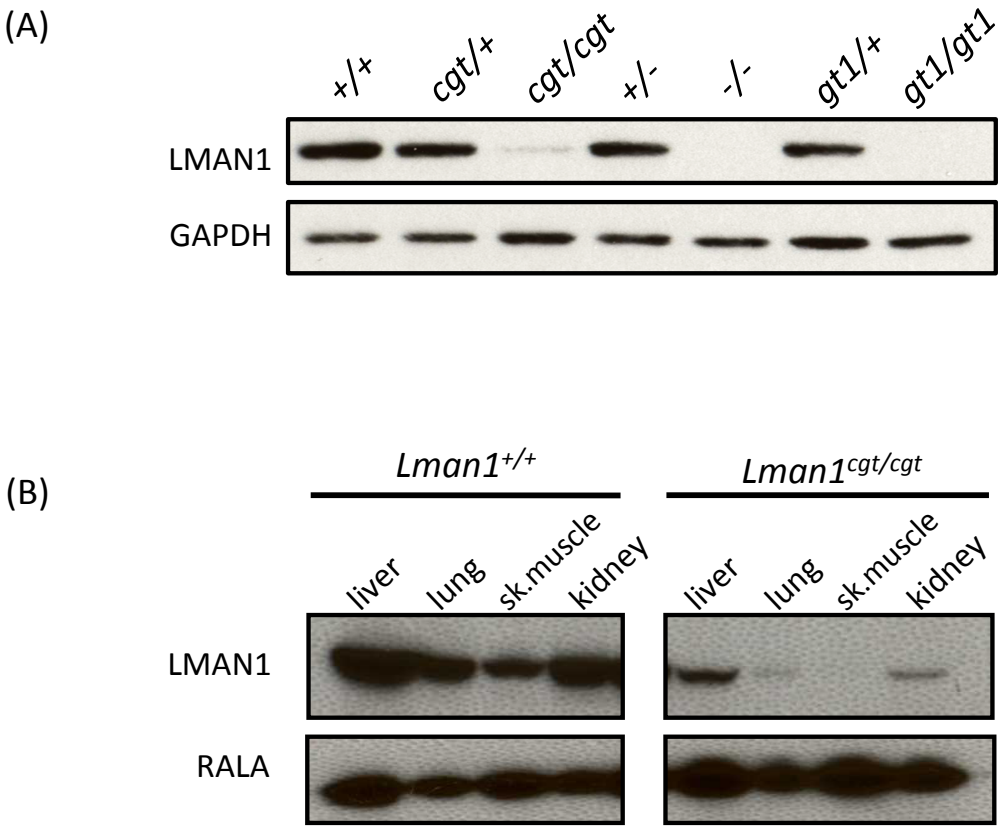


Figure 3-4: Western blot of LMAN1 expression in *Lman1* mutant mice

(A) Western blot of liver lysates from heterozygote and homozygote carriers of the *Lman1* mutant alleles: $Lman1^{cgt}$, $Lman1^{-}$, and $Lman1^{gt1}$. $Lman1^{cgt/cgt}$ mice are hypomorphic, retaining a low level of normal LMAN1 expression due to splicing around the gene-trap cassette. $Lman1^{-/-}$ and $Lman1^{gt1/gt1}$ mice demonstrate no detectable LMAN1 protein by Western blotting. Each lane corresponds to a single animal. **(B)** LMAN1 protein is detected in multiple tissues of $Lman1^{cgt/cgt}$ mice at proportional, but reduced, levels relative to wild-type mice.

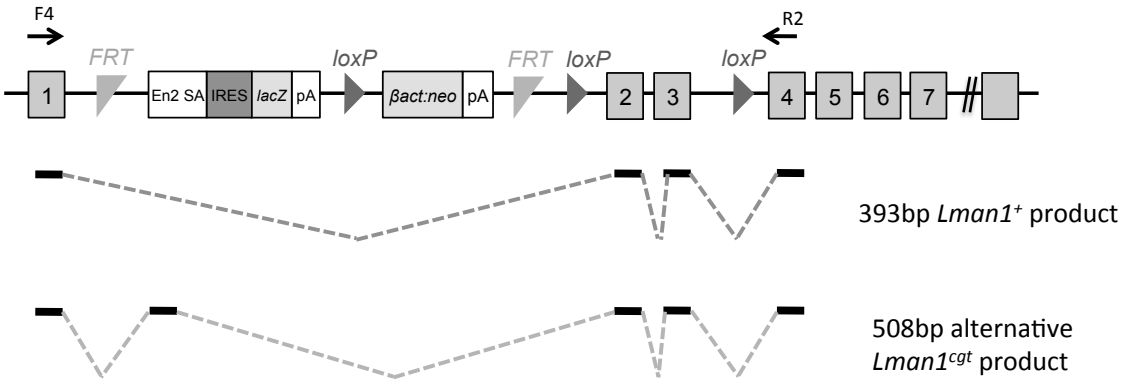


Figure 3-5: Alternative splicing of the *Lman1*^{cgt} allele

A 508bp alternatively spliced transcript was observed in *Lman1*$^{cgt/cgt}$ mice by RT-PCR with primers F4/R2 (Figure 3-3B), generated by a splicing event in which 115 nucleotides from the mouse *En2* exon 2 in the KOMP targeting construct was incorporated into the transcript. The strong splice acceptor sequence of the *En2* exonic sequence was used to drive this alternative splicing event. An internal pseudo splice-donor sequence exists immediately following the 115 bp of sequence incorporated into the alternative splicing product, facilitating splicing out of the *En2* exon and into *Lman1* exon 2.

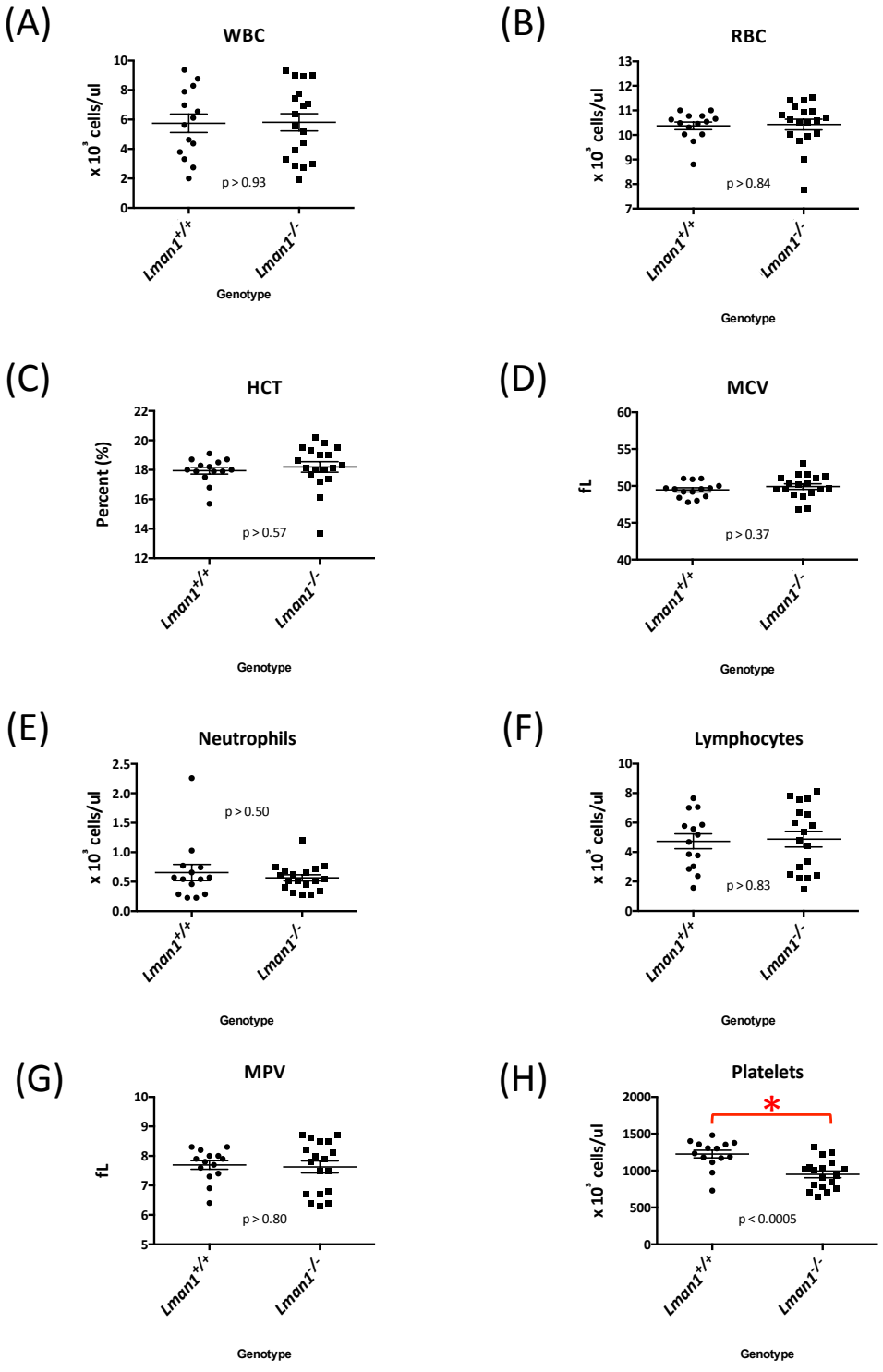


Figure 3-6: *Lman1*$^{-/-}$ mice are thrombocytopenic

Complete blood count analyses for *Lman1*$^{-/-}$ mice and wild-type control mice. **(A)** White blood cells (WBC), **(B)** red blood cells (RBC), **(C)** hematocrit (HCT), **(D)** mean

corpuscular volume (MCV), **(E)** neutrophils, **(F)** lymphocytes, and **(G)** mean platelet volume (MPV) were indistinguishable between wild-type control mice and *Lman1*$^{-/-}$ mice. **(H)** Platelet count was statistically significantly reduced in *Lman1*$^{-/-}$ mice relative to wild-type control mice ($p < 0.0005$). Given that 8 different CBC parameters were compared for the wild-type and *Lman1*$^{-/-}$ mice, a Bonferroni correction was applied to account for multiple observations. A p-value < 0.00625 was therefore considered statistically significant for these analyses, and significance is indicated by a (*) in the figure. Each symbol represents an individual animal. Horizontal lines indicate mean and error bars indicate standard error of the mean for each genotype.

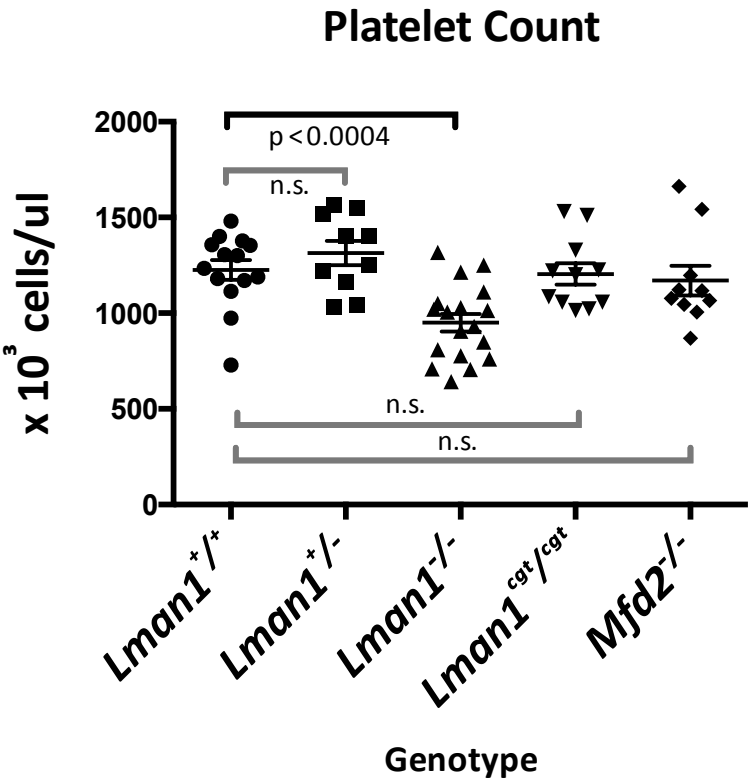


Figure 3-7: Thrombocytopenia is unique to *Lman1*$^{-/-}$ mice

Platelet counts for *Lman1*$^{+/+}$ mice, *Lman1*$^{+/-}$ mice, *Lman1*$^{-/-}$ mice, *Lman1*$^{cgt/cgt}$ mice, and *Mcfd2*$^{-/-}$ mice. Comparisons between individual genotypes were performed with a student's t-test. The thrombocytopenia phenotype is specifically observed in *Lman1*$^{-/-}$ mice ($p < 0.0004$). The differences among the mean platelet counts were statistically significant as determined by one-way ANOVA ($F(4,58) = 6.688$, $p < 0.0002$). Each symbol represents an individual animal. Horizontal lines indicate mean and error bars indicate standard error of the mean for each genotype. The platelet count data for *Lman1*$^{+/+}$ and *Lman1*$^{-/-}$ mice are the same as the data points shown in Figure 3-6H.

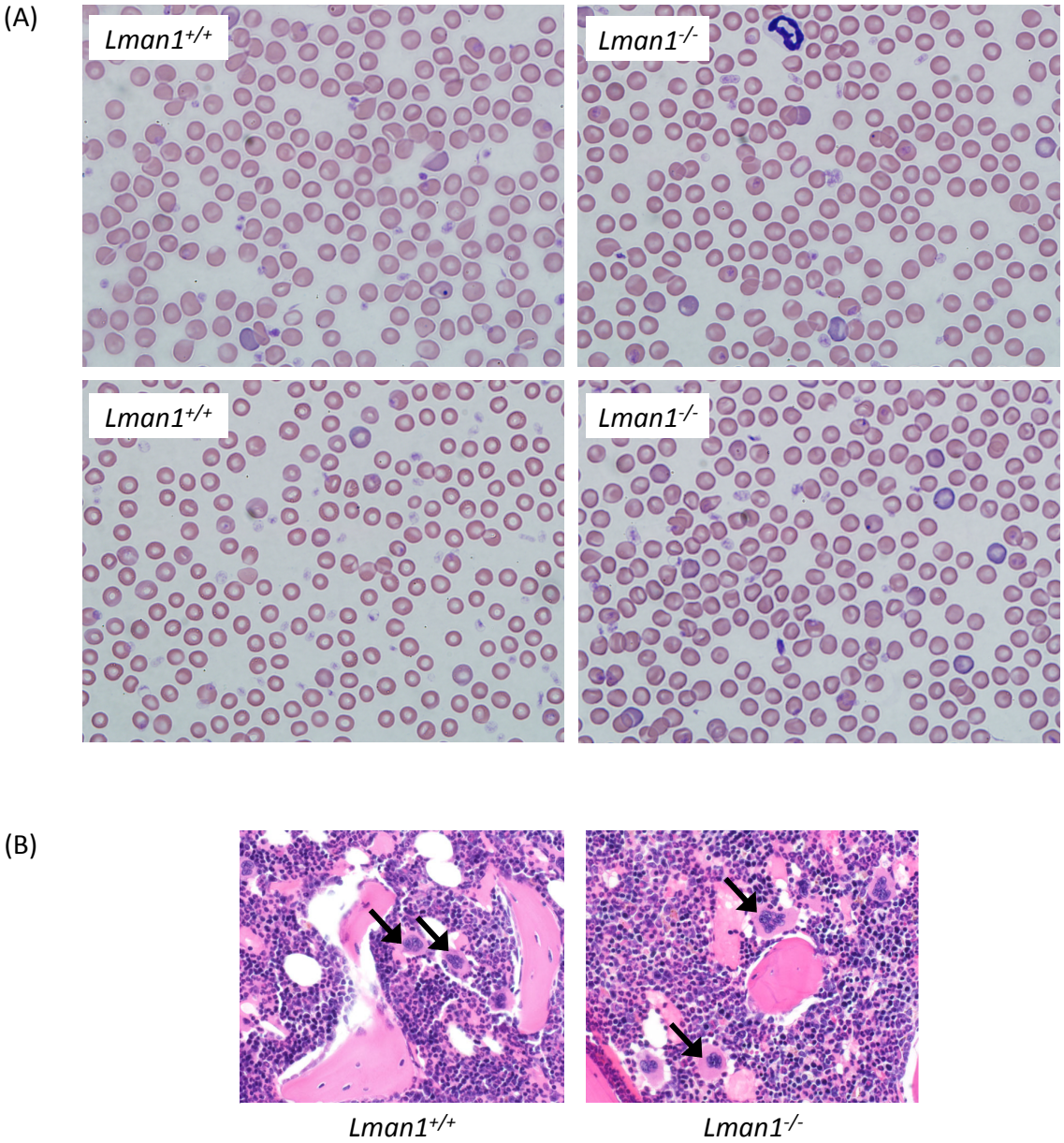


Figure 3-8: Peripheral blood smears and bone marrow of *Lman1*$^{-/-}$ mice

(**A)** Peripheral blood smears of *Lman1*$^{-/-}$ and wild-type mice. **(B)** Histology photomicrographs of bone marrow from an *Lman1*$^{-/-}$ mouse and wild-type control mouse were obtained to assess the bone marrow architecture and cellularity. Erythroid and myeloid progenitors are observed. Megakaryocyte (arrowheads) morphology appears normal in both images (H&E, 60X).

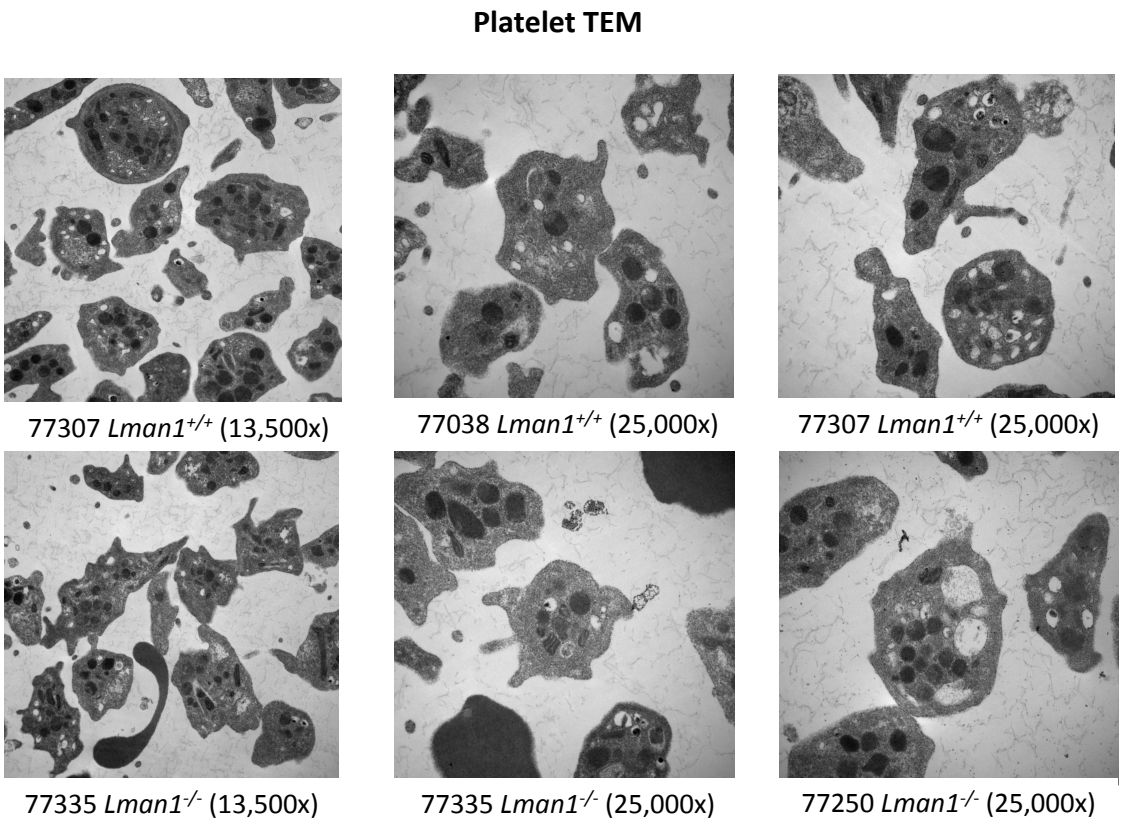


Figure 3-9: Transmission electron micrographs of platelets from *Lman1*$^{+/+}$ mice and *Lman1*$^{-/-}$ mice are indistinguishable

Platelets were isolated from two adult wild-type control mice (77307 and 77038) and two *Lman1*$^{-/-}$ mice (77335 and 77250). Samples were viewed at 13,500x and 25,000x direct magnification. Three blinded investigators assessed platelet and granule morphology.

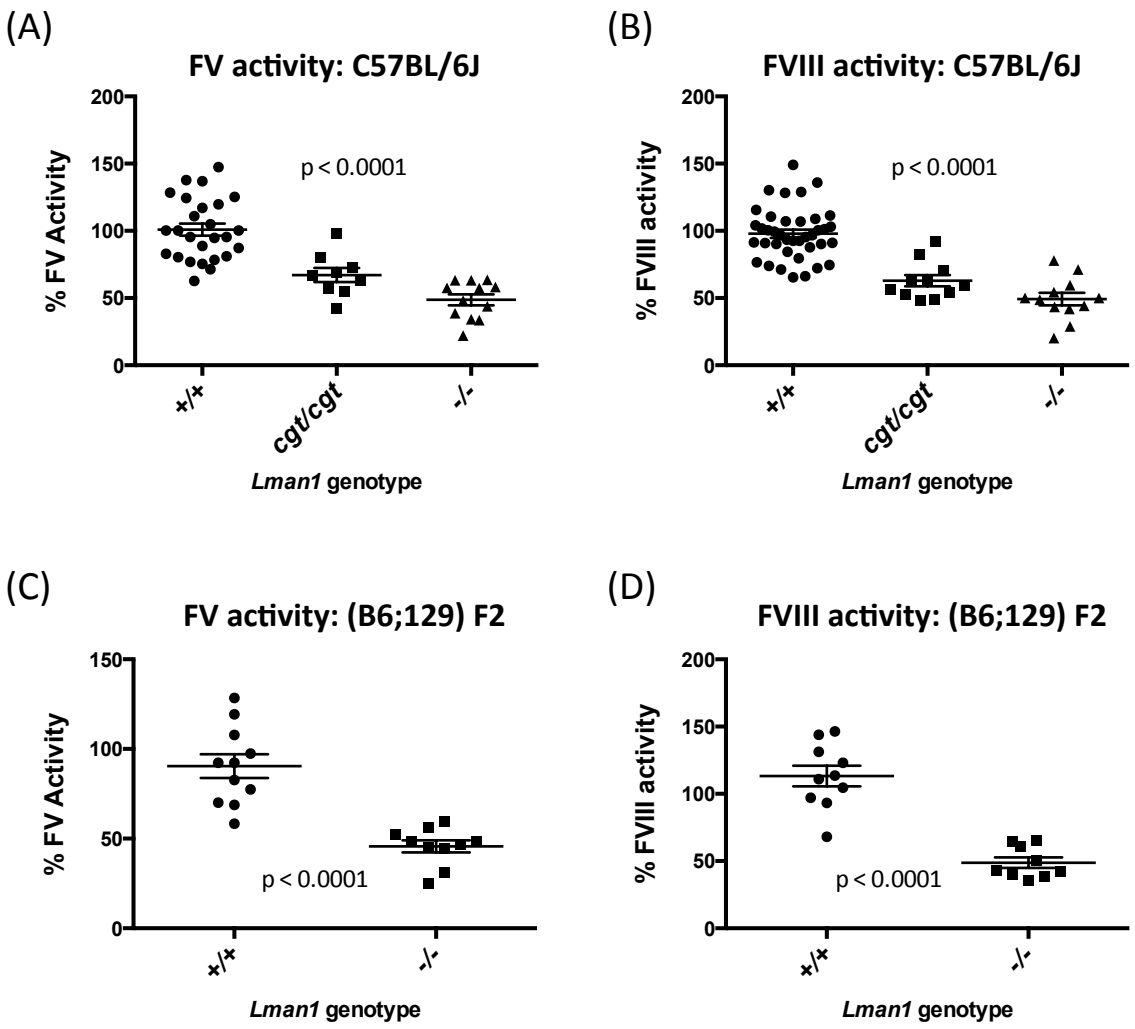


Figure 3-10: Plasma FV and FVIII activity levels for *Lman1*$^{-/-}$ mice on a C57BL/6J genetic background does not differ from mixed B6;129 genetic background

(A) Plasma FV activity measured by PT-based coagulation assay. *Lman1*$^{cgt/cgt}$ mice (C57BL/6J) exhibit a moderate reduction in plasma FV activity, relative to wild-type control mice (67.1% ± 5.3%). In contrast, *Lman1*$^{-/-}$ mice (C57BL/6J) demonstrate FV activity levels that are 48.7% ± 4.0% of normal, relative to wild-type, which is consistent with previously published reports of *Lman1*$^{gt1/gt1}$ mice.[95] The mean plasma FV activity levels for *Lman1*$^{+/+}$, *Lman1*$^{cgt/cgt}$, and *Lman1*$^{-/-}$ mice are statistically significantly different, as determined by one-way ANOVA ($F(2,44) = 31.33$, $p < 0.0001$). **(B)** Plasma FVIII activity measured by PTT-based coagulation assay. Plasma FVIII activity levels in *Lman1*$^{cgt/cgt}$ mice (C57BL/6J) are moderately reduced, relative to wild-type control mice (62.8% ± 4.2%). In contrast, *Lman1*$^{-/-}$ mice (C57BL/6J) demonstrate a ~50% reduction of FVIII activity levels (49.2% ± 4.6%), relative to wild-type, which is consistent with previously published reports of *Lman1*$^{gt1/gt1}$ mice.[95] The mean plasma FVIII activity levels for *Lman1*$^{+/+}$, *Lman1*$^{cgt/cgt}$, and *Lman1*$^{-/-}$ mice are statistically significantly

different, as determined by one-way ANOVA ($F(2,59) = 42.07$, $p < 0.0001$). **(C)** *Lman1*$^{-/-}$ F2 mice (B6;129) exhibit an ~50% reduction in plasma FV activity, relative to wild-type control mice (45.7% ± 3.3%). **(D)** Plasma FVIII activity levels in *Lman1*$^{-/-}$ F2 mice (B6;129) are also reduced to ~50% of normal levels (48.8% ± 4.0%). The reductions in plasma FV and FVIII activity in *Lman1*$^{-/-}$ mice on a mixed B6;129 genetic background are comparable to that observed on a pure C57BL/6J genetic background. Each symbol represents an individual animal. Horizontal lines indicate mean and error bars indicate standard error of the mean for each genotype.

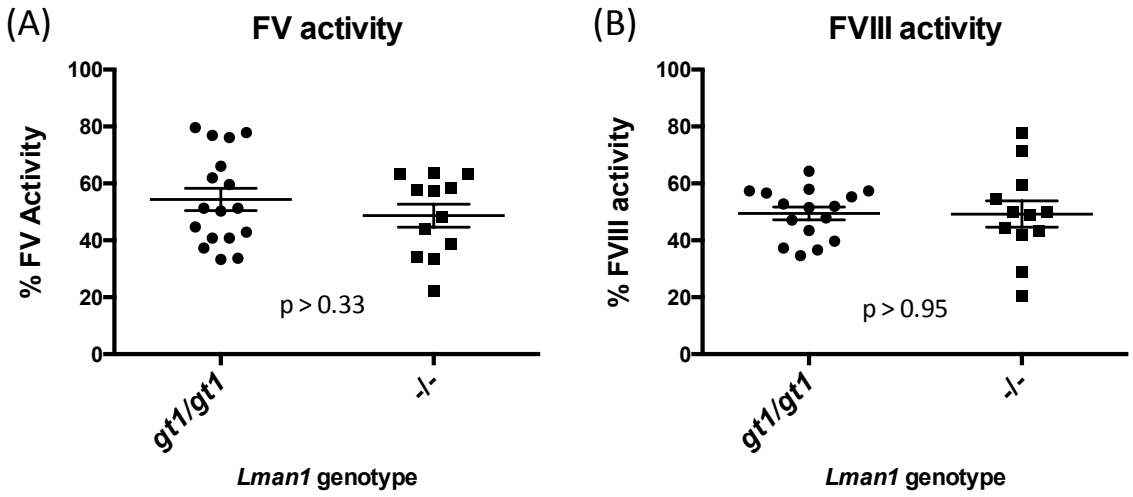


Figure 3-11: Plasma FV and FVIII activity levels are indistinguishable for *Lman1*$^{gt1/gt1}$ and *Lman1*$^{-/-}$ mice

(A) Plasma FV activity and **(B)** plasma FVIII activity for *Lman1*$^{gt1/gt1}$ and *Lman1*$^{-/-}$ mice are indistinguishable. Each symbol represents an individual animal. Horizontal lines indicate mean and error bars indicate standard error of the mean for each genotype.

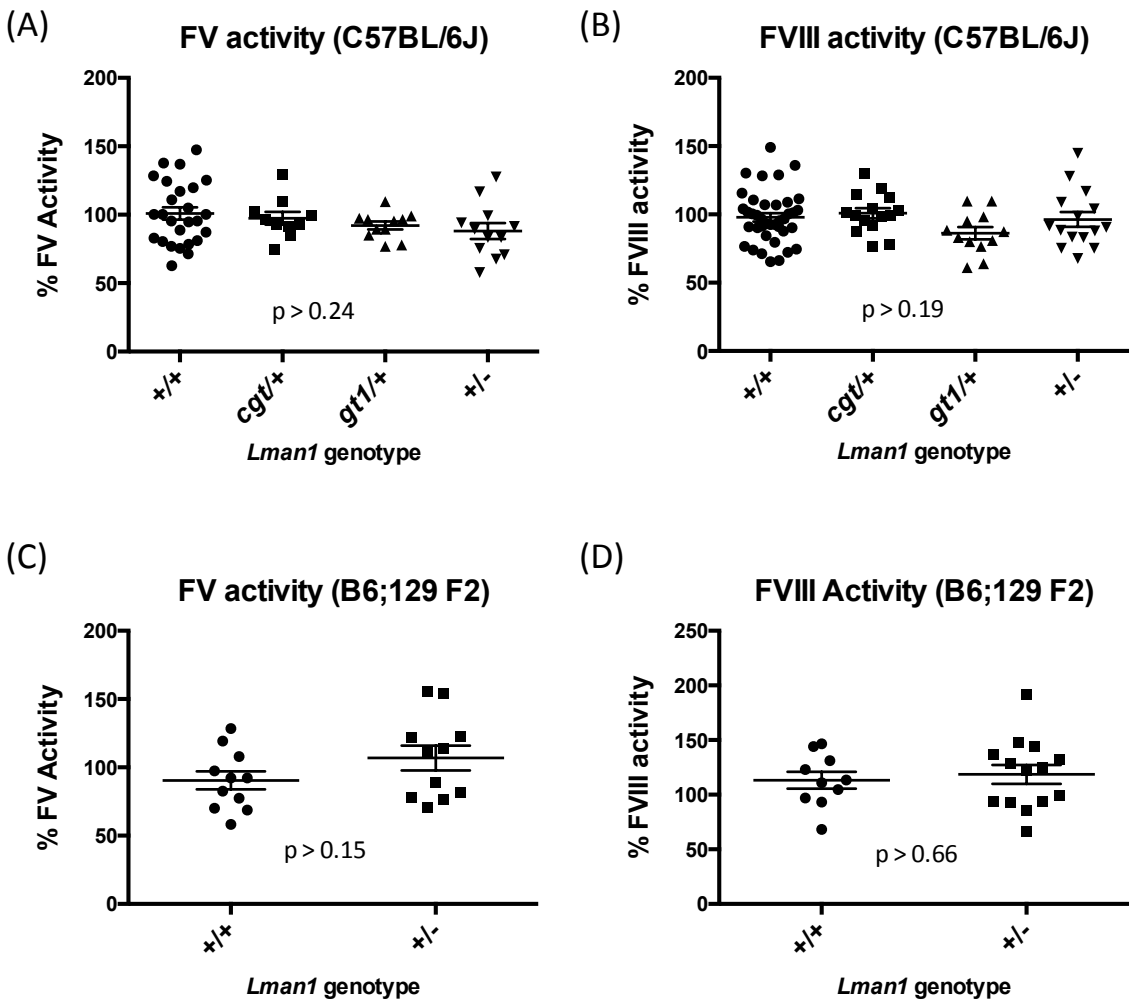


Figure 3-12: Plasma FV and FVIII activity levels for *Lman1* heterozygous mice on a C57BL/6J genetic background and mixed B6;129 genetic background are indistinguishable from wild-type controls

(A) Plasma FV activity and **(B)** plasma FVIII activity for heterozygous carriers of the $Lman1^{cgt}$ allele, the $Lman1^{gt1}$ allele, and the $Lman1^{-}$ allele on a C57BL/6J genetic background. Consistent with the autosomal recessive inheritance of F5F8D, *Lman1* heterozygous mice are indistinguishable from wild-type control mice. **(C)** Plasma FV activity and **(D)** plasma FVIII activity for $Lman1^{+/-}$ F2 mice (B6;129), relative to wild-type littermate controls. Consistent with the results from $Lman1^{+/-}$ mice on a pure C57BL/6J genetic background, $Lman1^{+/-}$ F2 mice (B6;129) have plasma FV and FVIII activity levels that are indistinguishable from wild-type control mice. Each symbol represents an individual animal. Horizontal lines indicate mean and error bars indicate standard error of the mean for each genotype.

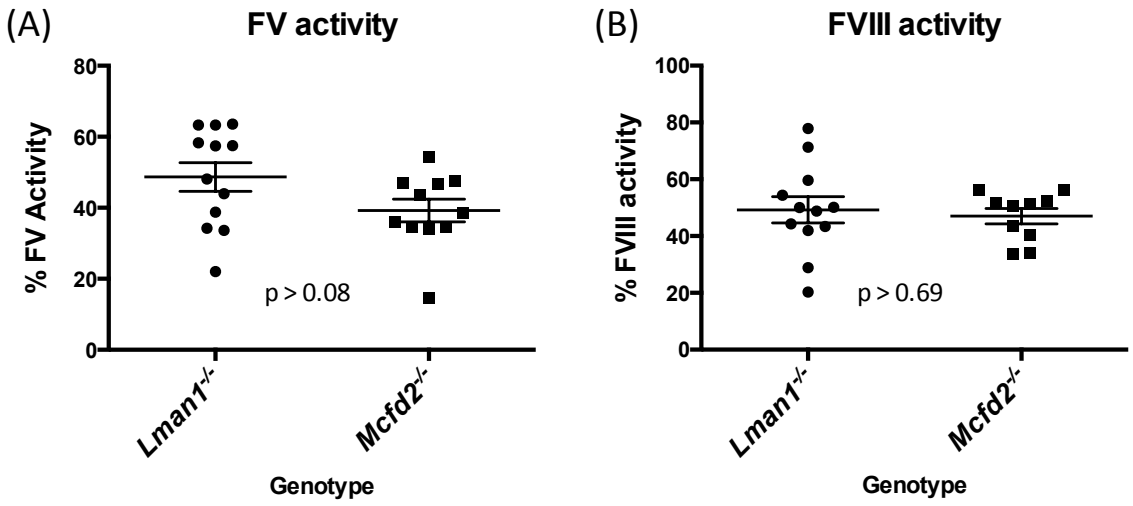


Figure 3-13: Plasma FV and FVIII activity for *Lman1*$^{-/-}$ and *Mcfd2*$^{-/-}$ mice

(A) Plasma FV activity and **(B)** plasma FVIII activity are indistinguishable between *Lman1*$^{-/-}$ mice and *Mcfd2*$^{-/-}$ mice. Each symbol represents an individual animal. Horizontal lines indicate mean and error bars indicate standard error of the mean for each genotype.

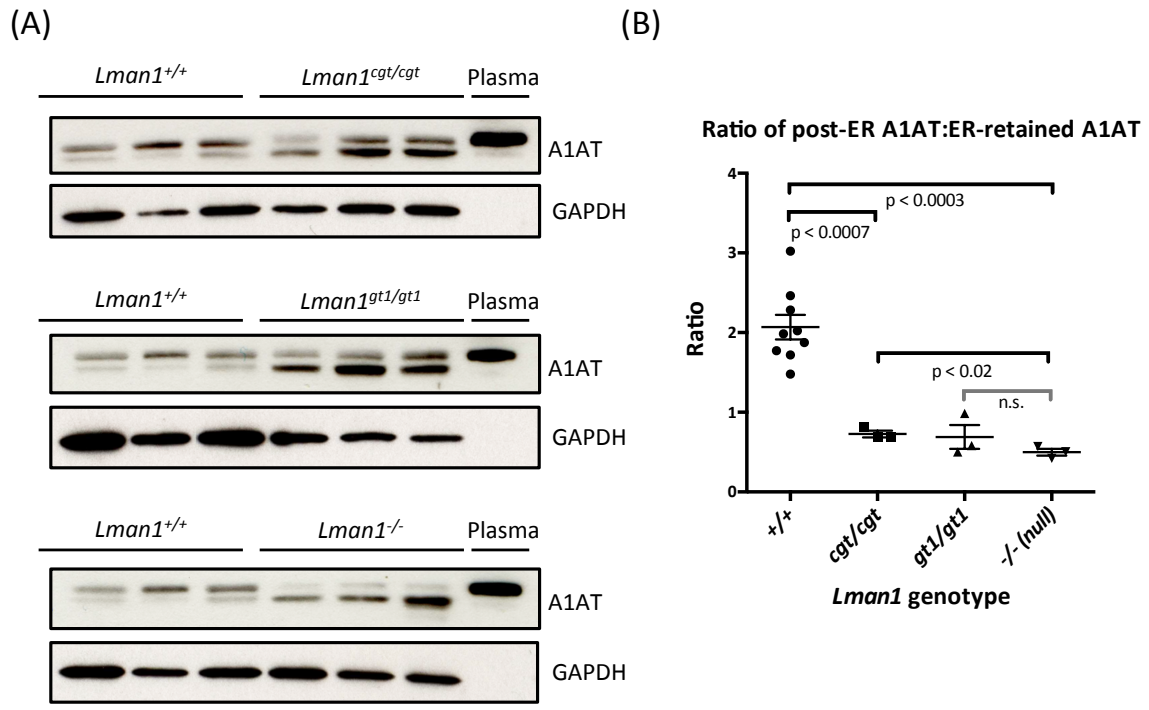
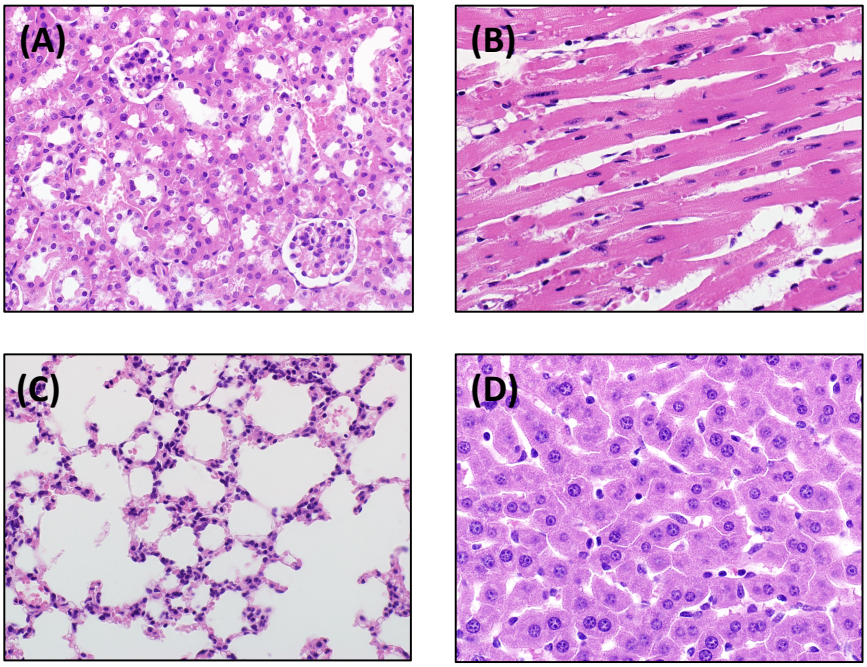


Figure 3-14: A1AT is retained in the ER of LMAN1 deficient mice

(A) Liver lysates from *Lman1*$^{cgt/cgt}$ mice, *Lman1*$^{gt1/gt1}$ mice, and *Lman1*$^{-/-}$ mice were compared to liver lysates from *Lman1*$^{+/+}$ control mice by Western blot to determine the extent of A1AT retention in the ER. The A1AT antibody detects a protein doublet on Western blot analysis. The lower band of the doublet corresponds to ER-retained A1AT, and was previously shown to be endoH sensitive.[95] The upper band in the doublet corresponds to post-ER A1AT, and is comparable in molecular weight to plasma A1AT. Each lane represents an individual mouse. **(B)** In order to quantify the retention of A1AT in the ER of *Lman1* mutant mice, relative to wild-type mice, the ratio of post-ER A1AT:ER-retained A1AT (upper band:lower band) was quantified for each sample using an IRDye-conjugated secondary antibody and the Odyssey Infrared Imaging System, with analysis performed on the Odyssey CLx Image Studio program. The ratio of post-ER A1AT:ER-retained A1AT is plotted for each animal. Ratio values < 1.0 indicate that the majority of A1AT in the liver sample is retained in the ER, and ratio values > 1.0 indicate that the majority of A1AT has undergone ER-to-Golgi transport (post-ER). The ratio of post-ER A1AT:ER-retained A1AT follows a similar relationship to *Lman1* gene expression levels as previously observed for FV and FVIII secretion in the *Lman1* mutant mice. *Lman1*$^{cgt/cgt}$ mice (~7% of normal *Lman1* expression) exhibit a slightly higher ratio of post-ER A1AT:ER-retained A1AT, as compared to *Lman1*$^{-/-}$ mice (zero *Lman1* expression). The ratios for *Lman1*$^{gt1/gt1}$ mice and *Lman1*$^{-/-}$ mice are indistinguishable. This indicates that *Lman1*$^{cgt/cgt}$ mice secrete more A1AT than do the *Lman1*$^{gt1/gt1}$ mice and *Lman1*$^{-/-}$ mice. The differences among the mean ratios were statistically significant as determined by one-way ANOVA ($F(3,14) = 23.24$, $p < 0.0001$). Horizontal lines indicate mean and error bars indicate standard error of the mean for each genotype.



Supplementary Figure 3-1: Adult *Lman1*$^{-/-}$ tissues are histologically normal

Adult *Lman1*$^{-/-}$ **(A)** kidney, **(B)** heart, **(C)** lung, and **(D)** liver tissues were evaluated by a pathologist blinded to mouse genotype. Images were taken using an Olympus BX43 microscope and an Olympus DP72 camera. Kidney and lung images are magnified 40X, heart and liver are 60X.

CHAPTER 4 :
PROTEOMIC APPROACHES TO IDENTIFY NOVEL LMAN1 FUNCTIONS

Note: Chapter 4 is a summary of our attempts to identify novel LMAN1 and MCFD2 interacting proteins through a number of in vitro and in vivo techniques. All experiments were performed by Lesley Everett. The Y2H assay was performed in collaboration with the Rual laboratory (UM Pathology). Mass-spectrometry was performed at the UM Pathology Department Proteomics facility, and bioinformatics analysis of the mass-spectrometry data was performed in collaboration with the Nesvizhskii laboratory (UM Pathology).

INTRODUCTION

Proteomic analysis of the secretory pathway

The protein secretory pathway is comprised of a number of specialized organelles essential for the folding, modification, and transport of secreted and membrane-bound proteins. Selective receptor-mediated ER-to-Golgi protein transport plays an important role in this pathway, but many of the protein-protein interactions and regulatory mechanisms utilized by ER cargo receptors are poorly defined. In particular, very little is understood about the specificity with which these receptors identify and recruit protein cargos from the ER lumen into budding COPII vesicles.[67,146] Furthermore, the mechanisms by which cargo receptors interact with cytoplasmic COPII vesicle components to facilitate proper cargo sorting are not well understood. It is likely that protein-protein interactions on both sides of the ER membrane affect the function of these ER cargo receptors. For example, SEC24 is the COPII coat component that interacts with transmembrane cargos and receptors in yeast and mammalian cells, and the transmembrane proteins serve as the communicating link between the inside and outside of COPII vesicles. Mammals have four SEC24 paralogs (A-D), and they are presumed to

fulfill distinct temporal or tissue-specific roles in the transport of protein cargos. This is consistent with the finding that mutations in individual COPII coat paralogs result in specific and distinct cellular defects and clinical disorders (Table 4-1).[67,147,148] These phenotypes may be due directly to the impaired secretion of a SEC24-specific transmembrane protein cargo. Alternatively, as in the case of SEC24A and the soluble PCSK9 cargo,[149] the defect may result from the impaired recruitment of an ER cargo receptor and its complement of soluble cargos (the ER cargo receptor for PCSK9 has yet to be identified). It is possible that the specific combination of SEC23 and SEC24 paralogs in a COPII coat may determine which ER cargo receptors and cargos are recruited into a budding vesicle. Although there is high-quality structural data revealing the general COPII coat framework and lattice structure,[150-156] little is known about the spectrum of possible interactions and potential binding preferences between the COPII coat component paralogs. The factors that determine the specific composition of COPII coat component paralogs in a given COPII vesicle are also unknown.

A more comprehensive understanding of the dynamic protein-protein interactions in ER-to-Golgi transport is required both to identify new cargo/receptor pairs and to identify the factors that determine which receptors and cargos are recruited into each COPII vesicle.[157] Numerous sensitive proteomic and bioinformatics techniques are now available to detect transient protein-protein interactions and to quantify impaired trafficking through the secretory pathway, complementing classical methods in genetics and molecular cell biology.[158-162] Quantitative proteomic methods such as SILAC, isotope-coded affinity tags, isobaric labeling, and label-free quantification can be used to determine quantitative differences in the abundance of protein cargos in different secretory compartments. High-throughput versions of classic genetic assays such as yeast two-hybrid (Y2H), fluorescent protein fragment complementation assays, and the split-ubiquitin method have been developed and are also suitable for these purposes.[163] These proteomic methods offer an unbiased, proteome-wide approach by which to dissect the components of the secretory pathway and the relationships among them. Most of these strategies have not yet been applied to the study of receptor-mediated cargo

transport and COPII vesicle trafficking. Such work will provide valuable insights into the function and regulation of ER cargo receptors.

Applications to the study of LMAN1 biology

Nearly 8,000 proteins traffic through the mammalian secretory pathway, and the majority are thought to rely on specialized ER cargo receptors for recruitment into COPII vesicles. It is therefore surprising that LMAN1/MCFD2 is the only known mammalian ER cargo receptor complex, and that only a handful of LMAN1-dependent cargos are known. LMAN1/MCFD2 functions to selectively recruit cargo proteins into budding COPII vesicles at the ER membrane for anterograde transport to the Golgi apparatus. The best characterized LMAN1-dependent secretory cargos are FV and FVIII, and three less characterized LMAN1 cargos are A1AT and cathepsins C and Z (CTSC, CTSZ).[85,164] Several lines of evidence suggest that additional LMAN1 cargos and MCFD2-like adaptors remain to be discovered, as well as novel ER cargo receptors for secretion of other proteins. Mass-spectrometry analyses and proteomic profiling of secretory organelles, plasma membranes, and secreted proteins from wild-type and LMAN1-deficient mice should facilitate the identification of novel LMAN1 cargos. The identification of novel LMAN1-dependent cargos and interacting proteins is of fundamental importance for elucidating other proteins that may contribute to F5F8D pathophysiology, and for understanding the ubiquitous and essential cellular process of protein secretion.

Previous characterization of LMAN1's protein-protein interactions and cargo selectivity relied on traditional methods such as crosslinking and immunoprecipitation or fluorescent complementation assays.[40,85,164] The limited scope and focus of these studies precluded the global identification and interrogation of LMAN1's full complement of binding partners. A global proteomics-based approach offers an unbiased method by which to identify additional LMAN1/MCFD2 cargos. A number of valuable reagents are available for this work, including *in vitro* and *in vivo* F5F8D models. The currently known LMAN1 cargos should serve as convenient "built in" positive controls. In the absence of LMAN1, its dependent cargos are expected to accumulate in the ER and to be

underrepresented in the Golgi or other post-Golgi compartments such as lysosomes, the plasma membrane, and secreted proteins. Quantifiable differences in LMAN1 cargo abundance in downstream secretory compartments of wild-type and LMAN1 deficient mice should be detectable by mass-spectrometry.

The identification of new LMAN1 cargos may provide insights into the binding motif and/or the specific recognition sequence utilized by LMAN1 for cargo recruitment, and this in turn may be used to predict other novel LMAN1 cargos based on protein sequence and structure. There is no clear homology sequence or structural domain shared between the five known LMAN1 cargos (FV, FVIII, A1AT, CTSC, CTSZ). The only other protein with sequence homology to FV and FVIII is ceruloplasmin, a copper-binding glycoprotein involved in iron transport across the cell membrane. FV, FVIII, and ceruloplasmin are evolutionarily related structures that likely resulted from duplications of ancestral precursor genes.[165] FV and FVIII share the protein domain structure A1-A2-B-A3-C1-C2, and ceruloplasmin contains three FV/FVIII type A domains. The A domains of FV and FVIII display ~30% sequence identity, and there is a comparable sequence identity (38-42%) between the FV and ceruloplasmin A1, A2, and A3 domains.

Preliminary characterization of samples from *Lman1*$^{gt1/gt1}$ mice was performed in our lab using a mass spectrometry-based approach (unpublished work).[166] Isobaric tags (iTRAQ reagents) were used to quantitatively analyze proteins from plasma, pancreatic zymogen granules, and lectin-affinity purified liver tissue lysates. Although these preliminary analyses presented technical challenges, results suggested that differences between the protein "secretome" of *Lman1*$^{gt1/gt1}$ and wild-type mice should allow us to identify additional LMAN1-dependent cargo proteins.[166]

Considering LMAN1's protein-protein interactions on the cytoplasmic side of the ER membrane, little is known about the factors regulating its interactions with the COPII coat for recruitment into budding vesicles. LMAN1's cytosolic tail domain contains a di-phenylalanine ER-exit motif (FF) for recruitment into vesicles by the SEC24 COPII coat component. Four mammalian SEC24 paralogs exist (A-D), and it is unknown whether

LMAN1 physically interacts with all 4 paralogs. However, Wendeler *et al*. reported that siRNA knockdown of each SEC24 paralog individually in HeLa cells had no effect on the ER-to-Golgi transport of LMAN1, implying that LMAN1 does not have a strong binding preference (or dependence) upon a specific SEC24 paralog for recruitment into COPII vesicles.[167] Similarly, it is unknown whether the specific combination of COPII coat components determines the content of the trafficking vesicle. As a first step in addressing this question, it is necessary to determine the different possible binding combinations of the COPII coat component paralogs, particularly between the inner coat components SEC23 (SEC23A and SEC23B) and SEC24 (SEC24 A-D).

Preliminary attempts to address these questions, and to identify novel LMAN1 cargos, were performed with a combination of *in vitro* experiments as well as cell-fractionation and mass-spectrometry studies.

MATERIALS AND METHODS

Cell culture and co-immunoprecipitation

COS cells were cultured in DMEM (Gibco) supplemented with 5% FBS and 1% pen-strep (Gibco) according to standard protocols. Cell transfections were performed with FuGENE HD transfection reagent (Promega) according to the manufacturer's instructions. Expression vectors encoding the following epitope-tagged murine proteins were used: RFP-SEC24A, RFP-SEC24B, RFP-SEC24C, RFP-SEC24D, FLAG-SEC23A, FLAG-SEC23B, and FLAG-LMAN1. Study of the interactions of LMAN1 and the SEC24 paralogs was performed by co-transfecting FLAG-LMAN1 individually with each of the RFP-tagged SEC24 constructs. Similarly, interactions between the SEC23 and SEC24 paralogs were explored by co-transfecting each FLAG-tagged SEC23 construct individually with each RFP-tagged SEC24 construct. DSP crosslinking (Pierce) of cell lysates was performed for 30 minutes at 4°C according to the manufacturer's instructions. Anti-FLAG and anti-RFP immunoprecipitations were

performed with EZview Red ANTI-FLAG M2 affinity gel beads (Sigma) and agarose conjugated anti-RFP beads (MBL International), respectively, according to the manufacturer's instructions.

Western blots

Antibodies used in Western blotting experiments are: a polyclonal antibody against RFP (Abcam), a polyclonal antibody against FLAG (Cell Signaling), a monoclonal antibody against GM130 (Santa Cruz Biotechnology), a polyclonal antibody against Erp44 (Cell Signaling), a polyclonal antibody against EEA1 (Abcam), a polyclonal antibody against LAMP-1 (Developmental Studies Hybridoma Bank), a monoclonal antibody against Complex II (Invitrogen), a polyclonal antibody against Annexin V (Cell Signaling), and a polyclonal antibody against ceruloplasmin (Abcam).

Yeast two-hybrid

In collaboration with Dr. Jean-François Rual (University of Michigan Department of Pathology), a high-throughput Y2H experiment[168-170] was performed to identify novel LMAN1 and MCFD2 interacting proteins. Three protein fragments of interest were cloned into the pDest22 and pDest32 Invitrogen Gateway Recombination cloning vectors according to the manufacturer's instructions in order to generate Gal4 DNA-binding domain (DB-X) and Gal4 activation domain (AD-Y) hybrid proteins. The three protein fragments that were tested are shown in Figure 4-1 and include: (1) the LMAN1 luminal domain without signal peptide, (2) the LMAN1 cytoplasmic tail, and (3) MCFD2 without signal peptide. The library of potential interacting proteins used by the Rual laboratory for all high-throughput Y2H screens is a human ORFeome library (v5.1), a cloned set of protein-encoding open reading frames representing approximately 11,200 human genes.[171,172] Relevant positive controls included in this hORFeome library are MCFD2, FVIII, and A1AT. The three protein fragments of interest were tested by Y2H in a primary screening in the DB-X conformation (protein fragment as "bait") against mini-libraries, each containing a pool of 12 AD-Y fusion proteins, by yeast-mating in 96-well format. As a secondary screening, the DB-X proteins were tested individually against 12

AD-Y fusions of each mini-pool, for those mini-pools in which hits were observed in the primary screening. The detailed methods, experimental controls, and validation steps for this high-throughput Y2H platform were previously described.[169]

Lman1 deficient mice

Lman1$^{-/-}$ mice (described in Chapter 3) were used as the source of plasma and tissue samples for several proteomic experiments. Plasma samples from *Lman1*$^{-/-}$ mice and wild-type controls were collected and snap frozen as described in Chapter 2. Equal volumes of mouse plasma from individual animals were run on Novex Bis-Tris 4-12% gels (Invitrogen) with MOPS buffer (Invitrogen), and the gels were subsequently stained with SYPRO Ruby Protein Gel Stain (Invitrogen) for fluorescent protein detection according to the manufacturer's instructions. Gels were imaged on a Typhoon 9410 Molecular Imager (Amersham Biosciences). Western blot analysis of murine plasma samples was performed by SDS-PAGE and transfer to nitrocellulose membranes. Tissues were collected as described in Chapter 3, except that samples were processed immediately (see "Cell Fractionation") rather than frozen. Mouse liver tissue samples were prepared from two mice per genotype. Littermate controls were used to reduce the potential effects of different diet, age, and housing conditions between experimental mice.

Cell fractionation

Collaborations with several University of Michigan faculty members (Yanzhuag Wang, Haoxing Xu, and Alexey Nesvizhskii) were established to optimize secretory organelle purification techniques and to conduct bioinformatics analyses of proteomics data. Standard differential density centrifugation protocols for ER, Golgi, and lysosome fractions of cultured cells and murine tissues were utilized.[173-175] Since two known LMAN1 cargos are lysosomal cathepsins (CTSC and CTSZ) and could therefore serve as a positive control in the cell fractionation prep, in preliminary studies we focused on the optimization of lysosome purification from mouse liver tissues. Multiple variables were tested in the optimization of this protocol, including fresh vs. frozen liver samples, the

extent of tissue homogenization, the volume of post-nuclear supernatant applied to the gradient, and the storage conditions of the final isolated lysosome prep. Briefly, samples were homogenized and lysed in HM Buffer (0.25M sucrose, 1mM EDTA, 10mM HEPES, pH 7.0), and the resulting lysate was centrifuged at 4200 rpm for 10 minutes at 4°C. The post-nuclear supernatant was applied to an 18% Percoll (Sigma) solution in HM buffer and centrifuged for 1 hour at 31,300 rpm at 4°C to separate the membrane fractions into light, medium, and heavy populations. The heavy membrane population was centrifuged again at 14,000 rpm for 10 minutes to concentrate and pellet the remaining organelles, which were subsequently applied to an Iodixanol (Sigma) gradient (8-27%) to separate lysosomes and mitochondria (the two primary components of the heavy membrane fraction). The resulting Iodixanol gradient fractions were analyzed by Western blot for relative abundance of lysosome and mitochondrial content, and the fractions with the highest enrichment for lysosomal proteins were applied to SDS-PAGE gels for electrophoresis. Gel lanes were cut into 5-8 pieces and submitted for mass-spectrometry.

Protein Identification by LC-Tandem Mass-spectrometry

Mass-spectrometry analysis was performed at the Fred Hutchinson Cancer Research Center Proteomics Facility or at the University of Michigan Molecular Diagnostics Laboratory. Purified samples (lysosome fractions or platelet-poor plasma) were resolved on a SDS-PAGE and proteins visualized using Colloidal Coomassie stain. Unique bands were excised and destained with 30% methanol for 4 h. Upon reduction and alklylation of the cysteines, proteins were digested overnight with sequencing grade, modified trypsin (Promega). Resulting peptides were resolved on a nano-capillary reverse phase column (Picofrit column, New Objective) using a 1% acetic acid/acetonitrile gradient at 300 nl/min and directly introduced in to a linear ion-trap mass spectrometer (LTQ XL, ThermoFisher). Data-dependent MS/MS spectra on the 5 most intense ion from each full MS scan were collected (relative CE ~35%). Proteins were identified by searching the data against Human IPI database (v 3.41) appended with decoy (reverse) sequences using X!Tandem/Trans-Proteomic Pipeline (TPP) software suite. All proteins with a ProteinProphet probability score of >0.8 (error rate <1%) were considered positive

identifications and manually verified. Only proteins with 2 or more unique peptides were included in the analysis. The Abacus computational tool for spectral count data for label-free quantitative proteomic analysis was used to normalize the spectral count data from each of the plasma band samples.[176] Preliminary statistical analysis of the normalized spectral counts was performed with a student's t-test, with $p < 0.05$ considered statistically significant. For a more appropriate and useful analysis of this data, the normalized spectral counts for the plasma samples were also submitted to QSPEC (http://www.nesvilab.org/qspec.php/), a software program for the analysis of differential protein expression/abundance using label-free spectral count data. QSPEC calculates the ratio of likelihoods (Bayes Factor) for differential expression/abundance for each protein based on certain model assumptions (Poisson-family distributions for count data and Gaussian distributions for intensity data). The QPSEC method is more appropriate than conventional statistic-based tests such as the t-test for our data given that there are only three replicate samples per genotype.

As part of my graduate training, I completed a graduate course in Proteome Informatics (taught by Dr. Alexey Nesvizhskii), as well as a weeklong proteomics and mass-spectrometry course at the Fred Hutchinson Cancer Research Center Proteomics Facility.

RESULTS

LMAN1 interacts with all four mammalian SEC24 paralogs

As demonstrated by co-expression and co-immunoprecipitation experiments *in vitro,* LMAN1 physically interacts with all four murine SEC24 paralogs (A-D) in approximately equal proportion (Figure 4-2). Similarly, both murine SEC23A (Figure 4-3A) and SEC23B (Figure 4-3B) paralogs physically interact with each of the four murine SEC24 paralogs (A-D), also in essentially equal proportions. Therefore, eight possible SEC23-SEC24 combinations are available for the generation of the heterodimer pairs that compose the inner COPII coat membrane (SEC23A-SEC24A, SEC23A-

SEC24B, SEC23A-SEC24C, SEC23A-SEC24D, SEC23B-SEC24A, SEC23B-SEC24B, SEC23B-SEC24C, SEC23B-SEC24D).

Novel LMAN1/MCFD2 interacting proteins were not identified by yeast two-hybrid

Protein fragments corresponding to the LMAN1 luminal domain, the LMAN1 cytosolic tail domain, and the MCFD2 protein were tested in a high throughput Y2H assay against a human ORFeome library of potential interacting proteins. In the primary screening of each fragment against mini-pools of 12 ORFs per pool, five potential hits were identified for the LMAN1 cytosolic domain. The identify of these potential interacting proteins was not determined in the primary screen because the ORFs are screened in groups of 12; one of the ORFs in a mini-pool of 12 ORFs was responsible for each of the primary screen hits. However, upon more stringent secondary screening of the LMAN1 cytosolic tail domain individually against each member of the ORF mini-pools for direct binary interactions, no hits were confirmed. The secondary screening was performed twice to confirm this result. No hits were identified in the primary screening for either the LMAN1 luminal domain or the MCFD2 fragment. Of note, no known "positive control" protein-protein interactions were identified in this Y2H screen, including LMAN1-MCFD2 interactions, LMAN1-FVIII interactions, MCFD2-FVIII interactions, or LMAN1-A1AT interactions.

Comparison of serum proteins in wild-type and *Lman1*$^{-/-}$ mice

Platelet-poor plasma from wild-type and *Lman1*$^{-/-}$ mice was compared by SDS-PAGE, and a protein of ~ 100 kDa molecular weight was observed to be of lower abundance (~50% or less than observed in wild-type) in *Lman1*$^{-/-}$ mouse plasma samples as compared to that of wild-type samples (Figure 4-4). Since ceruloplasmin, a potential candidate LMAN1 cargo with homology to FV and FVIII, is also ~100 kDa, we compared its abundance in the plasma of wild-type and *Lman1*$^{-/-}$ mice by Western blot. Plasma ceruloplasmin levels were indistinguishable between wild-type and *Lman1*$^{-/-}$ mice (Figure 4-5). The ~100 kDa band of interest (indicated by a black arrow in Figure 4-4) was excised from three wild-type and three *Lman1*$^{-/-}$ lanes for mass-spectrometry

analysis. The normalized spectral counts for each of the 24 proteins that were detected in 3 or more of the six excised gel bands are listed in Table 4-2. Based on a preliminary comparison of the spectral counts for each protein in *Lman1*$^{-/-}$ vs. *Lman1*$^{+/+}$ samples (by t-test), only one protein in the list, SERPINA3K, was found to have a statistically significantly lower number of spectral counts in the *Lman1*$^{-/-}$ plasma samples as compared to the wild-type plasma samples ($p < 0.05$). Complement Component C3 was also marginally significantly reduced in *Lman1*$^{-/-}$ plasma samples as compared to wild-type samples. For a more appropriate and computationally rigorous analysis, the normalized spectral counts for the plasma proteins were also analyzed by QSPEC. According to QSPEC, SERPINA3K and Complement Component C3 were both highly significantly under-represented in the *Lman1*$^{-/-}$ plasma samples as compared to the wild-type samples (Table 4-3), and these were the only two proteins that were under-represented in the *Lman1*$^{-/-}$ plasma samples. To further investigate these proteins as putative LMAN1 cargos, we are currently comparing the abundance of SERPINA3K and C3 in wild-type and *Lman1*$^{-/-}$ mouse plasma samples by Western blot.

Optimization of cell fractionation and organelle isolation

Lysosome isolation (Figure 4-6A) from cultured cells (not shown) and mouse tissues was optimized. Cellular membrane fractions were successfully separated by differential density centrifugation such that lysosome-enriched, mitochondrial-depleted fractions were consistently isolated (Figure 4-6B). When these lysosome-enriched fractions were pooled and then subjected to Western blot analysis to confirm the enrichment and purity of the sample, the preps were demonstrated to be highly enriched for lysosomal proteins and largely free of golgi, ER, endosome, mitochondrial, and plasma membrane protein contaminants (Figure 4-7). The purified lysosome samples were separated by SDS-PAGE (Figure 4-8) and were submitted for mass spectrometry analysis. Data analysis was performed with Dr. Alexey Nesvizhskii's bioinformatics laboratory at the University of Michigan.[177,178] Unfortunately, despite the high lysosomal enrichment and the lack of other cellular contaminants observed for these lysosome preps by Western blot analysis (Figure 4-7), bioinformatics analysis of the mass-spectrometry data revealed a surprising relatively small number of distinct lysosomal proteins as compared to other membrane-

associated and secretory proteins in these submitted samples. Several well-established lysosomal marker proteins were not detected in the samples, or detected at very low abundance. For example, only 3-5 spectral counts were identified in each sample for lysosome-associated membrane glycoprotein-1 (LAMP1), and only 5-6 spectral counts were identified in each sample for lysosome-associated membrane glycoprotein-2 (LAMP2). Similarly, both wild-type and *Lman1*$^{-/-}$ samples had a comparable low number of spectral counts for cathepsin Z (1-2 counts), cathepsin H (2-3 counts), and cathepsin D (5-16 counts). Cathepsin C was not detected in either the wild-type or *Lman1*$^{-/-}$ samples. The low abundance of lysosomal proteins and the presence of numerous other secretory proteins (many of which do not have cellular localization information listed in the protein databases) precluded the meaningful statistical comparison of these samples and the identification of any lysosomal proteins that were potentially differentially secreted in wild-type and *Lman1*$^{-/-}$ mice.

DISCUSSION

A combination of simple *in vitro* experiments and more technically challenging cell-fractionation and mass-spectrometry studies were performed for the purpose of identifying novel LMAN1-dependent secretory cargos, and to characterize LMAN1's interactions with the inner layer of COPII vesicle coats. These experiments were met with variable success, and emphasize that in order to develop a complete picture of LMAN1's interacting partners and dependent secretory cargos, we will likely require the detection of subtle quantitative differences in protein trafficking and the identification of transient protein interactions within secretory compartments.

We initially demonstrated by co-immunoprecipitation that LMAN1 physically interacts with all four mammalian SEC24 paralogs in approximately equal proportion. This finding suggests that the recruitment of LMAN1 and its cargos to budding COPII vesicles via interaction with the SEC24 COPII coat component is not likely a selective sorting step. This finding is consistent with the results of Wendeler *et al.*'s siRNA knockdown

of the SEC24 paralogs, which revealed no difference in ER-to-Golgi transport of LMAN1 in the absence of any individual SEC24 paralog.[167] Although LMAN1 does not demonstrate any SEC24 binding specificity, it is possible that other ER cargo receptors do use SEC24 binding as a sorting mechanism. We also demonstrated by co-immunoprecipitation that 8 possible SEC23-SEC24 heterodimer building blocks are available for the synthesis of new COPII vesicles. SEC23A and SEC23B are both able to form heterodimers with all four SEC24 paralogs (A-D). The biologic significance and functional differences between these 8 different SEC23-SEC24 heterodimer pairs is currently unknown. However, given the highly specific and distinct phenotypes that result from deficiency of these individual COPII coat component paralogs (Table 4-1), it is clear that these proteins (and perhaps the combinations in which they interact) play critical and complex roles in protein secretion that are essential for the maintenance of cell function.[149,179-182] Since LMAN1 binds all four SEC24 paralogs, it presumably binds all 8 heterodimer pairs, although we did not test this directly.

A Y2H assay was performed in collaboration with the Rual laboratory (UM Pathology Department) to identify novel LMAN1 and MCFD2 interacting proteins. Although the Y2H approach has its own technical limitations, especially for transmembrane proteins, recent advances in the field have enabled the development of high-quality, efficient, and reliable binary protein-protein interactome mapping strategies.[183,184] Y2H has led to the characterization of thousands of binary protein interactions and has been extremely useful for interrogating developmental and disease mechanisms. The Rual laboratory utilizes a unique Y2H protein interaction mapping platform[169] that can be used in both 1) the *discovery* mode to uncover new protein interactors and novel molecular mechanisms in which the proteins of interest are involved; and 2) the *validation* mode to dissect the molecular architecture of specific protein complexes by systematically testing pairs of interactors and by refining the precise molecular organization of the protein complexes, *i.e.*, determining which proteins (or domains) interact directly with each other in a given complex. Dr. Rual's laboratory has utilized this platform to systematically screen 25 million pairs of proteins for their ability to physically interact with each other, resulting in the detection of 2,800 high-quality Y2H protein interactions.

Though these methods would seem ideal for the purpose of identifying and confirming novel LMAN1- or MCFD2- interacting proteins, we did not detect any reproducible interactions that passed the second, more stringent screening test for either LMAN1 or MCFD2. Similarly, no known LMAN1 interacting proteins (MCFD2, FVIII, or A1AT) were identified in this screen. This negative result is not unexpected, given that Y2H is not an extremely sensitive method, and typically only about half of the baits tested generate a hit in the primary screen.[185] Also, Y2H has a low sensitivity for membrane proteins.[186] Since Y2H is based upon the reconstitution of a transcription factor, all protein-protein interactions must occur within the nucleus. Membrane proteins must therefore be tested in fragments (each as a DB-X or AD-Y fusion) that localize to the nucleus. As a result, some membrane protein fragments are misfolded or lack critical domains or residues necessary for their endogenous protein-protein interactions. Differences between the nuclear environment and the membrane protein's endogenous cellular location may prevent it from interacting with binding partners. In the case of LMAN1 and MCFD2, it is likely that the nuclear environment is insufficient to allow the protein-protein interactions that usually take place within the ER lumen.

Another high throughput strategy to identify protein-protein interactome networks in a more direct manner is affinity purification followed by mass spectrometry (AP-MS), which detects biochemically stable, copurifying protein complexes. Both direct and indirect protein associations may be detected by AP-MS. However, AP-MS is not well suited to membrane proteins, which are particularly difficult to isolate due to their hydrophobic nature. In addition, the biochemical isolation of membrane proteins such as LMAN1 usually disturbs any associated proteins in complex with the membrane protein.[187]

In addition to these *in vitro* experiments, our LMAN1 deficient mouse models serve as a valuable resource for proteomics samples, particularly the plasma and secretory tissues. We separated platelet-poor plasma samples by SDS-PAGE and observed an ~100 kDa band that was visibly lower in abundance in the plasma of *Lman1*$^{-/-}$ mice as compared to

wild-type mice. This band could potentially correspond to a plasma protein that depends on LMAN1 for secretion (i.e. a novel cargo). It could also potentially represent a protein which is differentially present in the *Lman1*$^{-/-}$ plasma samples as an indirect result of LMAN1 deficiency, but that is not an LMAN1 cargo itself.

Considering the major plasma glycoproteins of ~100 kDa molecular weight, ceruloplasmin was the most obvious potential LMAN1 cargo candidate. Ceruloplasmin shares a significant sequence and domain structure homology with FV and FVIII. However, plasma ceruloplasmin levels were indistinguishable between wild-type and *Lman1*$^{-/-}$ mice. Ceruloplasmin does not contain a glycosylated B domain (as do FV and FVIII), and the B domain has been shown to be required for efficient LMAN1-FVIII interactions,[42] potentially explaining why ceruloplasmin is not also an LMAN1 cargo. Consistent with our data, it was previously shown that ceruloplasmin does not co-immunoprecipitate with LMAN1 or MCFD2.[32,42] In addition, low plasma ceruloplasmin levels have not previously been reported in F5F8D patients.[188]

In order to identify the ~100 kDa protein of interest and to determine if there is a statistically significant difference in its abundance in the plasma between wild-type mice and *Lman1*$^{-/-}$ mice, three gel bands for each genotype were submitted for LC-MS/MS analysis. Of the ~30 proteins reliably detected in 3 or more of the 6 submitted samples, only two proteins demonstrated a statistically significant difference between the wild-type and the *Lman1*$^{-/-}$ samples (Table 4-3). The first protein that was detected in significantly lower abundance in *Lman1*$^{-/-}$ samples as compared to *Lman1*$^{+/+}$ samples was SERPINA3K, also known as serine protease inhibitor A3K. Serine protease inhibitors (otherwise known as SERPINS) are highly abundant in the plasma, and most are targeted toward the serine proteases involved in phagocytosis, coagulation, complement activation, and fibrinolysis. We would expect this SERPIN to likely be identified in complex with its target protease, since protease:protease inhibitor complexes are irreversibly bound and would remain in complex on an SDS-PAGE gel. Therefore, we would expect that SERPINA3K's target protease could also potentially be identified in the mass-spectrometry samples. Consistent with this expectation, the second under-

represented protein in the *Lman1*$^{-/-}$ plasma samples was Complement Component C3, a serine protease. As far as we are aware, SERPINA3K and Complement Component 3 are not known to function together, but many protease:protease inhibitor combinations are not currently known, or have been inaccurately catalogued.

SERPINA3K is a particularly interesting candidate LMAN1 cargo. It is a member of the "clade A" serpins, which comprises alpha-1 antitrypsin (SERPINA1) and alpha-1 antichymotrypsin (SERPINA3). Whereas alpha-1 antitrypsin and alpha-1 antichymotrypsin are each encoded by a single gene in humans, in the mouse they are represented by clusters of 5 and 14 genes, respectively, resulting from multiple gene duplication events.[189] Of interest, alpha-1 antitrypsin (SERPINA1) is one of LMAN1's reported cargo proteins, in addition to FV and FVIII.[85] The duplication and divergence of these clade A SERPINS in mice as compared to humans (and their potential differences in LMAN1-dependent secretion) represent an intriguing possible explanation for the different phenotypes observed in mice and humans with LMAN1 deficiency. To further assess SERPINA3K and/or Complement Component C3 as putative LMAN1 cargos, we are currently comparing the abundance of SERPINA3K and C3 in wild-type and *Lman1*$^{-/-}$ plasma samples and liver lysates.

Finally, our preliminary attempts at cell fractionation and secretory organelle isolation for mass-spectrometry were uninformative, and further optimization is required before we can put this approach into practice. For example, as demonstrated by Western blot analysis, the lysosome isolation protocols we utilized resulted in highly enriched, nearly pure lysosomal fractions that appeared to be of sufficient purity and enrichment for useful mass-spectrometry analysis. However, the protein identifications resulting from these samples demonstrated surprisingly few lysosomal proteins, and were convoluted by many other non-lysosomal secretory and membrane proteins. In the future, it may be useful to develop a more stringent standard for ensuring that the lysosome fractions are free of other cellular contaminants. In addition, we may consider using exclusively cultured cells rather than animal tissues for organelle isolations, since this is a cleaner system for

cell fractionation. However, the limitations of using an *in vitro* rather than *in vivo* system include the fact that (1) cells cultured *in vitro* may behave differently in a culture dish than in the physiologic context of an entire organism, (2) *in vitro* culture is an artificial system in which cells are not exposed to the same signals and stimuli as they are in the context of living organisms (i.e. cytokines and hormone signaling, sheer stress and flow conditions in a vessel, etc.), and (3) *in vitro* culture conditions cannot capture inter-animal variation.

Given that the protein secretion defect in LMAN1-deficient mice (and in humans with *LMAN1* mutations) is expected to vary between individuals (see Chapter 3), we ultimately need a precise and sensitive method by which to detect and quantify potentially subtle defects in protein transport. One approach that would provide quantitative information about the relative abundance of potential LMAN1 cargos in the secretory pathway is SILAC labeling[190] (stable isotope labeling by amino acids in cell culture), which provides highly accurate quantitation for cells grown in culture[160] as well as for metabolic labeling of whole animals.[191] 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. Our lab has successfully used SILAC in cell culture, and we have also generated a set of $^{13}C_6$-lysine wild-type C57BL/6J mice that may be used as controls for quantitative MS/MS comparisons with *Lman1*$^{-/-}$ mice in the future.

Subsequently, any putative LMAN1-dependent cargos will be experimentally validated to demonstrate cellular co-localization with LMAN1 (immunohistochemistry, immunofluorescence) as well as physical cargo-receptor interactions (immunoprecipitation and quantitative western blots, chemical crosslinking). In addition, several experimental models of F5F8D (including knockout mice, immortalized patient lymphocyte cell lines, *Lman1*$^{-/-}$ murine embryonic fibroblasts, and a dominant-negative LMAN1 cell line) are available to assess and quantify impaired trafficking of putative cargos of interest. The dominant-negative LMAN1 is a mutant form of LMAN1 that is incapable of leaving the ER.[87] In cells expressing this dominant-negative mutant, known

and novel LMAN1 cargos would be expected to accumulate in the ER (and to be under-represented in downstream secretory compartments) as compared to their localization in cells expressing wild-type LMAN1.

We anticipate that in the future, with further optimization of cell fractionation, it will be possible to compare the proteome of purified ER, Golgi, lysosome, plasma membrane, and secreted proteins for wild-type mice, *Lman1*$^{-/-}$ mice, *Mcfd2*$^{-/-}$ mice, and doubly deficient *Lman1*$^{-/-}$, *Mcfd2*$^{-/-}$ mice. We expect to identify new cargos that rely on LMAN1 for efficient ER-to-Golgi transport. We also expect to confirm several previously suspected or poorly characterized LMAN1 cargos such as CTSC, CTSZ, and A1AT.

Such work will be necessary to elucidate LMAN1's complete role in the secretory pathway. Further, it will provide a model for understanding selective receptor-mediated transport by other ER receptors, and will help explain why defects in protein trafficking lead to disease.

Notable challenges still remain in the field of organelle proteomics, such as how to most accurately measure changes in protein distribution under different cellular conditions or genetic backgrounds. However, technological and analytic advances in organellar proteomics have increased the reliability and resolving power of MS-based studies. Ultimately, these advances will enable us to make unbiased, proteome-wide characterizations of organellar proteomic profiles and to develop a more comprehensive, accurate picture of the subcellular localization of entire proteomes under different experimental conditions.[158] These developments have important basic science implications, as well as clinical and translation relevance. Ultimately, the proteomic characterization of secretory organelles and of ER receptor/cargo pairs may enable the therapeutic restoration of individual pathway defects to alleviate a wide range of human inherited diseases associated with protein misfolding and trafficking, including F5F8D.[147,158]

Table 4-1: Congenital disorders of the early secretory pathway

Type	*S. cerevisae*	Mammal	Human disease or Mouse Phenotype
COPII vesicle	*Sar1p*	*Sar1a* *Sar1b*	Embryonic lethal (E<10.5) CRD; Lethal by weaning
	Sec23p	*Sec23a* *Sec23b*	CLSD; Embryonic lethal CDAII: Perinatal lethality
	Sec24p *Lst1p* *Iss1p*	*Sec24a* *Sec24b* *Sec24c* *Sec24d*	Low cholesterol Late embryonic lethal, neural tube defect Early embryonic lethal Early embryonic lethal
	Sec13p	*Sec13*	None known
	Sec31p	*Sec31a* *Sec31b*	None known None known
Cargo receptor/adaptor	NA	*Lman1*	F5F8D: F5F8D, perinatal lethality
	NA	*Mcfd2*	F5F8D; F5F8D

CRD, Chylomicron retention disease; CLSD, Cranio-lenticulo-sutural dysplasia; CDAII, Congenital dyserythropoietic anemia type II; F5F8D, Combined deficiency of coagulation factors V and VIII

Table 4-2: Spectral counts for proteins identified in excised gel bands from mouse plasma samples

Protein	*Lman1*$^{-/-}$ (#1)	*Lman1*$^{-/-}$ (#2)	*Lman1*$^{-/-}$ (#3)	*Lman1*$^{+/+}$ (#1)	*Lman1*$^{+/+}$ (#2)	*Lman1*$^{+/+}$ (#3)	p-value
ITIH3	9.805	3.793	0.000	2.619	0.000	0.000	0.287
C4B	2.629	0.000	2.243	3.687	2.349	2.975	0.203
A2M	22.738	8.210	21.556	22.474	26.407	26.078	0.196
CFH	6.174	3.981	7.024	9.896	7.354	5.590	0.289
FGA	10.134	3.360	6.588	9.281	12.933	3.932	0.569
ITIH4	88.610	49.775	0.000	75.873	0.000	82.429	0.866
CP (ceruloplasmin)	9.466	3.255	9.572	8.429	8.456	5.714	0.966
ALB	127.748	78.491	164.563	87.319	80.559	69.741	0.156
C5	8.839	2.630	15.469	10.897	12.148	14.774	0.403
SERPINA3K	273.259	205.081	121.236	289.055	429.646	403.291	0.047
SERPINA3M	16.396	16.914	9.326	8.759	19.529	14.845	0.969
MUG1	10.834	7.784	12.325	9.922	15.901	11.912	0.363
FGB	4.749	3.675	5.403	15.224	6.364	8.601	0.114
ITIH1	12.593	3.898	1.433	4.037	2.250	3.421	0.469
ITIH2	7.245	3.737	15.109	6.773	2.157	5.466	0.344
TF	31.137	30.430	54.064	30.205	32.208	23.741	0.296
C3	0.000	29.759	53.133	73.756	57.678	97.015	0.064
GPLD1	0.000	9.446	13.889	10.871	4.848	15.968	0.621
VTN	2.390	3.698	0.000	0.000	0.000	2.164	0.372
LIFR	12.552	7.284	0.000	10.059	10.279	0.000	0.975
DOCK8	0.544	0.421	1.238	0.436	0.972	0.493	0.759
AMBP	6.546	5.065	3.723	7.868	0.000	0.000	0.416
FGG	13.099	4.054	5.961	8.398	9.362	0.000	0.682
HRG	4.262	0.000	2.424	0.000	3.808	0.000	0.617

The spectral counts for each protein that was identified in at least 3 of the 6 excised gel bands submitted for mass-spectrometry are listed. Spectral counts were normalized with the Abacus computational tool. A student's t-test was used for statistical analysis, with $p < 0.05$ considered statistically significant. Only one protein, SERPINA3K, was statistically significantly lower in abundance in the *Lman1*$^{-/-}$ plasma samples than in the *Lman1*$^{+/+}$ control samples in this analysis. Complement component C3 was also under-represented in *Lman1*$^{-/-}$ plasma samples (borderline significant).

Table 4-3: QSPEC analysis reveals two proteins that are significantly under-represented in *Lman1*$^{-/-}$ plasma as compared to *Lman1*$^{+/+}$ plasma

Protein	Z-statistic	False Discovery Rate	Function
Serine protease inhibitor A3K (SERPINA3K)	-4.703	0.000004	Protease inhibitor
Complement component 3	-4.198	0.000097	Serine protease

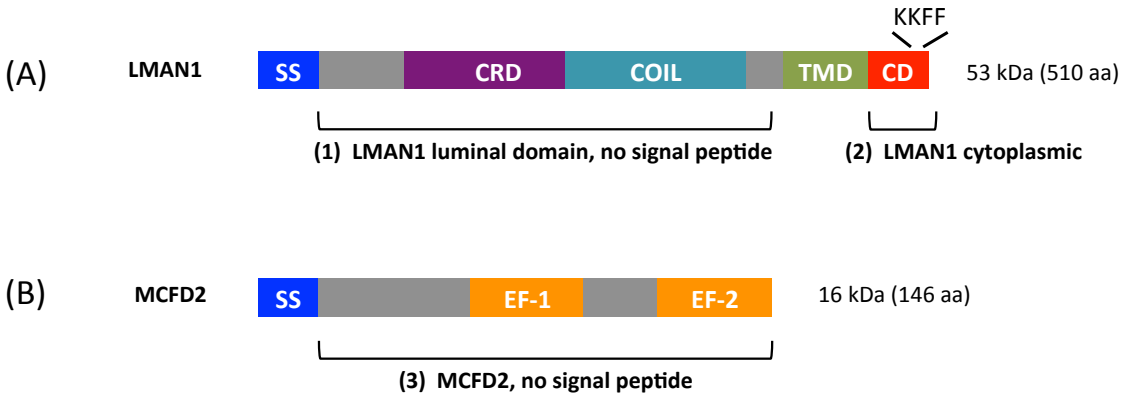


Figure 4-1: LMAN1 and MCFD2 protein fragments for yeast two-hybrid assay

(A) LMAN1 is a type 1 transmembrane protein. SS, signal peptide sequence; CRD, carbohydrate recognition domain; COIL, coiled coil domain; TMD, transmembrane domain; CD, cytoplasmic domain; KKFF, di-phenylalanine ER exit motif, di-lysine retrieval signal. The luminal domain and the cytoplasmic domain were tested as separate Y2H "baits" to eliminate technical difficulties associated with testing transmembrane proteins. **(B)** MCFD2 is a soluble protein with a signal peptide sequence (SS) and two EF-hand domains. It binds to LMAN1 via a Ca^{2+}-dependent interaction. The protein fragments tested by Y2H are indicated by brackets.

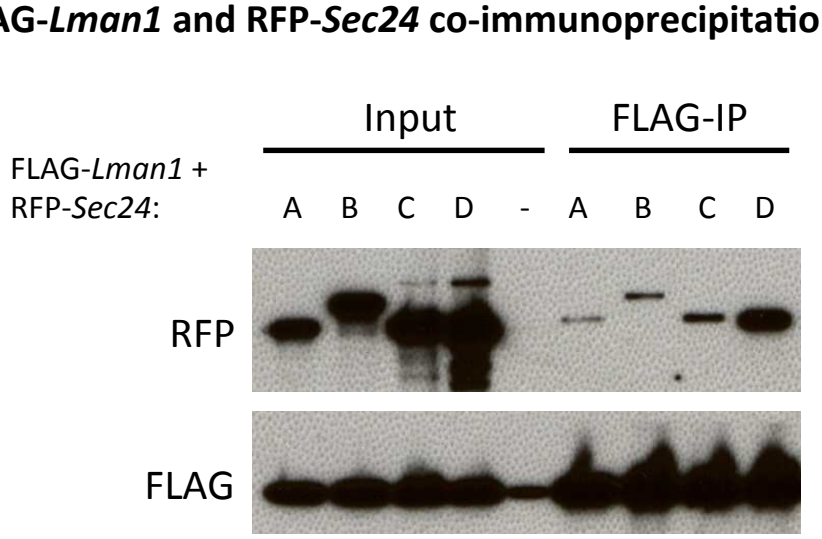


Figure 4-2: LMAN1 interacts with all four mammalian SEC24 paralogs

COS cells were co-transfected with FLAG-tagged LMAN1 and one of four RFP-tagged SEC24 paralogs (A-D). As indicated in lanes 1-4 of the Western blot, the co-transfections were successful, with high expression of both FLAG-LMAN1 and each RFP-SEC24. A negative control transfection (lane 5 under "Input," indicated by "-") was transfected only with FLAG-LMAN1. An anti-FLAG co-immunoprecipitation experiment was performed ("FLAG-IP" lanes) to isolate LMAN1 and its binding partners in complex. FLAG-LMAN1 was highly enriched and successfully precipitated from the total cell lysate. RFP-SEC24A (lane 6), RFP-SEC24B (lane 7), RFP-SEC24C (lane 8), and RFP-SEC24D (lane 9) co-immunoprecipitated with LMAN1, indicating that all four SEC24 paralogs physically interact with LMAN1.

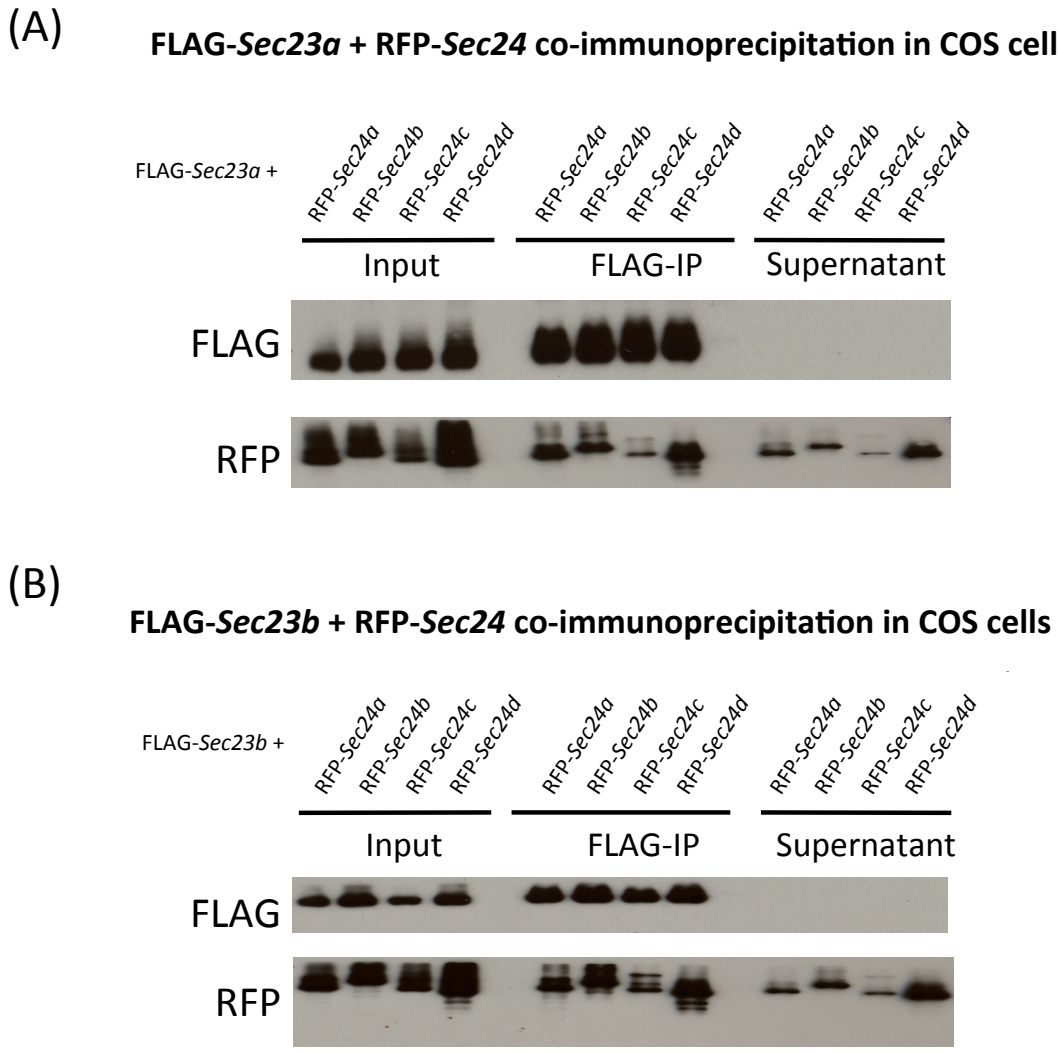


Figure 4-3: SEC23A and SEC23B both interact with all four SEC24 paralogs

COS cells were co-transfected with **(A)** FLAG-tagged SEC23A and each of four RFP-tagged SEC24 paralogs (A-D) individually, or with **(B)** FLAG-tagged SEC23B and each of the four RFP-SEC24 paralogs (A-D) individually. As indicated in lanes 1-4 of each Western blot ("Input"), FLAG-SEC23 and each RFP-SEC24 were successfully co-expressed. An anti-FLAG co-immunoprecipitation experiment was performed ("FLAG-IP" lanes) to isolate SEC23 and its interacting proteins. As demonstrated by Western blot analysis of the co-immunoprecipitation product ("FLAG-IP"), the FLAG-SEC23 proteins were highly enriched and successfully precipitated from the total cell lysate. Similarly, RFP-SEC24A (lane 5), RFP-SEC24B (lane 6), RFP-SEC24C (lane 7), and RFP-SEC24D (lane 8) each co-immunoprecipitated with SEC23A and SEC23B, indicating that each of the SEC24 paralogs physically interacts with both SEC23 paralogs.

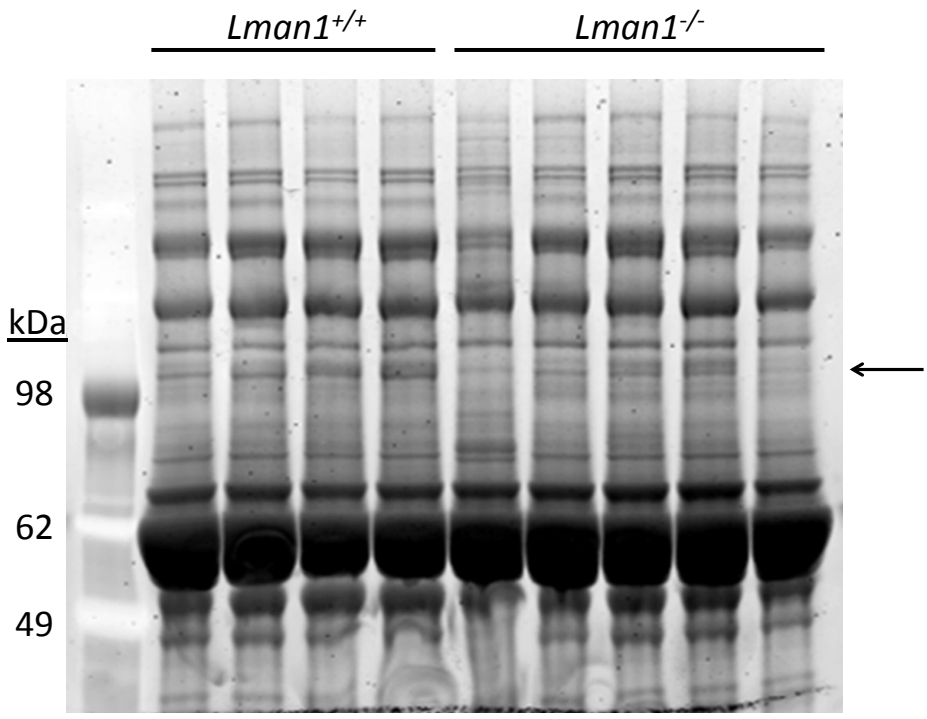


Figure 4-4: Differential abundance of a plasma protein in *Lman1*$^{+/+}$ and *Lman1*$^{-/-}$ mice

Platelet-poor plasma from wild-type mice and *Lman1*$^{-/-}$ mice was analyzed by SDS-PAGE to potentially identify plasma proteins with differential abundance. As indicated by the black arrow, a band of ~ 100 kDa was of lower abundance in the plasma samples from *Lman1*$^{-/-}$ mice as compared to *Lman1*$^{+/+}$ mice. Each lane represents a different animal. The band of interest was excised from three *Lman1*$^{+/+}$ lanes and three *Lman1*$^{-/-}$ lanes. These 6 gel fragments were submitted for mass-spectrometry in order to identify the protein, and to determine if its abundance is different between *Lman1*$^{+/+}$ mice and *Lman1*$^{-/-}$ mice.

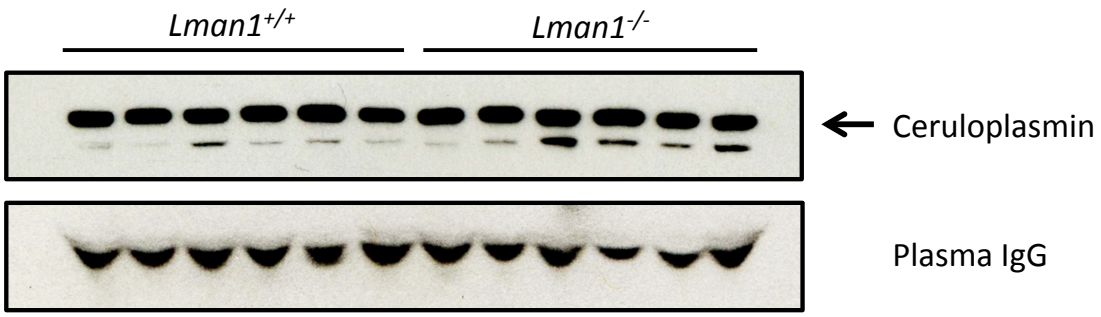


Figure 4-5: Plasma ceruloplasmin is not reduced in *Lman1*$^{-/-}$ mice

Platelet-poor plasma from wild-type mice and *Lman1*$^{-/-}$ mice was analyzed by Western blot to determine whether ceruloplasmin plasma levels are reduced in *Lman1*$^{-/-}$ mice. The primary ceruloplasmin band (indicated by the arrow) represents full-length native ceruloplasmin, while the slightly lower molecular weight band represents a degradation product. Although there is a slight trend for *Lman1*$^{-/-}$ mice to have more of the degradation product as compared to wild-type mice, overall the abundance of ceruloplasmin in *Lman1*$^{+/+}$ and *Lman1*$^{-/-}$ plasma samples is indistinguishable. Each lane represents a different animal.

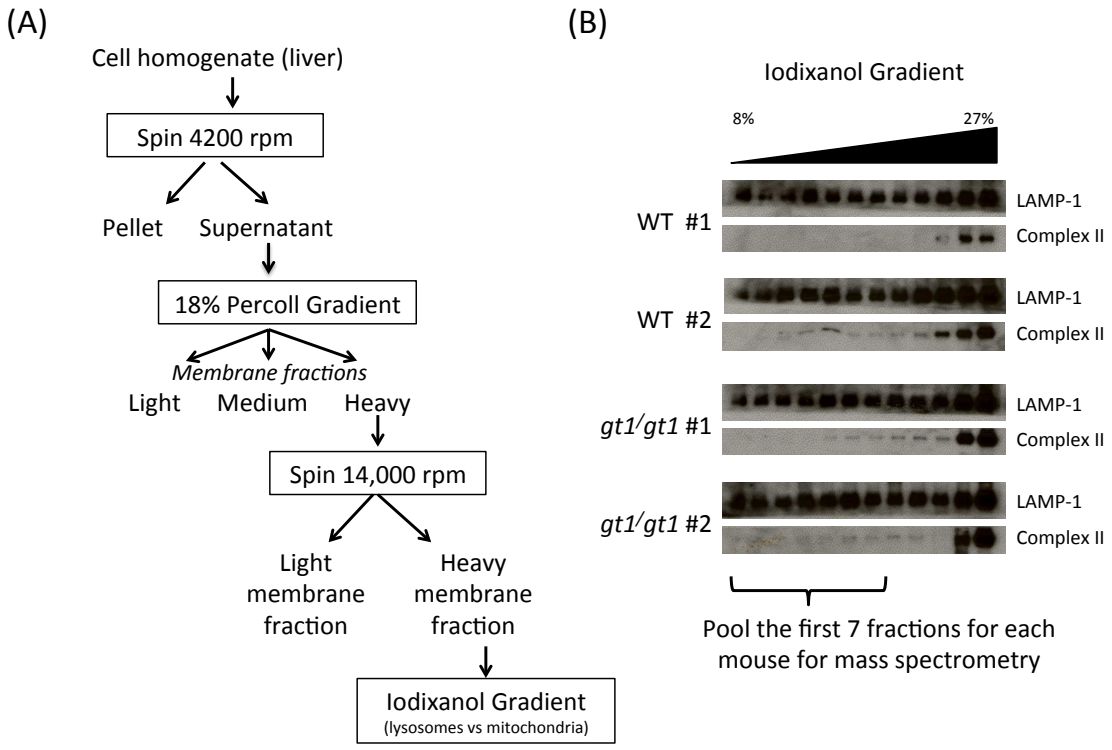


Figure 4-6: Isolation of liver lysosomes

(A) Schematic of the cell fractionation procedure to isolate lysosomes from mouse liver tissues. Liver tissue samples were pulverized and homogenized, and the lysate was centrifuged to obtain a post-nuclear supernatant. This supernatant was applied to an 18% Percoll solution for further centrifugation, and the resulting heavy membrane fraction was further centrifuged and fractionated on an iodixanol Gradient. **(B)** The fractions containing heavy membrane structures from **(A)** were applied to an iodixanol gradient (8-27%) in order to separate lysosomes and mitochondria. Western blot analysis of the resulting fractions was used to identify which fractions demonstrated high lysosome enrichment, but little to no mitochondrial contamination. A LAMP-1 antibody was used as a lysosome marker, and a Complex II antibody was used as a mitochondrial marker. The first 7 to 8 fractions from the iodixanol gradient (~8-15%) were reproducibly found to be highly enriched for lysosomes and largely depleted of mitochondria. The first 7 fractions of each prep were therefore pooled and submitted for mass-spectrometry analysis.

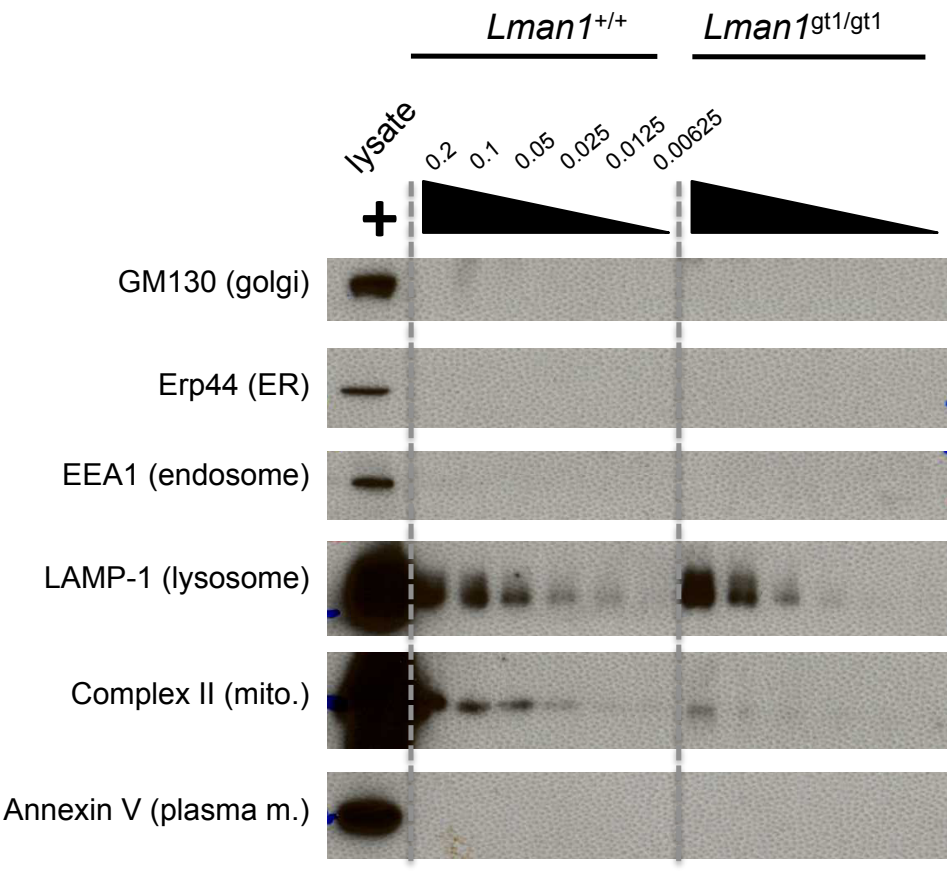


Figure 4-7: Lysosome preps are highly enriched and appear nearly pure

Lysosomes were isolated from fresh mouse liver tissue by differential density centrifugation on sucrose and iodixanol gradients. Serial dilutions of the resulting lysosome fraction (from 20% down to 0.625%) were analyzed by Western blot for a number of cell markers to confirm lysosome enrichment and sample purity. The positive control represents total liver lysate (input).

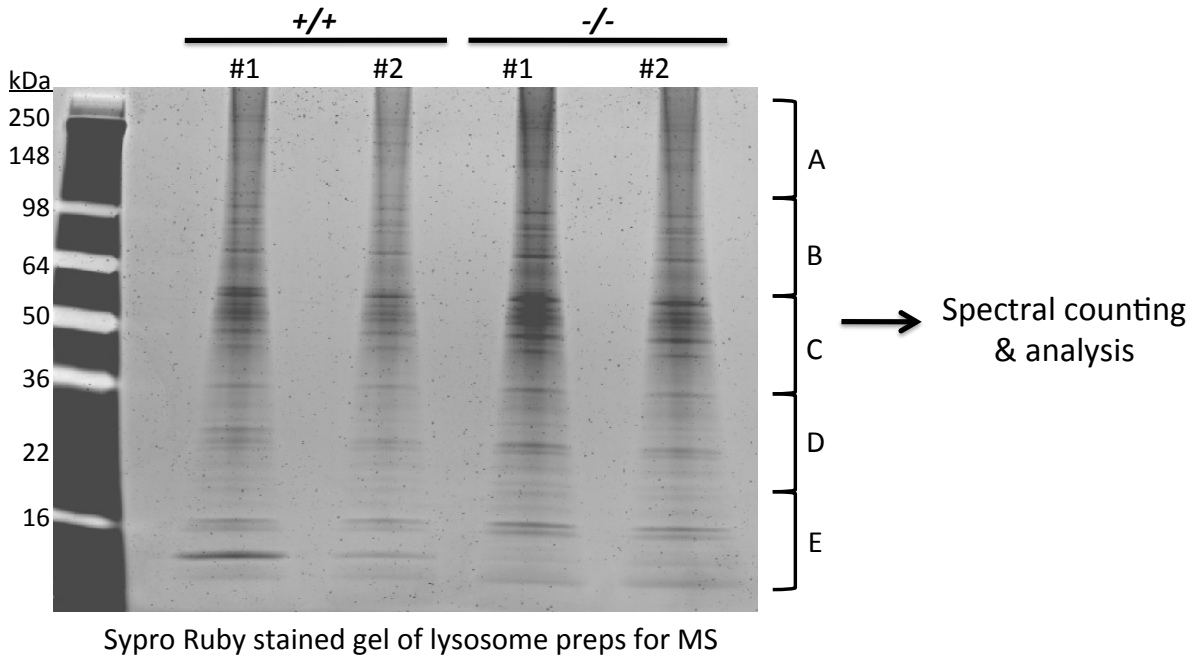


Figure 4-8: Gel fractionation of pooled lysosome preps for mass-spectrometry

The final product of the lysosome prep from mouse liver tissues (i.e. the pooled product of iodixanol fractions #1-7 in Figure 4-6B) was submitted to SDS-PAGE in order to assess for sample complexity. Although some subtle differences between the four lysosomal preps are detectable visually, there are no obvious differences in protein content or abundance between the wild-type and *Lman1*$^{-/-}$ samples. However, the complexity of these samples was quite high (i.e. there are many proteins in each prep), so each gel lane was cut into 5 different gel slices and submitted to mass-spectrometry individually.

CHAPTER 5 :
CONCLUSIONS AND FUTURE DIRECTIONS

CHARACTERIZATION OF AN ALLELIC SERIES OF MURINE *LMAN1* MUTATIONS

An LMAN1-deficient mouse model (*Lman1*$^{gt1/gt1}$) with <1% of normal *Lman1* gene expression was previously shown to reproduce the human F5F8D phenotype, but with important phenotypic differences as compared to human F5F8D patients: less severe FV and FVIII reduction and partial lethality.[95] To address these differences and to further characterize the role of the LMAN1 cargo receptor, we generated an independent, hypomorphic conditional gene-trap allele (*Lman1*^{cgt}) and a true null allele with no *Lman1* expression (*Lman1*$^{-}$). We demonstrated a correlation between *Lman1* gene expression and the levels of LMAN1-dependent cargo secretion for FV, FVIII, and A1AT. *Lman1*$^{cgt/cgt}$ mice retain ~7% of normal *Lman1* expression and have plasma FV and FVIII activity levels that are ~70% of wild-type levels. In contrast, null *Lman1*$^{-/-}$ mice exhibit plasma FV and FVIII levels that are ~50% of wild-type levels, a reduction of FV and FVIII plasma activity that is less severe than that observed in human F5F8D patients, with ~10-15% of normal levels. This finding demonstrates a threshold effect for LMAN1-mediated secretion, such that total loss of LMAN1 function in the mouse only results in ~50% reduction of FV and FVIII secretion. These data indicate that an unknown parallel or compensatory mechanism exists in mice for the secretion of the other ~50% of the plasma FV and FVIII. The correlation between *Lman1* expression and LMAN1-dependent cargo secretion also suggests that genetic variation at the *LMAN1* locus may account for variation in FV and FVIII levels in human populations, as well as for some patients with mild deficiencies in FV and FVIII not previously identified as F5F8D.

We also demonstrated that *Lman1*$^{-/-}$ mice exhibit a strain-specific, partially penetrant, perinatal lethality, confirming the previous report of lethality in *Lman1*$^{gt1/gt1}$ mice. Finally, *Lman1*$^{-/-}$ mice exhibit a moderate, isolated thrombocytopenia that is not observed in hypomorphic *Lman1*$^{cgt/cgt}$ mice, heterozygous *Lman1*$^{+/-}$ mice, or *Mcfd2*$^{-/-}$ mice, suggesting a role for a novel platelet-specific LMAN1 cargo, or an unknown function of LMAN1 in platelet development. Additional studies to characterize *Lman1*$^{-/-}$ mice for a defect in platelet secretory granules or the development of megakaryocytes and platelets are currently underway.

ENDOTHELIAL CELLS ARE THE PRIMARY SOURCE OF FVIII BIOSYNTHESIS *IN VIVO*

The primary cellular source of FVIII biosynthesis has been controversial, with contradictory evidence supporting an endothelial or hepatocyte origin. *Lman1* conditional knockout mice were studied to determine the relative contribution of hepatocytes and endothelial cells to the plasma FVIII pool. Mice deficient for LMAN1 in hepatocytes exhibit low plasma FV activity levels, but normal FVIII plasma activity levels. In contrast, mice deficient for LMAN1 in the endothelium exhibit normal plasma FV levels, but low FVIII levels, with FVIII reduction comparable to that observed in the ubiquitous *Lman1* knockout mice. These data indicate that endothelial cells are the primary biosynthetic source of murine FVIII, and that hepatocytes make no significant contribution to the plasma FVIII pool. Utilizing RiboTag mice and polyribosome immunoprecipitation, we performed endothelial cell-specific mRNA isolation and qPCR analyses to confirm that endothelial cells highly express *F8*, and to explore the heterogeneity of *F8* expression in different vascular beds. Endothelial cells from multiple, but not all, tissues contribute to the plasma FVIII pool in the mouse.

PROTEOMIC APPROACHES TO IDENTIFY LMAN1 FUNCTIONS AND INTERACTING PROTEINS

The identification of new LMAN1 cargos and interacting proteins that regulate LMAN1 function is important to further define LMAN1's role as an ER cargo-receptor. We demonstrated by co-immunoprecipitation that LMAN1 interacts with all four SEC24 paralogs (A-D), which are critical components of the inner COPII coat. We also demonstrated that both SEC23 paralogs (SEC23A and SEC23B) interact with all four SEC24 paralogs, implying that 8 possible SEC23-SEC24 heterodimer combinations are available as building blocks for COPII vesicles. The biologic significance and potential functional differences between these heterodimers are unknown. Different heterodimers could serve tissue- or cargo-specific functions to facilitate the selective recruitment of secretory cargos into budding COPII vesicles. We also performed a high-throughput yeast two-hybrid study to identify potentially novel protein-protein interactions for the LMAN1 luminal domain, the LMAN1 cytoplasmic domain, and the MCFD2 protein. Unfortunately this screen did not identify reproducible interactions (including interactions with known LMAN1 cargo proteins), perhaps due to the technical limitations of performing a Y2H assay for a membrane protein that is primarily located in the ER lumen (where protein-protein interactions form under very different conditions than would occur in the yeast nucleus).

Taking advantage of our mouse model of F5F8D, we attempted several cell fractionation procedures from liver lysates of F5F8D mice and wild-type controls in order to isolate ER, Golgi, and lysosome secretory compartments for mass-spectrometry analysis. Although we successfully enriched for lysosomes and appeared to eliminate nearly all other cellular contaminants from our preps, as demonstrated by Western blot analyses, the mass-spectrometry analyses of these samples revealed the presence of many non-lysosomal membrane and secretory proteins that prevented the useful comparison of lysosomal contents across different samples. Further optimization to achieve cleaner, simpler cell fractions for mass-spectrometry is necessary to resolve this issue, perhaps prepared from cultured cells rather than from tissues. Finally, we observed an ~100 kDa

protein of interest upon SDS-PAGE analysis of plasma samples from wild-type and *Lman1*$^{-/-}$ mice. The apparent protein was present in lower abundance in the plasma of *Lman1*$^{-/-}$ mice as compared to that of wild-type mice. Mass-spectrometry analysis of this ~100 kDa gel band from both wild-type and *Lman1*$^{-/-}$ mice demonstrated that two proteins were significantly under-represented in the plasma of *Lman1*$^{-/-}$ mice. The first protein was SERPINA3K (also known as "serine (or cysteine) peptidase inhibitor, clade A, member 3K," or as alpha-1 antichymotrypsin). Interestingly, SERPINA3K is a member of the same clade of SERPINS (serine proteinase inhibitors) as the highly homologous protein alpha-1 antitrypsin (A1AT, also known as SERPINA1), another previously-reported LMAN1 dependent cargo (see Chapter 4). The second under-represented protein in the *Lman1*$^{-/-}$ plasma samples was Complement Component C3, a serine protease. We are currently working to confirm these preliminary results and to further explore SERPINA3K and/or Complement Component C3 as putative LMAN1 cargos.

FUTURE DIRECTIONS

Our incomplete understanding of LMAN1's role in the secretory pathway is an important scientific problem. In the future, our efforts will focus on addressing several fundamental questions in LMAN1 biology and F5F8D pathophysiology. As described in the following sections, the molecular basis for the apparent differences in F5F8D presentation in mice and humans may be explained by identifying a compensatory or parallel mechanism for LMAN1 cargo secretion. It is also possible that the impaired secretion of a novel LMAN1 cargo that is unique to mice results in perinatal lethality, which would account for the lack of known perinatal death or elevated miscarriage rates associated with F5F8D in humans. In addition, further study will reveal the molecular etiology of the thrombocytopenia observed in *Lman1*$^{-/-}$ mice, which has not been observed in human patients. However, a relatively subtle thrombocytopenia may have been missed in humans for several reasons, including the small number of patients, the

wide range of platelet count values that are considered clinically normal, and the inter-laboratory variability in CBC analyses and blood analyzers. However, identification of a platelet defect in human F5F8D patients would be an important finding both for the further characterization of LMAN1-dependent protein secretion, and for understanding platelet biology.

Identify alternate mechanism(s) for LMAN1 cargo secretion

Complete LMAN1 or MCFD2 deficiency in both mice and humans does not cause FV and FVIII plasma levels to be reduced to zero, but only to ~50%, so an alternate secretion mechanism for these cargos clearly exists. Though a novel ER cargo receptor, similar in function to LMAN1, could be responsible for this alternate secretion pathway, no other mammalian ER cargo receptors are currently known, despite significant efforts to identify them. Physical cargo/receptor interactions are transient protein-protein interactions, unlike a stable complex that may be isolated intact, and may be difficult to detect. We do not currently know whether an alternate FV and FVIII secretion mechanism is utilized constitutively, or only in the absence of LMAN1 as an upregulated compensatory mechanism. Similarly, do other LMAN1 cargos also utilize the same alternate mechanism? Given that some controversy still remains regarding the two models of ER-to-Golgi transport (bulk-flow vs. receptor-mediated),[192] we must also consider the possibility that in the absence of LMAN1, at least some FV and FVIII is secreted via bulk-flow rather than by receptor-mediated ER-to-Golgi transport. However, the finding that 50% of FV and FVIII are secreted to the plasma in *Lman1*$^{-/-}$ mice makes bulk-flow a less likely explanation, because bulk-flow alone is unlikely to be able to support such a high rate of secretion.[71]

Lman1$^{-/-}$ and *Mcfd2*$^{-/-}$ mice both exhibit FV and FVIII levels that are ~50% of normal. Since LMAN1 and MCFD2 work together in the same receptor complex (i.e. same "pathway"), double knockout mice are expected to also exhibit ~50% of normal FV and FVIII levels. If the double knockout mice exhibit a more severe reduction in FV and FVIII plasma levels as compared to the single knockout mice, this will indicate that

LMAN1 and MCFD2 are not epistatic, and suggest an alternative interaction for one or both of these proteins.

As an initial approach to identify potential alternate secretion mechanisms for LMAN1 cargos, and to understand the cellular response to LMAN1 deficiency, RNAseq could be used to determine the gene expression changes that occur in the absence of LMAN1. If an alternate receptor complex or secretory pathway is upregulated in LMAN1-deficient cells, it is possible that differences in the RNAseq profiles of wild-type cells and *Lman1*$^{-/-}$ cells will reveal these compensatory cellular changes. If no differences are observed at the RNA level, this does not rule out a compensatory upregulation at the protein level (due to increased translation and/or decreased protein degradation). RNAseq and/or qPCR analysis will also reveal potential upregulation of unfolded protein response (UPR) genes and markers of ER stress resulting from impaired secretion of cargos, which might serve as a cellular trigger for an alternate secretion mechanism for LMAN1 cargos. It is possible that LMAN1 cargos are upregulated themselves in the absence of LMAN1, as a way to compensate for loss of their usual secretory receptor complex. This finding might suggest a feedback loop or sensing mechanism so that LMAN1 cargos are upregulated when their secretion is impaired. In this case, RNAseq would reveal upregulation of *F5, F8,* and *A1AT* in *Lman1*$^{-/-}$ cells as compared to wild-type cells.

Whole-genome functional screens represent another appealing approach to identify the alternate secretion mechanism for FV and FVIII. Whole-genome CRISPR (clustered regularly interspaced short palindromic repeats)-Cas9 knockout screens have recently been developed to efficiently and systematically knockout almost every gene in the genome for use in high-throughput studies.[193,194] This approach was used to target 18,080 genes in human cells to identify genes that are essential for cell viability in cancer and pluripotency, as well as to identify genes involved in resistance to vemurafenib.[193] This approach could be applied to LMAN1-deficient cells (hepatocytes for FV, endothelial cells for FVIII) to identify novel components of the alternate secretion pathway for FV and FVIII. In this system, LMAN1-deficient cells would still secrete ~50% of normal levels of FV or FVIII. However, in the presence of a CRISPR-induced

genetic mutation that inhibits the alternate FV/FVIII secretion pathway, the secretion of these coagulation factors would drop from ~50% down to as low as 0% secretion. This approach for a secreted protein such as FV or FVIII would require some technical adaptation in order to determine whether the protein of interest was secreted or not in a high-throughput manner. One possible approach would be to engineer a glycosylphosphatidylinositol (GPI) anchor onto the secreted cargo, which would cause it to be bound to the cell membrane. Flow-cytometry cell sorting could then be used to sort the cells based on the presence or absence of the GPI-anchored cargo on the cell surface. Cells without the GPI-anchored protein on their surface would represent cells in which a CRISPR-induced mutation had impaired the alternate secretion mechanism for the LMAN1 cargo. Next-generation sequencing of the clones that did not display the GPI–anchored cargo on the cell surface could reveal genes responsible for the alternate FV and FVIII secretion pathway.

Elucidate the etiology of *Lman1*$^{-/-}$ thrombocytopenia

The finding that *Lman1*$^{-/-}$ mice are moderately thrombocytopenic is a surprising result, suggesting that LMAN1 may serve as a cargo receptor for a platelet- or megakaryocyte-specific secreted protein that is critical for platelet development or survival in the plasma. Alternatively, LMAN1 could serve as an ER cargo receptor for a growth factor that is secreted from another (non-megakaryocyte) cell type. A number of studies could help elucidate the etiology of this platelet defect.

First, it is critical to determine if the defect is cell intrinsic (i.e. due to a defect in the megakaryocyte/platelet lineage) as opposed to a defect in another cell type (bone marrow hematopoietic cells, bone marrow stromal cells, or other cell types) that have a secondary, downstream effect of thrombocytopenia. We could utilize conditional *Lman1* knockout mice (Chapter 2) in which *Lman1* is specifically deleted in the megakaryocyte/platelet lineage (*Pf4*-Cre) or the hematopoietic compartment (*Vav1*-Cre). If the thrombocytopenia phenotype results from an intrinsic defect in megakaryocytes or platelets, then the phenotype should be reproduced in the *Lman1* platelet (*Pf4*) and hematopoietic (*Vav1*) knockout mice. Alternatively, if thrombocytopenia is observed in

the hematopoietic compartment knockout mice, but not the megakaryocyte/platelet-specific knockout mice, this finding would indicate that the defect arises in the bone marrow, but is not intrinsic to the megakaryocyte/platelet lineage. If hematopoietic (*Vav1*) and megakaryocyte/platelet (*Pf4*) conditional knockout mice have normal platelet counts, this would suggest that the thrombocytopenia observed in *Lman1*$^{-/-}$ mice is an indirect effect of a defect in another cell type. An intriguing possible explanation would be a defect in the secretion of thrombopoietin (also known as megakaryocyte growth and development factor) from the liver as the cause of *Lman1*$^{-/-}$ thrombocytopenia. Thrombopoietin is a glycoprotein synthesized in the liver that simulates the production and differentiation of megakaryocytes, the cells that generate platelets. To explore this possibility, we would measure platelet counts in conditional knockout mice with *Lman1* specifically deleted in the hepatocytes. A low platelet count in hepatocyte-specific *Lman1* knockout mice would suggest that an LMAN1 cargo secreted from the liver is responsible for the thrombocytopenia of *Lman1*$^{-/-}$ mice. These tissue-specific knockout studies are currently underway.

If the thrombocytopenia defect is intrinsic to the megakaryocyte/platelet lineage, it would be important to determine if the thrombocytopenia is associated with a secretory defect in megakaryocytes/platelets. Platelets are highly secretory cells that release the contents of alpha-granules and dense-granules into the plasma upon activation. The contents of alpha-granules and dense-granules are all secretory cargos that must be properly synthesized and transported from the ER to the Golgi and the secretory granules. A defect in the ER-to-Golgi transport of these platelet secretory cargos would be expected to be identified by comparing the contents of platelet alpha granules[195] and dense granules[196] from wild-type control mice and *Lman1*$^{-/-}$ mice by mass-spectrometry.

Megakaryocyte differentiation studies[197] in cell culture for wild-type and *Lman1*$^{-/-}$ cells could be performed to assess a potential defect in cell lineage, such as in pro-platelet formation. Similarly, FACS analysis of bone marrow cells from wild-type and *Lman1*$^{-/-}$ mice may reveal a potential block in platelet differentiation in *Lman1*$^{-/-}$ mice, although this approach may not reveal a defect since the thrombocytopenia phenotype in *Lman1*$^{-/-}$

mice is mild. A block in platelet differentiation would indicate that LMAN1 plays a role in platelet formation. These experiments may help distinguish between impaired platelet differentiation and premature platelet destruction.

Any putative LMAN1-dependent cargos in platelets or megakaryocytes could be validated in our murine F5F8D mice, as well as in human patients and control subjects. The finding of a platelet defect in F5F8D patients would be an important clinical and translational result. In order to validate the putative cargo, immunofluorescence studies could be used to demonstrate accumulation in the ER in the absence of LMAN1. *In vitro* budding assays could be performed to determine whether incorporation of the putative cargo into COPII vesicles is dependent upon LMAN1.

Identification of novel LMAN1 cargos and interacting proteins

Since purification of transmembrane proteins with their associated binding partners is technically difficult and not well suited to affinity-purification mass-spectrometry, another sensitive method of capturing and identifying LMAN1's protein-protein interactions is desirable. Biotinylation of LMAN1 proximal and interacting proteins (cargos, adaptors, regulators) *in vivo* can be achieved by utilizing a promiscuous biotin ligase fusion protein (LMAN1-BirA*), a technique based on the fusion of a modified biotin ligase (BirA*) to a protein of interest.[198] The BirA* attaches a biotin to the LMAN1-fusion protein and its interacting proteins (proximity-dependent biotinylation). The strong interaction between biotin and streptavidin can then be utilized to affinity-purify labeled proteins, followed by mass-spectrometry for identification. *In vivo* biotinylation is particularly appropriate for the identification of novel LMAN1 cargos and interacting proteins because it is suitable for detecting interacting and neighboring proteins in the native cellular environment, and the biotinylation is efficient enough to label proteins that only transiently interact with LMAN1 (such as cargos or adaptors). In fact, *in vivo* biotinylation has already been successfully applied for the biotinylation of proteins within the secretory pathway.[199] The *in vivo* biotinylation technique could be applied not only to LMAN1 (to identify new cargos, LMAN1-adaptors, and regulators of its function), but also to MCFD2 in order to potentially identify other ER cargo receptors

that it might pair with, and/or other secretory cargos that are MCFD2-dependent for secretion. Similarly, FV-BirA* and FVIII-BirA* fusion proteins might be used to biotinylate other cargo receptors or components that are responsible for the alternate, LMAN1-independent secretion of FV and FVIII. One potential limitation of the *in vivo* biotinylation approach is that the BirA* fusion may cause the targeted protein to mis-fold, and/or may impair its endogenous protein-protein interactions. The effect of the BirA* fusion on the function, cellular localization, and binding pattern of LMAN1 and any other targeted proteins would therefore have to be evaluated on a case-by-case basis to ensure the validity of this experimental approach.

Finally, since LMAN1 has a luminal carbohydrate recognition domain and is a mannose-specific lectin, glycopeptide enrichment of plasma or cell samples represents another useful mass-spectrometry based approach to enrich a biological sample for presumed LMAN1 cargos. Glycoproteins are often low in abundance in a biologic sample or complex mixture, and glycopeptide enrichment compensates for the otherwise low number of spectral counts assigned to glycoproteins in mass-spectrometry studies. Glycopeptide enrichment, in combination with quantitative proteomic methods such as SILAC (discussed in Chapter 4), are expected to generate high-quality, targeted proteomic profiling data for the characterization of individual secretory compartments in mammalian cells.

Although further work and technical optimization of proteomic-based assays is required to identify novel LMAN1 secretory cargos, these approaches are expected to reveal LMAN1's full complement of cargos and to further elucidate the fundamental cellular process of receptor-mediated cargo secretion.

REFERENCES

1. Rosendaal FR, Reitsma PH. Genetics of venous thrombosis. *J Thromb Haemost.* Jul 2009;7 Suppl 1:301-304.

2. Kyrle PA, Minar E, Hirschl M, et al. High plasma levels of factor VIII and the risk of recurrent venous thromboembolism. *N Engl J Med.* Aug 17 2000;343(7):457-462.

3. Zhang B, Cunningham MA, Nichols WC, et al. Bleeding due to disruption of a cargo-specific ER-to-Golgi transport complex. *Nat Genet.* Jun 2003;34(2):220-225.

4. Nichols WC, Seligsohn U, Zivelin A, et al. Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII. *Cell.* Apr 3 1998;93(1):61-70.

5. Oeri J, Matter M, Isenschmid H, Hauser F, Koller F. [Congenital factor V deficiency (parahemophilia) with true hemophilia in two brothers]. *Bibliotheca paediatrica.* 1954;58:575-588.

6. Spreafico M, Peyvandi F. Combined FV and FVIII deficiency. *Haemophilia.* Nov 2008;14(6):1201-1208.

7. Mannucci PM, Duga S, Peyvandi F. Recessively inherited coagulation disorders. *Blood.* Sep 1 2004;104(5):1243-1252.

8. Segal A, Zivelin A, Rosenberg N, Ginsburg D, Shpilberg O, Seligsohn U. A mutation in LMAN1 (ERGIC-53) causing combined factor V and factor VIII deficiency is prevalent in Jews originating from the island of Djerba in Tunisia. *Blood Coagul Fibrinolysis.* Jan 2004;15(1):99-102.

9. Seligsohn U, Zivelin A, Zwang E. Combined factor V and factor VIII deficiency among non-Ashkenazi Jews. *N Engl J Med.* Nov 4 1982;307(19):1191-1195.

10. Zhang B, Ginsburg D. Familial multiple coagulation factor deficiencies: new biologic insight from rare genetic bleeding disorders. *J Thromb Haemost.* Sep 2004;2(9):1564-1572.

11. *World Federation of Hemophilia Report on the Annual Global Survey 2011.* 2012.

12. Exome Variant Server (NHLBI GO Exome Sequencing Project S, WA.). 2014.

13. Mansouritorgabeh H, Rezaieyazdi Z, Pourfathollah AA, Rezai J, Esamaili H. Haemorrhagic symptoms in patients with combined factors V and VIII deficiency in north-eastern Iran. *Haemophilia.* May 2004;10(3):271-275.

14. Peyvandi F, Tuddenham EG, Akhtari AM, Lak M, Mannucci PM. Bleeding symptoms in 27 Iranian patients with the combined deficiency of factor V and factor VIII. *Br J Haematol.* Mar 1998;100(4):773-776.

15. Kappeler F, Klopfenstein DR, Foguet M, Paccaud JP, Hauri HP. The recycling of ERGIC-53 in the early secretory pathway. ERGIC-53 carries a cytosolic

endoplasmic reticulum-exit determinant interacting with COPII. *J Biol Chem.* Dec 12 1997;272(50):31801-31808.

16. Nufer O, Guldbrandsen S, Degen M, et al. Role of cytoplasmic C-terminal amino acids of membrane proteins in ER export. *J Cell Sci.* Feb 1 2002;115(Pt 3):619-628.

17. Klumperman J, Schweizer A, Clausen H, et al. The recycling pathway of protein ERGIC-53 and dynamics of the ER-Golgi intermediate compartment. *J Cell Sci.* Nov 1998;111 (Pt 22):3411-3425.

18. Schweizer A, Fransen JA, Bachi T, Ginsel L, Hauri HP. Identification, by a monoclonal antibody, of a 53-kD protein associated with a tubulo-vesicular compartment at the cis-side of the Golgi apparatus. *J Cell Biol.* Nov 1988;107(5):1643-1653.

19. Hauri H, Appenzeller C, Kuhn F, Nufer O. Lectins and traffic in the secretory pathway. *FEBS Lett.* Jun 30 2000;476(1-2):32-37.

20. Fiedler K, Simons K. A putative novel class of animal lectins in the secretory pathway homologous to leguminous lectins. *Cell.* Jun 3 1994;77(5):625-626.

21. Itin C, Roche AC, Monsigny M, Hauri HP. ERGIC-53 is a functional mannose-selective and calcium-dependent human homologue of leguminous lectins. *Mol Biol Cell.* Mar 1996;7(3):483-493.

22. Fiedler K, Simons K. Characterization of VIP36, an animal lectin homologous to leguminous lectins. *J Cell Sci.* Jan 1996;109 (Pt 1):271-276.

23. Neve EP, Svensson K, Fuxe J, Pettersson RF. VIPL, a VIP36-like membrane protein with a putative function in the export of glycoproteins from the endoplasmic reticulum. *Experimental cell research.* Aug 1 2003;288(1):70-83.

24. Nufer O, Mitrovic S, Hauri HP. Profile-based data base scanning for animal L-type lectins and characterization of VIPL, a novel VIP36-like endoplasmic reticulum protein. *J Biol Chem.* May 2 2003;278(18):15886-15896.

25. Yerushalmi N, Keppler-Hafkemeyer A, Vasmatzis G, et al. ERGL, a novel gene related to ERGIC-53 that is highly expressed in normal and neoplastic prostate and several other tissues. *Gene.* Mar 7 2001;265(1-2):55-60.

26. Satoh T, Suzuki K, Yamaguchi T, Kato K. Structural Basis for Disparate Sugar-Binding Specificities in the Homologous Cargo Receptors ERGIC-53 and VIP36. *PLoS One.* 2014;9(2):e87963.

27. Reiterer V, Nyfeler B, Hauri HP. Role of the lectin VIP36 in post-ER quality control of human alpha1-antitrypsin. *Traffic.* Aug 2010;11(8):1044-1055.

28. Kamiya Y, Kamiya D, Yamamoto K, Nyfeler B, Hauri HP, Kato K. Molecular basis of sugar recognition by the human L-type lectins ERGIC-53, VIPL, and VIP36. *J Biol Chem.* Jan 25 2008;283(4):1857-1861.

29. Nawa D, Shimada O, Kawasaki N, Matsumoto N, Yamamoto K. Stable interaction of the cargo receptor VIP36 with molecular chaperone BiP. *Glycobiology.* Sep 2007;17(9):913-921.

30. Arshad N, Ballal S, Visweswariah SS. Site-specific N-linked glycosylation of receptor guanylyl cyclase C regulates ligand binding, ligand-mediated activation and interaction with vesicular integral membrane protein 36, VIP36. *J Biol Chem.* Feb 8 2013;288(6):3907-3917.

31. Zhang B, McGee B, Yamaoka JS, et al. Combined deficiency of factor V and factor VIII is due to mutations in either LMAN1 or MCFD2. *Blood.* Mar 1 2006;107(5):1903-1907.
32. Zheng C, Liu HH, Zhou J, Zhang B. EF-hand domains of MCFD2 mediate interactions with both LMAN1 and coagulation factor V or VIII. *Blood.* Feb 4 2010;115(5):1081-1087.
33. Wigren E, Bourhis JM, Kursula I, Guy JE, Lindqvist Y. Crystal structure of the LMAN1-CRD/MCFD2 transport receptor complex provides insight into combined deficiency of factor V and factor VIII. *FEBS Lett.* Mar 5 2010;584(5):878-882.
34. Kawasaki N, Ichikawa Y, Matsuo I, et al. The sugar-binding ability of ERGIC-53 is enhanced by its interaction with MCFD2. *Blood.* Feb 15 2008;111(4):1972-1979.
35. Nyfeler B, Zhang B, Ginsburg D, Kaufman RJ, Hauri HP. Cargo selectivity of the ERGIC-53/MCFD2 transport receptor complex. *Traffic.* Nov 2006;7(11):1473-1481.
36. Zhang B. Recent developments in the understanding of the combined deficiency of FV and FVIII. *Br J Haematol.* Apr 2009;145(1):15-23.
37. Nufer O, Kappeler F, Guldbrandsen S, Hauri HP. ER export of ERGIC-53 is controlled by cooperation of targeting determinants in all three of its domains. *J Cell Sci.* Nov 1 2003;116(Pt 21):4429-4440.
38. Peyvandi F, Kunicki T, Lillicrap D. Genetic sequence analysis of inherited bleeding diseases. *Blood.* Nov 14 2013;122(20):3423-3431.
39. Zhang B, Spreafico M, Zheng C, et al. Genotype-phenotype correlation in combined deficiency of factor V and factor VIII. *Blood.* Jun 15 2008;111(12):5592-5600.
40. 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.* Jul 8 2005;280(27):25881-25886.
41. Karimi M, Cairo A, Safarpour MM, et al. Genotype and phenotype report on patients with combined deficiency of factor V and factor VIII in Iran. *Blood Coagul Fibrinolysis.* Jan 2 2014.
42. Cunningham MA, Pipe SW, Zhang B, Hauri HP, Ginsburg D, Kaufman RJ. LMAN1 is a molecular chaperone for the secretion of coagulation factor VIII. *J Thromb Haemost.* Nov 2003;1(11):2360-2367.
43. Moussalli M, Pipe SW, Hauri HP, Nichols WC, Ginsburg D, Kaufman RJ. Mannose-dependent endoplasmic reticulum (ER)-Golgi intermediate compartment-53-mediated ER to Golgi trafficking of coagulation factors V and VIII. *J Biol Chem.* Nov 12 1999;274(46):32539-32542.
44. Miao HZ, Sirachainan N, Palmer L, et al. Bioengineering of coagulation factor VIII for improved secretion. *Blood.* May 1 2004;103(9):3412-3419.
45. Kaufman RJ. Post-translational modifications required for coagulation factor secretion and function. *Thromb Haemost.* Jun 1998;79(6):1068-1079.
46. Chiu HC, Schick PK, Colman RW. Biosynthesis of factor V in isolated guinea pig megakaryocytes. *J Clin Invest.* Feb 1985;75(2):339-346.

47. Camire RM, Pollak ES, Kaushansky K, Tracy PB. Secretable human platelet-derived factor V originates from the plasma pool. *Blood.* Nov 1 1998;92(9):3035-3041.
48. Gould WR, Simioni P, Silveira JR, Tormene D, Kalafatis M, Tracy PB. Megakaryocytes endocytose and subsequently modify human factor V in vivo to form the entire pool of a unique platelet-derived cofactor. *J Thromb Haemost.* Mar 2005;3(3):450-456.
49. Sun H, Yang TL, Yang A, Wang X, Ginsburg D. The murine platelet and plasma factor V pools are biosynthetically distinct and sufficient for minimal hemostasis. *Blood.* Oct 15 2003;102(8):2856-2861.
50. Kaufman RJ, Wasley LC, Dorner AJ. Synthesis, processing, and secretion of recombinant human factor VIII expressed in mammalian cells. *J Biol Chem.* May 5 1988;263(13):6352-6362.
51. Kaufman RJ, Wasley LC, Davies MV, Wise RJ, Israel DI, Dorner AJ. Effect of von Willebrand factor coexpression on the synthesis and secretion of factor VIII in Chinese hamster ovary cells. *Mol Cell Biol.* Mar 1989;9(3):1233-1242.
52. Bontempo FA, Lewis JH, Gorenc TJ, et al. Liver transplantation in hemophilia A. *Blood.* Jun 1987;69(6):1721-1724.
53. Marchioro TL, Hougie C, Ragde H, Epstein RB, Thomas ED. Hemophilia: role of organ homografts. *Science.* Jan 10 1969;163(3863):188-190.
54. Scharrer I, Encke A, Hottenrott C. Clinical cure of haemophilia A by liver transplantation. *Lancet.* Oct 1 1988;2(8614):800-801.
55. Gordon FH, Mistry PK, Sabin CA, Lee CA. Outcome of orthotopic liver transplantation in patients with haemophilia. *Gut.* May 1998;42(5):744-749.
56. Webster WP, Zukoski CF, Hutchin P, Reddick RL, Mandel SR, Penick GD. Plasma factor VIII synthesis and control as revealed by canine organ transplantation. *Am J Physiol.* May 1971;220(5):1147-1154.
57. Madeira CL, Layman ME, de Vera RE, Fontes PA, Ragni MV. Extrahepatic factor VIII production in transplant recipient of hemophilia donor liver. *Blood.* May 21 2009;113(21):5364-5365.
58. Norman JC, Covelli VH, Sise HS. Transplantation of the spleen: experimental cure of hemophilia. *Surgery.* Jul 1968;64(1):1-14.
59. Webster WP. Allotransplantation of spleen in hemophilia. *North Carolina medical journal.* 1967;28:505.
60. Veldkamp JJ AE, van de Torren K, van der Does JA, van Tilburg NH, Pauwels EK. Extrahepatic factor VIII synthesis: lung transplants in hemophilic dogs. *Transplantation.* 1974;18:56-62.
61. Wion KL, Kelly D, Summerfield JA, Tuddenham EG, Lawn RM. Distribution of factor VIII mRNA and antigen in human liver and other tissues. *Nature.* Oct 24-30 1985;317(6039):726-729.
62. Zelechowska MG, van Mourik JA, Brodniewicz-Proba T. Ultrastructural localization of factor VIII procoagulant antigen in human liver hepatocytes. *Nature.* Oct 24-30 1985;317(6039):729-730.
63. Do H, Healey JF, Waller EK, Lollar P. Expression of factor VIII by murine liver sinusoidal endothelial cells. *J Biol Chem.* Jul 9 1999;274(28):19587-19592.

64. Groth CG, Hathaway WE, Gustafsson A, et al. Correction of coagulation in the hemophilic dog by transplantation of lymphatic tissue. *Surgery.* May 1974;75(5):725-733.
65. Lee MC, Miller EA, Goldberg J, Orci L, Schekman R. Bi-directional protein transport between the ER and Golgi. *Annual review of cell and developmental biology.* 2004;20:87-123.
66. Barlowe C. Signals for COPII-dependent export from the ER: what's the ticket out? *Trends in cell biology.* Jun 2003;13(6):295-300.
67. Baines AC, Zhang B. Receptor-mediated protein transport in the early secretory pathway. *Trends Biochem Sci.* Aug 2007;32(8):381-388.
68. Wieland FT, Gleason ML, Serafini TA, Rothman JE. The rate of bulk flow from the endoplasmic reticulum to the cell surface. *Cell.* Jul 17 1987;50(2):289-300.
69. Martinez-Menarguez JA, Geuze HJ, Slot JW, Klumperman J. Vesicular tubular clusters between the ER and Golgi mediate concentration of soluble secretory proteins by exclusion from COPI-coated vesicles. *Cell.* Jul 9 1999;98(1):81-90.
70. Bednarek SY, Ravazzola M, Hosobuchi M, et al. COPI- and COPII-coated vesicles bud directly from the endoplasmic reticulum in yeast. *Cell.* Dec 29 1995;83(7):1183-1196.
71. Barlowe C, Orci L, Yeung T, et al. COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell.* Jun 17 1994;77(6):895-907.
72. Herrmann JM, Malkus P, Schekman R. Out of the ER--outfitters, escorts and guides. *Trends in cell biology.* Jan 1999;9(1):5-7.
73. Malkus P, Jiang F, Schekman R. Concentrative sorting of secretory cargo proteins into COPII-coated vesicles. *J Cell Biol.* Dec 23 2002;159(6):915-921.
74. Mizuno M, Singer SJ. A soluble secretory protein is first concentrated in the endoplasmic reticulum before transfer to the Golgi apparatus. *Proc Natl Acad Sci U S A.* Jun 15 1993;90(12):5732-5736.
75. Balch WE, McCaffery JM, Plutner H, Farquhar MG. Vesicular stomatitis virus glycoprotein is sorted and concentrated during export from the endoplasmic reticulum. *Cell.* Mar 11 1994;76(5):841-852.
76. Rexach MF, Latterich M, Schekman RW. Characteristics of endoplasmic reticulum-derived transport vesicles. *J Cell Biol.* Sep 1994;126(5):1133-1148.
77. Kuehn MJ, Schekman R. COPII and secretory cargo capture into transport vesicles. *Current opinion in cell biology.* Aug 1997;9(4):477-483.
78. Oprins A, Rabouille C, Posthuma G, Klumperman J, Geuze HJ, Slot JW. The ER to Golgi interface is the major concentration site of secretory proteins in the exocrine pancreatic cell. *Traffic.* Nov 2001;2(11):831-838.
79. Muniz M, Nuoffer C, Hauri HP, Riezman H. The Emp24 complex recruits a specific cargo molecule into endoplasmic reticulum-derived vesicles. *J Cell Biol.* Mar 6 2000;148(5):925-930.
80. Belden WJ, Barlowe C. Role of Erv29p in collecting soluble secretory proteins into ER-derived transport vesicles. *Science.* Nov 16 2001;294(5546):1528-1531.
81. Fiedler K, Veit M, Stamnes MA, Rothman JE. Bimodal interaction of coatomer with the p24 family of putative cargo receptors. *Science.* Sep 6 1996;273(5280):1396-1399.

82. Schimmoller F, Singer-Kruger B, Schroder S, Kruger U, Barlowe C, Riezman H. The absence of Emp24p, a component of ER-derived COPII-coated vesicles, causes a defect in transport of selected proteins to the Golgi. *EMBO J.* Apr 3 1995;14(7):1329-1339.

83. Caldwell SR, Hill KJ, Cooper AA. Degradation of endoplasmic reticulum (ER) quality control substrates requires transport between the ER and Golgi. *J Biol Chem.* Jun 29 2001;276(26):23296-23303.

84. Otte S, Barlowe C. Sorting signals can direct receptor-mediated export of soluble proteins into COPII vesicles. *Nat Cell Biol.* Dec 2004;6(12):1189-1194.

85. Nyfeler B, Reiterer V, Wendeler MW, et al. Identification of ERGIC-53 as an intracellular transport receptor of alpha1-antitrypsin. *J Cell Biol.* Feb 25 2008;180(4):705-712.

86. Carrell RW, Lomas DA. Alpha1-antitrypsin deficiency--a model for conformational diseases. *N Engl J Med.* Jan 3 2002;346(1):45-53.

87. Vollenweider F, Kappeler F, Itin C, Hauri HP. Mistargeting of the lectin ERGIC-53 to the endoplasmic reticulum of HeLa cells impairs the secretion of a lysosomal enzyme. *J Cell Biol.* Jul 27 1998;142(2):377-389.

88. Appenzeller C, Andersson H, Kappeler F, Hauri HP. The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. *Nat Cell Biol.* Oct 1999;1(6):330-334.

89. Chen Y, Hojo S, Matsumoto N, Yamamoto K. Regulation of Mac-2BP secretion is mediated by its N-glycan binding to ERGIC-53. *Glycobiology.* Jul 2013;23(7):904-916.

90. Roeckel N, Woerner SM, Kloor M, et al. High frequency of LMAN1 abnormalities in colorectal tumors with microsatellite instability. *Cancer research.* Jan 1 2009;69(1):292-299.

91. Cortini M, Sitia R. ERp44 and ERGIC-53 synergize in coupling efficiency and fidelity of IgM polymerization and secretion. *Traffic.* May 2010;11(5):651-659.

92. Mattioli L, Anelli T, Fagioli C, Tacchetti C, Sitia R, Valetti C. ER storage diseases: a role for ERGIC-53 in controlling the formation and shape of Russell bodies. *J Cell Sci.* Jun 15 2006;119(Pt 12):2532-2541.

93. Morais VA, Brito C, Pijak DS, 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.* Sep 2006;1762(9):802-810.

94. Klaus JP, Eisenhauer P, Russo J, et al. The intracellular cargo receptor ERGIC-53 is required for the production of infectious arenavirus, coronavirus, and filovirus particles. *Cell host & microbe.* Nov 13 2013;14(5):522-534.

95. Zhang B, Zheng C, Zhu M, et al. Mice deficient in LMAN1 exhibit FV and FVIII deficiencies and liver accumulation of {alpha}1-antitrypsin. *Blood.* Sep 22 2011;118(12):3384-3391.

96. Martin-Bermudo MD, Brown NH. The localized assembly of extracellular matrix integrin ligands requires cell-cell contact. *J Cell Sci.* Nov 2000;113 Pt 21:3715-3723.

97. Prout M, Damania Z, Soong J, Fristrom D, Fristrom JW. Autosomal mutations affecting adhesion between wing surfaces in Drosophila melanogaster. *Genetics.* May 1997;146(1):275-285.

98. Haberichter SL, Montgomery, R.R. . The biology of von Willebrand factor and factor VIII-regulated release. *Haematologica Reports.* 2005;1(6):9-14.
99. Lamont PA, Ragni MV. Lack of desmopressin (DDAVP) response in men with hemophilia A following liver transplantation. *J Thromb Haemost.* Oct 2005;3(10):2259-2263.
100. Jacquemin M, Neyrinck A, Hermanns MI, et al. FVIII production by human lung microvascular endothelial cells. *Blood.* Jul 15 2006;108(2):515-517.
101. Kelly DA, Summerfield JA, Tuddenham EG. Localization of factor VIIIC: antigen in guinea-pig tissues and isolated liver cell fractions. *Br J Haematol.* Apr 1984;56(4):535-543.
102. Exner T, Rickard KA, Kronenberg H. Measurement of factor VIII CAg by immunoradiometric assay in human tissue extracts. *Thromb Res.* Nov 15 1983;32(4):427-436.
103. Stel HV, van der Kwast TH, Veerman EC. Detection of factor VIII/coagulant antigen in human liver tissue. *Nature.* Jun 9-15 1983;303(5917):530-532.
104. Hollestelle MJ, Thinnes T, Crain K, et al. Tissue distribution of factor VIII gene expression in vivo--a closer look. *Thromb Haemost.* Sep 2001;86(3):855-861.
105. Yadav N, Kanjirakkuzhiyil S, Ramakrishnan M, Das TK, Mukhopadhyay A. Factor VIII can be synthesized in hemophilia A mice liver by bone marrow progenitor cell-derived hepatocytes and sinusoidal endothelial cells. *Stem cells and development.* Jan 2012;21(1):110-120.
106. Bi L, Lawler AM, Antonarakis SE, High KA, Gearhart JD, Kazazian HH, Jr. Targeted disruption of the mouse factor VIII gene produces a model of haemophilia A. *Nat Genet.* May 1995;10(1):119-121.
107. de Lange WJ, Halabi CM, Beyer AM, Sigmund CD. Germ line activation of the Tie2 and SMMHC promoters causes noncell-specific deletion of floxed alleles. *Physiological genomics.* Sep 17 2008;35(1):1-4.
108. Yang TL, Cui J, Rehumtulla A, et al. The structure and function of murine factor V and its inactivation by protein C. *Blood.* Jun 15 1998;91(12):4593-4599.
109. Adeghe AJ, Cohen J. A better method for terminal bleeding of mice. *Laboratory animals.* Jan 1986;20(1):70-72.
110. Tchaikovski SN, van Vlijmen BJ, Rosing J, Tans G. Development of a calibrated automated thrombography based thrombin generation test in mouse plasma. *J Thromb Haemost.* Oct 2007;5(10):2079-2086.
111. Sanz E, Yang L, Su T, Morris DR, McKnight GS, Amieux PS. Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. *Proc Natl Acad Sci U S A.* Aug 18 2009;106(33):13939-13944.
112. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc.* 2008;3(6):1101-1108.
113. Kouadjo KE, Nishida Y, Cadrin-Girard JF, Yoshioka M, St-Amand J. Housekeeping and tissue-specific genes in mouse tissues. *BMC genomics.* 2007;8:127.
114. Yang TL, Pipe SW, Yang A, Ginsburg D. Biosynthetic origin and functional significance of murine platelet factor V. *Blood.* Oct 15 2003;102(8):2851-2855.

115. Tang Y, Harrington A, Yang X, Friesel RE, Liaw L. The contribution of the Tie2+ lineage to primitive and definitive hematopoietic cells. *Genesis.* Sep 2010;48(9):563-567.
116. Tomer A. Human marrow megakaryocyte differentiation: multiparameter correlative analysis identifies von Willebrand factor as a sensitive and distinctive marker for early (2N and 4N) megakaryocytes. *Blood.* Nov 1 2004;104(9):2722-2727.
117. Murphy GJ, Leavitt AD. A model for studying megakaryocyte development and biology. *Proc Natl Acad Sci U S A.* Mar 16 1999;96(6):3065-3070.
118. Koni PA, Joshi SK, Temann UA, Olson D, Burkly L, Flavell RA. Conditional vascular cell adhesion molecule 1 deletion in mice: impaired lymphocyte migration to bone marrow. *The Journal of experimental medicine.* Mar 19 2001;193(6):741-754.
119. Chiasson VL, Talreja D, Young KJ, Chatterjee P, Banes-Berceli AK, Mitchell BM. FK506 binding protein 12 deficiency in endothelial and hematopoietic cells decreases regulatory T cells and causes hypertension. *Hypertension.* Jun 2011;57(6):1167-1175.
120. Ostronoff M, Ostronoff F, Campos G, et al. Allogeneic bone marrow transplantation in a child with severe aplastic anemia and hemophilia A. *Bone marrow transplantation.* Mar 2006;37(6):627-628.
121. Hollestelle MJ, Geertzen HG, Straatsburg IH, van Gulik TM, van Mourik JA. Factor VIII expression in liver disease. *Thromb Haemost.* Feb 2004;91(2):267-275.
122. Rapaport SI, Ames SB, Mikkelsen S, Goodman JR. Plasma clotting factors in chronic hepatocellular disease. *N Engl J Med.* Aug 11 1960;263:278-282.
123. Langley PG, Hughes RD, Williams R. Increased factor VIII complex in fulminant hepatic failure. *Thromb Haemost.* Oct 30 1985;54(3):693-696.
124. Nolan DJ, Ginsberg M, Israely E, et al. Molecular signatures of tissue-specific microvascular endothelial cell heterogeneity in organ maintenance and regeneration. *Developmental cell.* Jul 29 2013;26(2):204-219.
125. Aird WC. Phenotypic heterogeneity of the endothelium: I. Structure, function, and mechanisms. *Circulation research.* Feb 2 2007;100(2):158-173.
126. Aitsebaomo J, Portbury AL, Schisler JC, Patterson C. Brothers and sisters: molecular insights into arterial-venous heterogeneity. *Circulation research.* Oct 24 2008;103(9):929-939.
127. Kumaran V, Benten D, Follenzi A, Joseph B, Sarkar R, Gupta S. Transplantation of endothelial cells corrects the phenotype in hemophilia A mice. *J Thromb Haemost.* Sep 2005;3(9):2022-2031.
128. Follenzi A, Benten D, Novikoff P, Faulkner L, Raut S, Gupta S. Transplanted endothelial cells repopulate the liver endothelium and correct the phenotype of hemophilia A mice. *J Clin Invest.* Mar 2008;118(3):935-945.
129. Shahani T, Covens K, Lavend'homme R, et al. Human liver sinusoidal endothelial cells but not hepatocytes contain FVIII. *J Thromb Haemost.* Sep 30 2013.
130. Pandey GS, Yanover C, Miller-Jenkins LM, et al. Endogenous factor VIII synthesis from the intron 22-inverted F8 locus may modulate the immunogenicity

of replacement therapy for hemophilia A. *Nature medicine.* Oct 2013;19(10):1318-1324.

131. Aird WC. Endothelial cell heterogeneity. *Cold Spring Harbor perspectives in medicine.* Jan 2012;2(1):a006429.

132. Yamamoto K, de Waard V, Fearns C, Loskutoff DJ. Tissue distribution and regulation of murine von Willebrand factor gene expression in vivo. *Blood.* Oct 15 1998;92(8):2791-2801.

133. Aird WC, Edelberg JM, Weiler-Guettler H, Simmons WW, Smith TW, Rosenberg RD. Vascular bed-specific expression of an endothelial cell gene is programmed by the tissue microenvironment. *J Cell Biol.* Sep 8 1997;138(5):1117-1124.

134. Chi JT, Chang HY, Haraldsen G, et al. Endothelial cell diversity revealed by global expression profiling. *Proc Natl Acad Sci U S A.* Sep 16 2003;100(19):10623-10628.

135. Kane NM, Xiao Q, Baker AH, Luo Z, Xu Q, Emanueli C. Pluripotent stem cell differentiation into vascular cells: a novel technology with promises for vascular re(generation). *Pharmacology & therapeutics.* Jan 2011;129(1):29-49.

136. Bai H, Wang ZZ. Directing human embryonic stem cells to generate vascular progenitor cells. *Gene Ther.* Jan 2008;15(2):89-95.

137. Du LM, Nurden P, Nurden AT, et al. Platelet-targeted gene therapy with human factor VIII establishes haemostasis in dogs with haemophilia A. *Nature communications.* 2013;4:2773.

138. Chuah MK, Nair N, VandenDriessche T. Recent progress in gene therapy for hemophilia. *Human gene therapy.* Jun 2012;23(6):557-565.

139. Vandendriessche T, Chuah MK. Targeting endothelial cells by gene therapy. *Blood.* Sep 19 2013;122(12):1993-1994.

140. Zheng C, Page RC, Das V, et al. Structural characterization of carbohydrate binding by LMAN1 protein provides new insight into the endoplasmic reticulum export of factors V (FV) and VIII (FVIII). *J Biol Chem.* Jul 12 2013;288(28):20499-20509.

141. Zheng C, Liu HH, Yuan S, Zhou J, Zhang B. Molecular basis of LMAN1 in coordinating LMAN1-MCFD2 cargo receptor formation and ER-to-Golgi transport of FV/FVIII. *Blood.* Sep 3 2010.

142. Miller EA, Beilharz TH, Malkus PN, et al. Multiple cargo binding sites on the COPII subunit Sec24p ensure capture of diverse membrane proteins into transport vesicles. *Cell.* Aug 22 2003;114(4):497-509.

143. Lynch EC. Peripheral Blood Smear. In: Walker HK, Hall WD, Hurst JW, eds. *Clinical Methods: The History, Physical, and Laboratory Examinations*. 3rd ed. Boston1990.

144. Smith NL, Chen MH, Dehghan A, et al. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. *Circulation.* Mar 30 2010;121(12):1382-1392.

145. Gieger C, Radhakrishnan A, Cvejic A, et al. New gene functions in megakaryopoiesis and platelet formation. *Nature.* Dec 8 2011;480(7376):201-208.

146. Dancourt J, Barlowe C. Protein sorting receptors in the early secretory pathway. *Annu Rev Biochem.* 2010;79:777-802.
147. Routledge KE, Gupta V, Balch WE. Emergent properties of proteostasis-COPII coupled systems in human health and disease. *Mol Membr Biol.* Nov 2010;27(8):385-397.
148. Jensen D, Schekman R. COPII-mediated vesicle formation at a glance. *J Cell Sci.* Jan 1 2011;124(Pt 1):1-4.
149. Chen XW, Wang H, Bajaj K, et al. SEC24A deficiency lowers plasma cholesterol through reduced PCSK9 secretion. *eLife.* 2013;2:e00444.
150. Zanetti G, Prinz S, Daum S, et al. The structure of the COPII transport-vesicle coat assembled on membranes. *eLife.* 2013;2:e00951.
151. O'Donnell J, Maddox K, Stagg S. The structure of a COPII tubule. *J Struct Biol.* Feb 2011;173(2):358-364.
152. Russell C, Stagg SM. New insights into the structural mechanisms of the COPII coat. *Traffic.* Mar 2010;11(3):303-310.
153. Stagg SM, Gurkan C, Fowler DM, et al. Structure of the Sec13/31 COPII coat cage. *Nature.* Jan 12 2006;439(7073):234-238.
154. Stagg SM, LaPointe P, Razvi A, et al. Structural basis for cargo regulation of COPII coat assembly. *Cell.* Aug 8 2008;134(3):474-484.
155. Fath S, Mancias JD, Bi X, Goldberg J. Structure and organization of coat proteins in the COPII cage. *Cell.* Jun 29 2007;129(7):1325-1336.
156. Bhattacharya N, J OD, Stagg SM. The structure of the Sec13/31 COPII cage bound to Sec23. *Journal of molecular biology.* Jul 20 2012;420(4-5):324-334.
157. Gingras AC, Gstaiger M, Raught B, Aebersold R. Analysis of protein complexes using mass spectrometry. *Nat Rev Mol Cell Biol.* Aug 2007;8(8):645-654.
158. Yan W, Aebersold R, Raines EW. Evolution of organelle-associated protein profiling. *J Proteomics.* Feb 15 2009;72(1):4-11.
159. Schirmer EC, Gerace L. Organellar proteomics: the prizes and pitfalls of opening the nuclear envelope. *Genome Biol.* 2002;3(4):REVIEWS1008.
160. Ong SE, Blagoev B, Kratchmarova I, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics.* May 2002;1(5):376-386.
161. Foster LJ, De Hoog CL, Mann M. Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Proc Natl Acad Sci U S A.* May 13 2003;100(10):5813-5818.
162. Borner GH, Harbour M, Hester S, Lilley KS, Robinson MS. Comparative proteomics of clathrin-coated vesicles. *J Cell Biol.* Nov 20 2006;175(4):571-578.
163. Dreze M, et al. High-Quality Binary Interactome Mapping. In: Weissman J, Guthrie, C., Fink, G., ed. *Guide to Yeast Genetics: Functional Genomics, Proteomics, and Other Systems Analysis.* Vol 470: Elsevier Inc.; 2010:282-313.
164. Nyfeler B, Michnick SW, Hauri HP. Capturing protein interactions in the secretory pathway of living cells. *Proc Natl Acad Sci U S A.* May 3 2005;102(18):6350-6355.
165. Church WR, Jernigan RL, Toole J, et al. Coagulation factors V and VIII and ceruloplasmin constitute a family of structurally related proteins. *Proc Natl Acad Sci U S A.* Nov 1984;81(22):6934-6937.

166. Baines AC. *Characterization of mouse models to study protein trafficking in the early secretory pathway*. Ann Arbor: Cellular and Molecular Biology, University of Michigan; 2009.
167. Wendeler MW, Paccaud JP, Hauri HP. Role of Sec24 isoforms in selective export of membrane proteins from the endoplasmic reticulum. *EMBO Rep.* Mar 2007;8(3):258-264.
168. Yu H, Braun P, Yildirim MA, et al. High-quality binary protein interaction map of the yeast interactome network. *Science.* Oct 3 2008;322(5898):104-110.
169. Rual JF, Venkatesan K, Hao T, et al. Towards a proteome-scale map of the human protein-protein interaction network. *Nature.* Oct 20 2005;437(7062):1173-1178.
170. Fields S, Song O. A novel genetic system to detect protein-protein interactions. *Nature.* Jul 20 1989;340(6230):245-246.
171. Rual JF, Hirozane-Kishikawa T, Hao T, et al. Human ORFeome version 1.1: a platform for reverse proteomics. *Genome Res.* Oct 2004;14(10B):2128-2135.
172. hORFeome Database. 2013; http://horfdb.dfci.harvard.edu/. Accessed March 25, 2013, 2013.
173. Wang X, Zhang X, Dong XP, et al. TPC proteins are phosphoinositide- activated sodium-selective ion channels in endosomes and lysosomes. *Cell.* Oct 12 2012;151(2):372-383.
174. Rutkowski DT, Arnold SM, Miller CN, et al. Adaptation to ER stress is mediated by differential stabilities of pro-survival and pro-apoptotic mRNAs and proteins. *PLoS Biol.* Nov 2006;4(11):e374.
175. Graham JM. Isolation of Golgi membranes from tissues and cells by differential and density gradient centrifugation. *Curr Protoc Cell Biol.* May 2001;Chapter 3:Unit 3 9.
176. Fermin D, Basrur V, Yocum AK, Nesvizhskii AI. Abacus: a computational tool for extracting and pre-processing spectral count data for label-free quantitative proteomic analysis. *Proteomics.* Apr 2011;11(7):1340-1345.
177. Choi H, Ghosh D, Nesvizhskii AI. Statistical validation of peptide identifications in large-scale proteomics using the target-decoy database search strategy and flexible mixture modeling. *J Proteome Res.* Jan 2008;7(1):286-292.
178. Gingras AC, Nesvizhskii A. Protein complexes and interaction networks. *Proteomics.* May 2012;12(10):1475-1477.
179. Boyadjiev SA, Fromme JC, Ben J, et al. Cranio-lenticulo-sutural dysplasia is caused by a SEC23A mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking. *Nat Genet.* Oct 2006;38(10):1192-1197.
180. Schwarz K, Iolascon A, Verissimo F, et al. Mutations affecting the secretory COPII coat component SEC23B cause congenital dyserythropoietic anemia type II. *Nat Genet.* Aug 2009;41(8):936-940.
181. Baines AC, Adams EJ, Zhang B, Ginsburg D. Disruption of the Sec24d gene results in early embryonic lethality in the mouse. *PLoS One.* 2013;8(4):e61114.
182. Wansleeben C, Feitsma H, Montcouquiol M, Kroon C, Cuppen E, Meijlink F. Planar cell polarity defects and defective Vangl2 trafficking in mutants for the COPII gene Sec24b. *Development.* Apr 2010;137(7):1067-1073.

183. Snider J, Kittanakom S, Damjanovic D, Curak J, Wong V, Stagljar I. Detecting interactions with membrane proteins using a membrane two-hybrid assay in yeast. *Nat Protoc.* Jul 2010;5(7):1281-1293.
184. Vidal M, Cusick ME, Barabasi AL. Interactome networks and human disease. *Cell.* Mar 18 2011;144(6):986-998.
185. Sprinzak E, Sattath S, Margalit H. How reliable are experimental protein-protein interaction data? *Journal of molecular biology.* Apr 11 2003;327(5):919-923.
186. Stagljar I, Fields S. Analysis of membrane protein interactions using yeast-based technologies. *Trends Biochem Sci.* Nov 2002;27(11):559-563.
187. Kashino Y. Separation methods in the analysis of protein membrane complexes. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences.* Nov 25 2003;797(1-2):191-216.
188. *Hemostasis and thrombosis : basic principles and clinical practice.* 6th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2012.
189. Forsyth S, Horvath A, Coughlin P. A review and comparison of the murine alpha1-antitrypsin and alpha1-antichymotrypsin multigene clusters with the human clade A serpins. *Genomics.* Mar 2003;81(3):336-345.
190. Mann M. Functional and quantitative proteomics using SILAC. *Nat Rev Mol Cell Biol.* Dec 2006;7(12):952-958.
191. Wu CC, MacCoss MJ, Howell KE, Matthews DE, Yates JR, 3rd. Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal Chem.* Sep 1 2004;76(17):4951-4959.
192. Klumperman J. Transport between ER and Golgi. *Current opinion in cell biology.* Aug 2000;12(4):445-449.
193. Shalem O, Sanjana NE, Hartenian E, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science.* Jan 3 2014;343(6166):84-87.
194. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. *Science.* Jan 3 2014;343(6166):80-84.
195. Maynard DM, Heijnen HF, Horne MK, White JG, Gahl WA. Proteomic analysis of platelet alpha-granules using mass spectrometry. *J Thromb Haemost.* Sep 2007;5(9):1945-1955.
196. Hernandez-Ruiz L, Valverde F, Jimenez-Nunez MD, et al. Organellar proteomics of human platelet dense granules reveals that 14-3-3zeta is a granule protein related to atherosclerosis. *J Proteome Res.* Nov 2007;6(11):4449-4457.
197. Onodera K, Shavit JA, Motohashi H, Yamamoto M, Engel JD. Perinatal synthetic lethality and hematopoietic defects in compound mafG::mafK mutant mice. *EMBO J.* Mar 15 2000;19(6):1335-1345.
198. Roux KJ, Kim DI, Raida M, Burke B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol.* Mar 19 2012;196(6):801-810.
199. Predonzani A, Arnoldi F, Lopez-Requena A, Burrone OR. In vivo site-specific biotinylation of proteins within the secretory pathway using a single vector system. *BMC biotechnology.* 2008;8:41.