DEFINING THE ROLES OF MAMMALIAN COPII COMPONENTS SEC24C AND SEC24D

by

Elizabeth Janine Adams

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in
(Cellular and Molecular Biology)
in the University of Michigan
2014

Doctoral Committee:

Professor David Ginsburg, Chair
Professor J. Douglas Engel
Professor Robert S. Fuller
Professor Daniel J. Klionsky
Professor Billy Tsai

DEDICATION

To my parents, with my deepest love and gratitude.

ACKNOWLEDGMENTS

I would like to thank the members of my Dissertation Committee (David Ginsburg, Billy Tsai, Doug Engel, Dan Klionsky, and Bob Fuller) for their time, encouragement, suggestions, support, and helpful and enjoyable discussions. The mentorship I received from my committee members was above and beyond what I ever expected, and I am most grateful. I'm not sure how many PhD students could say that they looked forward to and enjoyed their committee meetings, but I know I certainly did. I would especially like to thank David Ginsburg for not only being an inspiring scientific role model, but an amazing mentor as well. Thank you for your words of encouragement, your patience and compassion, your support, and all of your advice over the past five years!

I would also like to thank the Cellular and Molecular Biology Graduate Program (Bob Fuller, Cathy Mitchell, and Jessica Schwartz) for all of their help, support, and funding.   A special thank you to Bob Fuller, who has just recently taken over as CMB Director, but has always been deeply involved in the interests of CMB students. I would like to thank a number of other people at the University of Michigan for all of their help with this work, including Thom Saunders, Elizabeth Hughes, Judy Poore, Marta Dzaman, and Sue O'Shea. Thank you to Patrick Hu, for continuing to be a trusted mentor to me

I would like to thank all of my friends I have made along the way, and especially those in Michigan for making my time in grad school one for the books! A special thank you to Jon, for being right beside me through all of the ups and downs of grad school. I am so grateful for your unwavering love and support.

And finally… I would like to thank my family for all their love and support and continuous encouragement throughout my life. Without them, I certainly would not be where I am today. My mom and dad have always encouraged me be the best I can possibly be, and to always believe in myself. Dave, Phill, and Amanda have been there for me through thick and thin, and I'm so grateful. I would also like to thank all of my aunts and uncles and cousins for all of their love and never-ending support. And of course, I would like to thank Nana, my traveling companion, and one of my very best friends, for inspiring me to live life to the fullest.

Table of Contents

LIST OF TABLES

ABSTRACT

**DEFINING THE ROLES OF MAMMALIAN COPII COMPONENTS SEC24C
AND SEC24D**

by

Elizabeth Janine Adams

Chair: David Ginsburg

Newly synthesized proteins are transported from the endoplasmic reticulum (ER) to the Golgi via COPII coated vesicles. The COPII coat is a highly conserved and tightly regulated complex composed of an inner layer containing the GTPase Sar1 and heterodimers of SEC23 and SEC24, and an outer coat made up of heterotetramers of SEC13 and SEC31. The SEC24 subunit is thought to be primarily responsible for recruitment of protein cargoes into nascent vesicles through direct contacts between SEC24 and the cytoplasmic portion of transmembrane cargo proteins or through contact with cargo adaptor proteins that are essential in linking soluble cargo proteins within the ER lumen to SEC24. Mammalian genomes encode four *Sec24* paralogs (*Sec24a-d)*, though little is known about their comparative functions. A number of biochemical and structural studies have identified a handful of paralog-specific interactions, but many more remain are likely uncharacterized, given the highly heterogeneous pool of proteins traversing the secretory pathway. Based on protein sequence, SEC24A/B are more closely related to one another than they are to SEC24C/D. Mice deficient for SEC24A

exhibit a mild hypocholesterolemic phenotype, and *Sec24b* null embryos die from a specific neural tube closure defect near the end of embryonic development.

Through the generation and characterization of a mouse model deficient in SEC24C, we demonstrate that SEC24C is required during early embryonic development, with SEC24C-deficient embryos dying around embryonic day 7.5. However, we found SEC24C to be dispensable in a number of tissues, likely as a result of compensation by other *Sec24* paralogs. A mouse model of SEC24D-deficiency revealed that the requirement for SEC24D begins prior to the blastocyst stage. Mice heterozygous for null alleles of *Sec24c* or *Sec24d*, as well as compound heterozygotes, exhibit normal growth, development and survival, and no obvious phenotypic abnormalities. We sought to determine the extent of functional overlap between SEC24C and SEC24D by generating a *Sec24c^{c-d}* allele, in which the SEC24C coding sequence has been largely replaced with SEC24D. Crossing mice with the *Sec24c^{c-d}* allele to *Sec24c* mice demonstrates that SEC24D, when its expression is driven by *Sec24c* regulatory elements, can rescue *Sec24c* null mice from embryonic lethality. However, *Sec24c^{c-d/c-d}* mice die shortly after birth, suggesting that the overlap in function between SEC24C and SEC24D is incomplete, consistent with the wide range of phenotypes observed in mouse models of SEC24 deficiency. Taken together, these results indicate that the four *Sec24* paralogs have developed unique functions over the course of vertebrate evolution, but have also maintained some overlap in function. Future work to understand the functional differences and cargo-specificities between the mammalian SEC24 paralogs will provide further insight into the dynamic process of cargo recruitment and COPII assembly.

CHAPTER I: INTRODUCTION

Abstract

Newly synthesized proteins are transported from the endoplasmic reticulum (ER) to the Golgi via COPII coated vesicles that assemble at specific sites on the ER. The COPII coat is a highly conserved and tightly regulated complex composed of an inner layer containing the GTPase Sar1 and heterodimers of SEC23 and SEC24, and an outer coat made up of heterotetramers of the cytosolic proteins SEC13 and SEC31. The SEC24 subunit of the inner coat complex is thought to be primarily responsible for recruitment of protein cargoes into nascent vesicles. This action is mediated by direct interaction of SEC24 with the cytoplasmic portion of transmembrane cargo proteins or through contact with cargo adaptor proteins that are essential in linking soluble cargo proteins within the ER lumen to SEC24. The mammalian genome encodes four *Sec24* paralogs (*Sec24a-d)*, and while their overall structure is conserved, little is known about their comparative functions. Based on protein sequence, SEC24A and B are more closely related to one another than they are to SEC24C and D. *Sec24b* and *Sec24d* knockout mice exhibit embryonic lethality, while *Sec24a* knockouts have low cholesterol levels due to reduced secretion of PCSK9.   A number of biochemical and structural studies have identified a handful of paralog-specific interactions, but it stands to reason that many more have yet to be characterized, given the highly heterogeneous pool of proteins

traversing the secretory pathway.   Understanding the overlap in function between the mammalian *Sec24* paralogs will provide further insight into the dynamic process of cargo recruitment and COPII assembly.

COPII Vesicle Formation

Nearly one-third of all proteins traverse the secretory pathway en route to their final destinations, either within the cell or beyond [1,2]. Included in this subset of proteins are all transmembrane proteins destined for certain intracellular compartments or the plasma membrane, as well as soluble proteins intended for secretion or to be located in some membrane bound intracellular compartments (Figure 1-1).   These newly synthesized proteins begin their journey in the endoplasmic reticulum (ER), where they are folded and undergo initial post-translational modifications. Once properly folded, these proteins are recruited to ribosome-free regions of the ER called ER-exit sites (ERES) [1,3,4,5]. It is here that the process of COPII vesicle formation occurs, and a cargo protein's journey through the secretory pathway begins.

Sar1-GTP, Sec23/24 and Sec13/31 are the minimal machinery required for COPII vesicle budding *in vitro* [6]. COPII vesicle formation is initiated by the activation of the small GTPase SAR1 by its guanine nucleotide exchange factor (GEF), Sec12, which facilitates the binding of GTP to the inactive, GDP-bound Sar1 (Figure 1-2). Sec12 is localized to ERES by the protein SEC16, a cytosolic factor required for the *in vivo* formation of COPII vesicles [7,8]. GTP binding to Sar1 induces a conformational change, leading to the insertion of an N-terminal, amphipathic helix into the ER membrane, initiating membrane curvature [9,10]. Activated Sar1 recruits the cytosolic heterodimer of Sec23/Sec24 to form the pre-budding complex via GTP-Sar1's interaction

with Sec23[11,12]. Transmembrane cargoes and adaptors interact with Sec24 and become concentrated in the nascent vesicle bud. Vesicle formation continues as the cytosolic heterotetramer of Sec13/31 binds to the inner layer of Sec23/24 via a Sec23-Sec31 interaction; this outer coat of Sec13/31 further drives membrane curvature initiated by Sar1-GTP [13].   As the self-assembly of the COPII coat nears completion, GTPase-activating protein (GAP) activity of Sec23 triggers GTP hydrolysis, and disassembly of the coat begins [14,15]. The final step in the COPII process is vesicle fission, and once free of the donor membrane, the COPII vesicle travels to the ER-Golgi Intermediate Complex (ERGIC).

Selective Recruitment Model

Despite the extensive studies of the early secretory pathway and the formation of the COPII coat, there remains at least one aspect of ER export that is still a point of contention, and that is the mechanism by which proteins destined for ER export end up in a newly forming COPII vesicle. Two models have been proposed: selective recruitment and bulk flow [16,17,18].   In the selective recruitment model, cargo molecules containing ER exit motifs interact with SEC24 directly or via cargo adaptors that function to recruit them to ERES and package them into nascent COPII vesicles by acting as a physical linker to the cytoplasmic coat. The bulk flow model proposed that secreted proteins are packaged into COPII vesicles in a stochastic manner, with unfolded and ER resident proteins precluded from these vesicles by their interaction with other ER protein machinery. Although there is some evidence to suggest that export of proteins from the ER by bulk flow may be sufficient for a number of highly abundant proteins [19,20,21], a growing body of evidence suggests that the packaging of the majority of cargo proteins

into newly forming vesicles is mediated by selective transport [22,23]. The random nature of bulk flow fails to explain the efficiency of ER export, and an increasing number of ER exit motifs have been identified (discussed in detail in a later section) further supporting the model of selective transport.

Expansion of SEC24 Paralogs

While the components of the COPII coat are highly conserved, multiple paralogs for several of these genes have arisen in higher eukaryotes (Figure 1-2), potentially driven by the increased complexity and diversity of cargoes passing through the early secretory pathway. A primary example of this evolutionary expansion is the major cargo-binding COPII component, SEC24. The yeast genome encodes the canonical Sec24p, but also two non-essential homologs, Iss1p and Lst1p, both of which are able to bind to Sec23p [24,25]. All three paralogs are able to bind to the SNARE Sed5/syntaxin-5 [25]. However, Lst1p can form heterodimers with yeast Sec23p and facilitate the formation of larger vesicles to accommodate lager cargoes such as ATPase Pma1 [24]. The mammalian genome encodes four Sec24p homologs (*Sec24a-d)*. Like Sec24p (and Lst1p, Iss1p), each of the four mouse paralogs contain a hypervariable N-terminal region, while the C-terminal half of the protein displays a higher degree of sequence conservation, a region containing several predicted conserved domains (Figure 1-3). Based on protein sequence identity, Sec24p is more closely related to mammalian SEC24A and SEC24B, and Iss1p, while Lst1p has more homology to SEC24C and SEC24D (Figure 1-4). Interestingly, the mammalian SEC24A/B subgroup is no more similar to SEC24C/D than is either subgroup to the ancestral Sec24p, suggesting an early duplication and divergence between these two subfamilies. The amplification of COPII

paralogs in higher eukaryotes might result in an increased capacity to handle the greater diversity of secretory proteins in mammalian cells compared to yeast, with various COPII paralogs potentially functioning in a tissue-specific and/or cell-type-specific fashion to direct selective transport of a wide range of cargo proteins.

In the mouse, a diverse range of phenotypes is observed in the absence of any single paralog of SEC24. Loss of SEC24A disrupts secretion of the regulatory protein PCSK9, resulting in low circulating cholesterol levels in SEC24A-deficient mice [26]. These mice are otherwise healthy with no other apparent phenotypic abnormalities. In contrast, loss of SEC24B, C, and D all result in lethality at different stages of development. A mutation in murine SEC24B that results in a truncated form of the protein was identified in a genetic screen as the cause of a very specific neural tube closure defect resulting from a block in secretion of the planar-cell-polarity protein VANGL2 [27]. Loss of SEC24C results in embryonic lethality ~E7.5 (described in more detail in Chapter III), while SEC24D-deficient embryos are lost prior to the blastocyst stage [28], discussed in Chapter II.

In both *C. elegans* and *D. melanogaster* two paralogs of Sec24p (*sec24.1/sec24.2* and *sec24/CG10882*) are present, while in fish there are four SEC24 paralogs (SEC24A-D), similar to mammals. Mutagenesis screens for craniofacial defects in zebrafish uncovered a role for the *Sec24* paralogs in proper secretion of extracellular matrix collagens [29,30]. The *bulldog* phenotype results from a frame shift mutation and early stop codon in the zebrafish *Sec24d* gene, with *bulldog* mutants exhibiting craniofacial defects due to abnormalities in collagen secretion [31]. A second mutant, *crusher*, due to a substitution in *Sec23a* [32] which closely resembles the *bulldog* phenotype, again

consistent with a critical role for COPII transport in the secretion of collagen and other extracellular matrix proteins in zebrafish. A medaka mutant, *vertebra imperfecta,* due to a nonsense mutation in SEC24D, also exhibits skeletal defects and abnormal craniofacial cartilage [33]. In *Arabadopsis thaliana,* the loss of SEC24 is lethal. However, a missense mutation (R693K) in the conserved binding pocket of Sec24 resulted in the accumulation of cargoes in an unusually distended ER [34].

To date there have been no human diseases associated with mutations in any of the four SEC24 paralogs. Given the frequency of potentially damaging alleles in the general population (0.07% for *SEC24A, 0.02%* for *SEC24B, 0.05%* for *SEC24C,* and 0.9% for *SEC24D,* data from Exome Variant Server, NHLBI), it remains possible that mutations in any human SEC24 paralog could be responsible for a human disease for which the genetic cause is unknown. The loss of a particular paralog in humans may also result in a very mild phenotype that wouldn't necessarily be seen clinically. Or it is possible that the complete loss of function of any particular paralog would result in embryonic lethality as is observed in the mice, thus no viable homozygotes or compound heterozygotes would ever be seen.

SEC24 recruits cargoes directly or via cargo adaptor molecules

SEC24 is the COPII component primarily responsible for the recruitment of cargo proteins into the newly forming vesicle bud. Given the topology of a COPII coated vesicle, the interaction between SEC24 and a potential cargo can occur in one of two ways: (i) directly, if the protein possesses a cytoplasmic tail that can come into contact with SEC24 via an ER exit motif, or (ii) indirectly if the potential cargo does not have a cytoplasmic tail to interact with cytosolic SEC24. Proteins in the second class require an

adaptor molecule to span the ER membrane and facilitate interaction with the cytosolic coat. Data from the human proteome suggests that approximately 11% of the 24,701 predicted full-length open reading frames in the human proteome are thought to be soluble, secreted proteins which would require a cargo adaptor, while about 20% are predicted to be transmembrane proteins [35], suggesting that thousands of distinct cargo molecules must be able to be recruited via interactions directly with SEC24 or via cargo adaptors. Once a COPII vesicle reaches the ERGIC or the Golgi compartment, many of these receptors are then recycled back to the ER for another round of cargo recruitment. Several cargo adaptors have been identified in yeast, and mutations in these proteins result in defective transport of specific cargoes reviewed in [36,37]. Similar results are also seen when homologous proteins in Drosophila are mutated [36]. However, the only clearly defined mammalian cargo adaptor is ERGIC-53, also known as LMAN1, part of the LMAN1-MCFD2 complex that is responsible for the selective transport of the blood coagulation proteins FV and FVIII [38,39]. The highly conserved p24 family of proteins have a similar topology to LMAN1, appear to cycle between the ER and the Golgi, and selective trafficking defects occur when these proteins are mutated. However, these proteins may play another role within the early secretory pathway, as there is no direct evidence that they function as professional cargo adaptors [36,40].

Many additional cargo adaptors may yet remain to be identified, each contributing to the COPII system's ability to accommodate the diverse array of secretory proteins exiting the ER. It is possible that these proteins have other primary functions within the cell and only transiently interact with soluble cargoes in the context of COPII vesicle formation. Alternatively a large class professional cargo adaptors may bind to their

ligands in the ER, facilitate transport to the ERGIC, followed by cargo release and recycling of the adaptor back to the ER.

The structure of SEC24

Detailed structural analysis of the COPII coat has provided critical insights into the formation, regulation, and evolution of this intricate cellular complex to handle the enriched cargo diversity found in mammals. Early electron microscopy (EM) experiments provided a first look at the overall shape of the COPII components and revealed that the inner coat has a bow-tie like shape, with Sec23p forming one half of the bow-tie and Sec24p the other [41]. Using full-length recombinant yeast Sec24p in complex with yeast Sec23p as well as the complex of Sec23p/Sar1, Bi *et al.* determined the structure of the pre-budding complex of Sar1/Sec23/Sec24 at a resolution of approximately 2.75Å [42] that fit nicely with the previously reported EM data. The initial 132 residues of Sec24p were disordered in the crystal, likely due to a high number of glutamine and proline residues in this N-terminal hypervariable region. In this study, the authors found that the pre-budding complex has a concave inner surface potentially contributing to the formation and maintenance of membrane curvature of the nascent vesicle bud. Residues on both Sec23p and Sec24p proximal to the ER membrane were overall positively charged, while the outer face of these proteins was much more acidic. This polarized characteristic is conserved from yeast to humans [42]. Structurally, Sec23p and Sec24p are quite similar to each other, with the same conserved domains including a zinc finger, a von Willebrand factor type A "trunk" domain, a β-barrel, a helical domain, and a gelsolin-like repeat (Figures 1-3 and 1-5), yet strikingly their overall sequence identity is low, at only 14% [42]. The trunk domain forms the

interface of Sec23p and Sec24p, specifically with a single β-strand (β-14) on both Sec23p and Sec24p and covers a 900Å^2 region on the surface of Sec24. Two critical residues in this region Phe385 and Pro 387 are conserved in all known Sec24 sequences from yeast to humans. The Sec23p trunk domain is also responsible for its interaction with Sar1p. However, the distance between these the Sec24p and Sar1p contact points on Sec23p is more than 25Å, sufficient for coordinated formation of the pre-budding complex [42].

Cargo binding sites on SEC24

In the selective transport model, each cargo must contain an export motif or signal that interacts with a cargo recognition site located within SEC24. Four such binding sites on Sec24p have previously been characterized, each critical for the recruitment of a specific set of cargoes (Figure 1-5) [23,43,44,45]. The "A-site" (also known as the Sed5 binding site) is located on the edge of the membrane-proximal region of SEC24, similar to the SAR1 binding site on SEC23, and recognizes cargo displaying the sorting signal YNNSNPF such as the yeast t-SNARE Sed5 [44]. A second site, the "B-site," was shown in structural studies near a basic pocket on the membrane face of SEC24 [44] [46] and is conserved from yeast to humans (though only found in human SEC24A/B) [23]. A third site on SEC24, the "C-site" is required for the binding of the t-SNARE Sec22p, though this interaction depends on the state of SNARE assembly, as the export signal is likely a folded epitope [44,45,46]; of note, the binding of Sec22p at this site is specific to mammalian SEC24A and SEC24B, as this binding site in SEC24C and D is occluded [45]. A fourth site (IxM binding site) is located on the opposite face of SEC24 from the SEC23 interface, about half the distance between the A-site and the B-site (Figure 1-5) [23]. This site is only found on human SEC24C and D; in human SEC24A and B, a loop

from the N-terminal region competes for binding at this site and prevents cargo binding

[23].   While these are the only characterized binding sites to date, it is likely that many

additional sites, as yet undefined, are required for any given subset of cargo molecules,

and they have yet to be defined.

ER exit signals on transmembrane cargo interact with the COPII coat

To be recruited by SEC24 and enriched at ER exit sites, cargo molecules or

adaptors are thought to require an ER export signal that binds to cargo recognition sites

on SEC24. In this way, ER resident proteins lacking an exit motif will be excluded from

the newly forming vesicle. Structural and biochemical studies have identified several

classes of signal motifs have [47]. A subset of transmembrane cargo proteins interacts

with the Sec23p–Sec24p subunit of COPII via a di-acidic D/E-X-D/E motif.   This signal

was originally identified in a series of experiments truncating the 29 amino acids in the

cytoplasmic tail of the vesicular stomatitis virus glycoprotein (VSV-G), followed by an

alanine scan that revealed the requirement for Asp21 and Glu23 for efficient transport of

VSV-G out of the ER [48]. Similar experiments demonstrated the requirement of the

acidic motif DxE on the yeast protein Sys1p for its secretion [49]. Later work

demonstrated that the DxE signal on VSV-G confers a specificity for human SEC24A/B

[23]. Di-acidic motifs have also been found to direct the export of several additional

proteins including the cystic fibrosis transmembrane conductance regulator (CFTR) [50],

the mammalian lysosomal acid phosphatase [48], and the mammalian potassium channels

Kir1.1 and Kir2.1 [51], and the yeast protein Gap1p [22].   Cargoes containing this motif

bind to the B-site on SEC24 [46].

A second class of ER exit motifs that also binds to the B-site on SEC24 is the peptide sequence LXXLE, which was originally found on the v-SNARE Bet1p [52]. A related signal LXXME was identified on the yeast t-SNARE Sed5, which is used in conjunction with YNNSNPF to form a bipartite signal to bind to the A-site of Sec24p in yeast [44]. A third motif (IxM) was identified on the human paralog of Sed5, syntaxin 5. This export signal was found to bind specifically to a groove on the surface of human SEC24C and D [23]; in a study to determine the paralog-specific interactions of SNARE proteins using an *in vitro* budding assay, SEC24C and SEC24D were shown to specifically recruit and package syntaxin 5 and membrin via their IxM exit motif [23], demonstrating overlap in function between SEC24C and D.

Di-hydrophobic export signals have also been identified for COPII mediated transport of several proteins. The FF motif present on LMAN1 is required for its efficient export out of the ER [53], with this motif also found on Prm8p [46]. The FF motif seems to function in the context of a larger export signal, because it is not sufficient to drive export of a reporter protein [53]. An LV motif is necessary for the recruitment and export of the p24 protein Emp24p [54] that has been shown to bind to SEC24's B site. The yeast protein Erv41p requires an IL signal on cytoplasmic tail for proper ER export, while the related protein Erv46p requires both an IL and FY signal [55].

Recent work indicates that the SEC24A/B and SEC-24C/D mammalian subgroups differ in their affinity for a subset of known sorting signals present on cargo proteins. The serotonin receptor (SERT) has been reported to be an exclusive cargo of SEC24C, with this specificity mediated by an export signal on the N-terminus of SERT [56]. Mutation of this site results in SERT preferentially binding to SEC24D instead of

SEC24C [57]. Additionally, it was found that ER export of the GABA transporter 1

(GAT1) is reduced by RNAi knockdown of SEC24D, and mutational analysis suggests

that an RL amino acid sequence motif located near the C-terminus of GAT1 and two

other members of the SLC6 family of neurotransmitter transporters might be a paralog-

specific export motif for SEC24D [58]. RNAi experiments in HeLa cells demonstrated

that SEC24A is required for the transport of reporter constructs containing the LL ER

export motif, while transport of II and VV containing signals have a preference for

SEC24C/D [59]. Finally, it should be noted that although direct cargo binding by COPII

is thought to be mediated by Sec23p–Sec24p, one study reports an interaction between

Sar1 and a specific class of Golgi glycosyltransferases that requires a di-basic R/K-X-

R/K ER export motif [60].

The role of cargoes in COPII vesicle formation

A number of reports suggest that cargoes and cargo adaptors play an active role in

COPII vesicle formation (reviewed in [61,62]). The amount of cargo present in the ER

can influence the number and size of ERESs, which can increase COPII budding activity

in a cell [63,64]. Additionally, cargoes may stabilize the pre-budding complex on the

ER membrane though an interaction with SEC24 despite the inactivation of SAR1

[65,66]. The size and topology of cargo also influence the size of COPII vesicles and the

architecture of the coat itself. Bulky and asymmetric cargo molecules, such as the p24

family and GPI-anchored proteins, oppose the membrane curvature generated by the

COPII coat, requiring more rigidity in the coat conferred by the outer coat component

SEC13 [67]. Larger cargoes such as pro-collagen require additional accessory proteins

such as TANGO, which can bind to both luminal collagen and SEC23, acting as a cargo

adaptor-like molecule, and regulating the formation of extended COPII vesicles or tubules [68].

Other roles for SEC24

In addition to providing a platform for cargo binding, SEC24 may also play a regulatory role as well. A number of Sec24p mutants have shed light on the functions of critical residues and the overall function of Sec24p. The essential ERES protein Sec16 decreases the GTPase activity within the coat by preventing recruitment of Sec31 to the Sec23/23/Sar1 complex. A recently classified mutant (*Sec24-m11)* contains the substitutions E504A and D505A near the A-site on Sec24p; these changes reduce the binding affinity of Sec24p for Sec16 by preventing this interaction between Sec16 and Sec31 and increasing the GTPase activity in the coat, resulting in the formation of smaller vesicles [69]. This observation raises the possibility that cargo binding to Sec24 could play a role in dictating the stability of the COII coat. Akt, a protein kinase central to many cellular processes including glucose metabolism, cell proliferation and apoptosis, was shown to phosphorylate mammalian SEC24C and Sec24D [70], suggesting another potential regulatory role for SEC24. This phosphorylation weakens SEC24 binding to SEC23, thus providing an opportunity to regulate the formation of COPII vesicles in response to cellular cues.   Finally, SEC24C has been implicated in the docking of specialized pre-chylomicron transport vesicles to the Golgi [71], suggesting a role for the COPII coat beyond the recruitment and formation of vesicles at the ER surface.

Concluding Remarks

The formation and regulation of COPII-mediated ER export is an essential and complex process. The fundamental roles of the core proteins in COPII biogenesis have been highly conserved, though there has been an expansion in the number of paralogs, perhaps to compensate for the amplification in the secretory demand of multicellular organisms. This increase in paralogs may also allow higher organisms to form different classes of COPII vesicles composed of different combinations of the coat proteins to deal with the dynamic needs of a given cell or tissue type, though this remains to be proven. Central to this cellular process is the cargo-binding subunit SEC24. Structural and biochemical studies have shed light on the mechanisms by which SEC24 can recognize specific classes ER exit motifs on transmembrane cargoes and adaptor molecules. However, given the great diversity and sheer number of cargo proteins that are transported through the secretory pathway, it is likely that many more cargo recognition sites on SEC24 remain to be discovered. Is the function of a given SEC24 paralog driven by its precise tissue-specific or cell-type specific expression patterns or are there key differences in the proteins themselves (perhaps in the N-terminal region where a lot of variability is present) that govern their ability to recruit particular cargoes into COPII vesicles? While there appears to be partial overlap in function among the mammalian SEC24 paralogs, increasing evidence for paralog-specific interactions are emerging (Table 1-1).
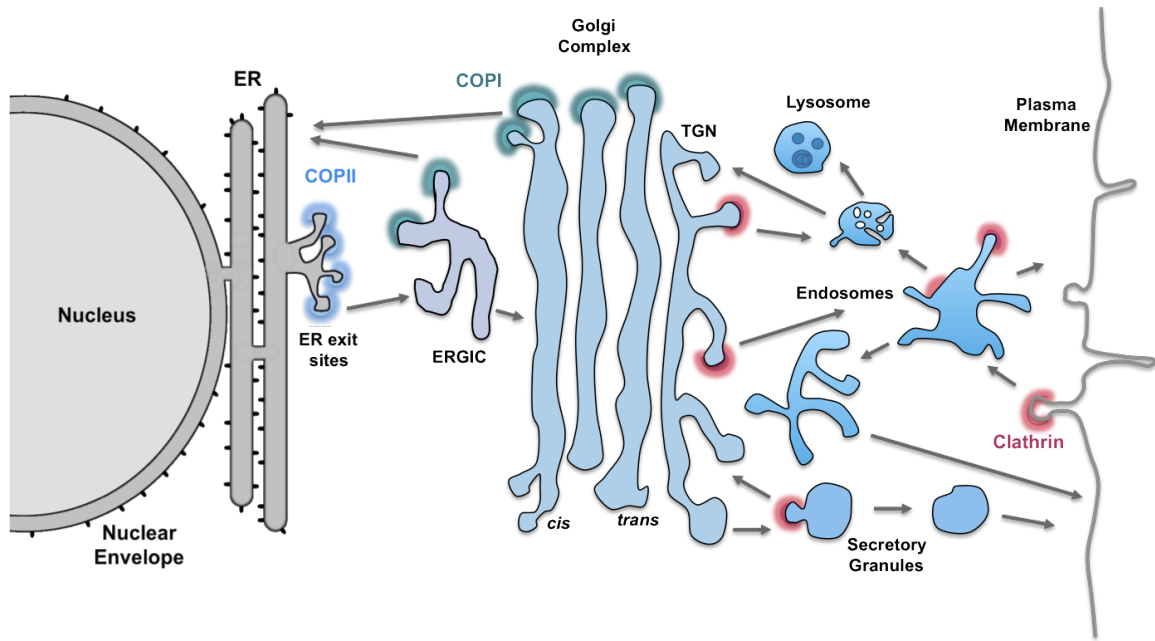
Figure 1-1: The mammalian secretory pathway

A schematic depicting the many compartments of the secretory pathway. Newly synthesized proteins enter the secretory pathway in the endoplasmic reticulum (ER), and travel to the ER-Golgi intermediate compartment (ERGIC) via COPII vesicles (blue). They are then transported to the Golgi, where they undergo further modifications before being trafficked to various intracellular compartments or are secreted outside the cell. COPI vesicles (green) direct retrograde transport from the ERGIC and Golgi to the ER. Transport later in the pathway is primarily mediated by clathrin coats (red). Arrows indicate directional vesicular transport steps. *(Figure adapted from Bonifacino and Glick, Cell, 2004).*
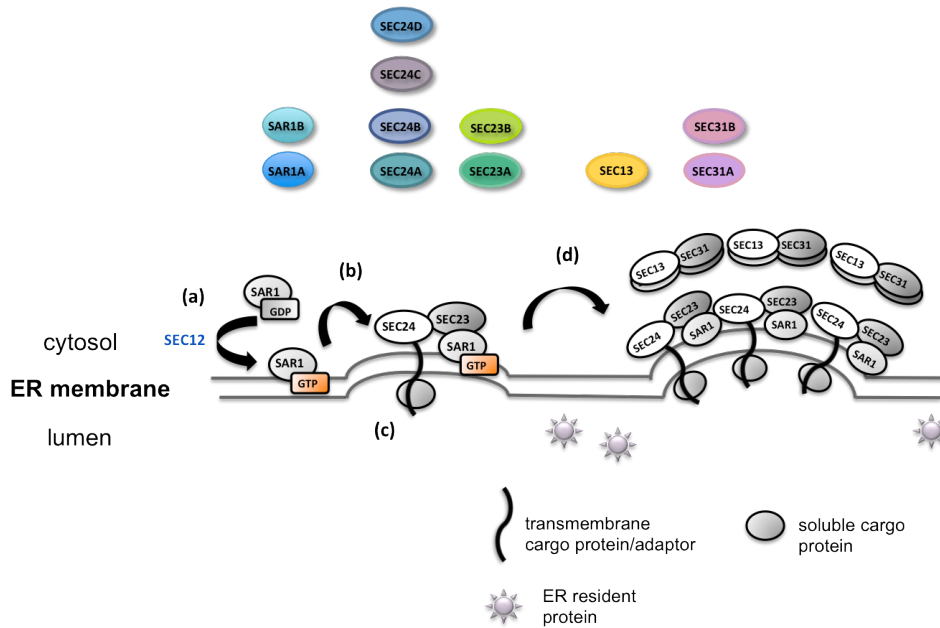
Figure 1-2: COPII vesicle composition and formation

Newly synthesized proteins exit the ER in COPII-coated vesicles, which are composed of the small GTP-binding protein SAR1, the heterodimeric SEC23-SEC24 complex and the heterotetrameric SEC13-SEC31 complex. The mammalian genome encodes multiple paralogs of many of the core components (shown in color). (a) The exchange of GDP for GTP, catalyzed by SEC12, a guanine nucleotide exchange factor located in the ER membrane, activates SAR1, causing it to insert an N-terminal amphipathic helix into the ER membrane. (b) SAR1-GTP recruits the inner coat complex, SEC23-SEC24. (c) SEC24 interacts with the cytoplasmic domains of transmembrane proteins, which may in turn serve as cargo receptors for soluble proteins. (d) The outer coat complex, SEC13-SEC31, polymerizes the COPII vesicle to drive budding from the ER membrane. (*Figure adapted from A. Baines, 2009)*
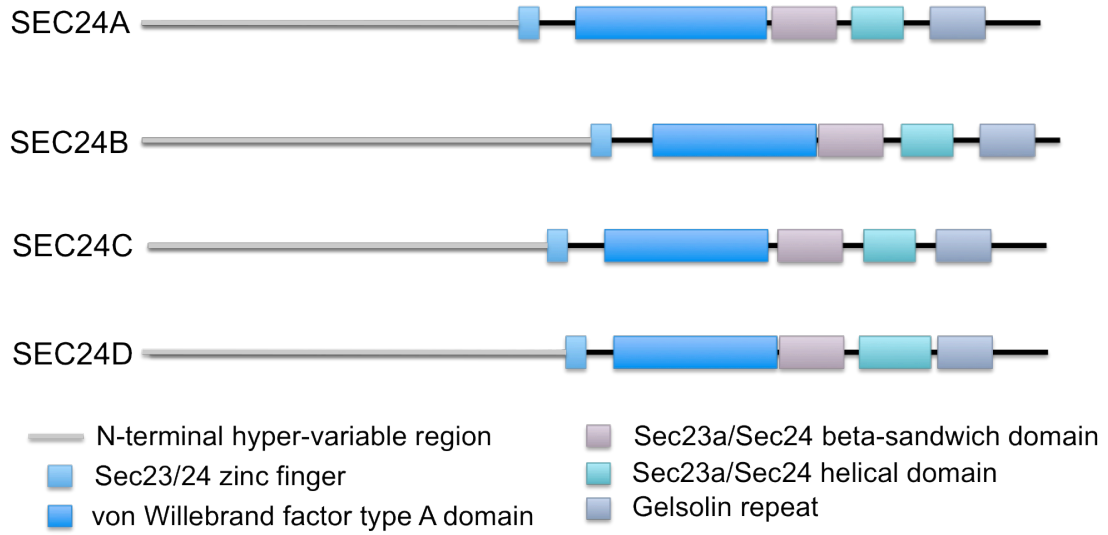
Figure 1-3: Domain Structure of the mammalian SEC24 paralogs

The mammalian SEC24 proteins are shown with the conserved domains highlighted. Though variable in overall protein length, all paralogs of SEC24 share the same conserved domain organization and an N-terminal hypervariable region.

A

58%

mSEC24C
mSEC24D

25% identity with Sec24p

mSEC24A
61%
mSEC24B

30% identity with Sec24p

26% identity

Sec24p

Iss1p

Lst1p

B

	Lst1p	mSEC24C	mSEC24D	mSEC24A	mSEC24B	Sec24p	Iss1p
Lst1p	100.00	25.83	25.96	23.63	23.83	22.35	23.90
mSEC24C		100.00	58.10	27.52	28.36	27.55	27.36
mSEC24D			100.00	27.26	28.47	26.99	25.91
mSEC24A				100.00	60.84	32.48	33.58
mSEC24B					100.00	32.57	33.54
Sec24p						100.00	60.49
Iss1p							100.00
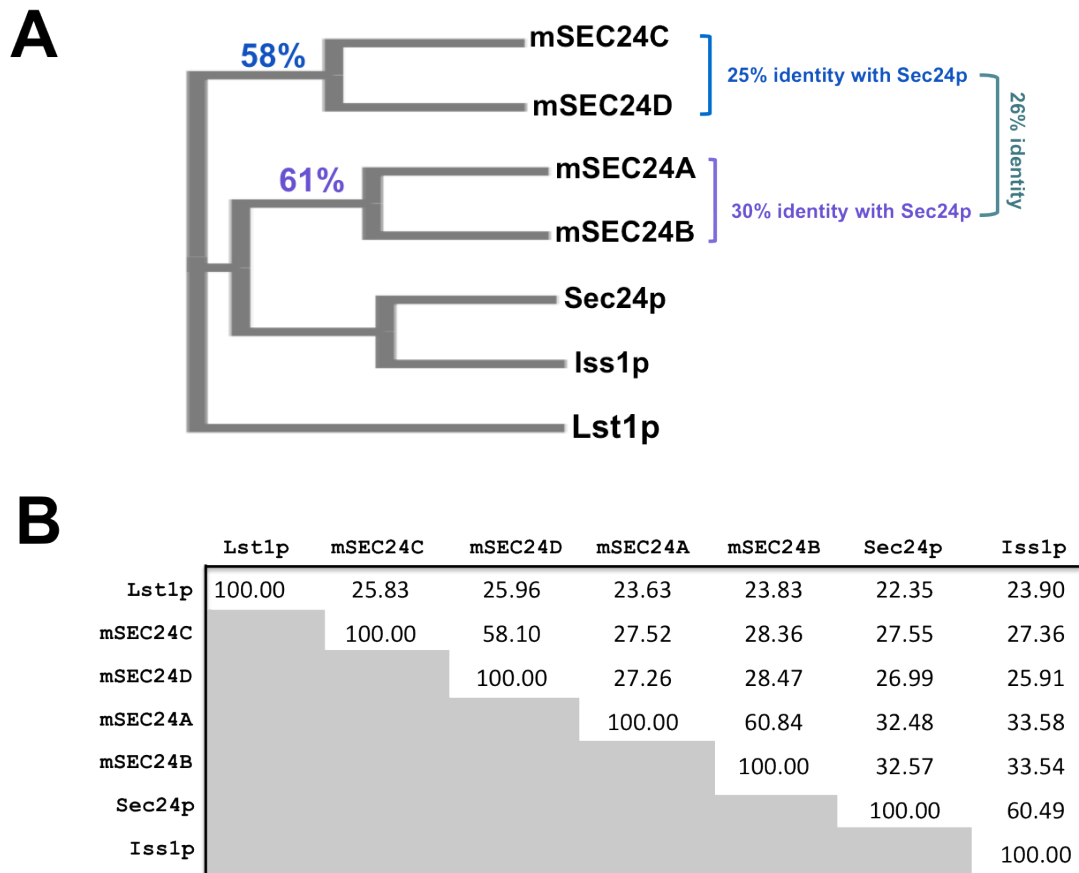
Figure 1-4: Evolution and divergence of the SEC24 paralogs

(A) Phylogram depicting the SEC24A/B and SEC24C/D subfamilies of the mammalian paralogs and their relationship to the ancestral yeast paralog Sec24p. (B) Table displaying percent of protein sequence identity between yeast and mammalian SEC24 paralogs, calculated using ClustalW2 alignment software.
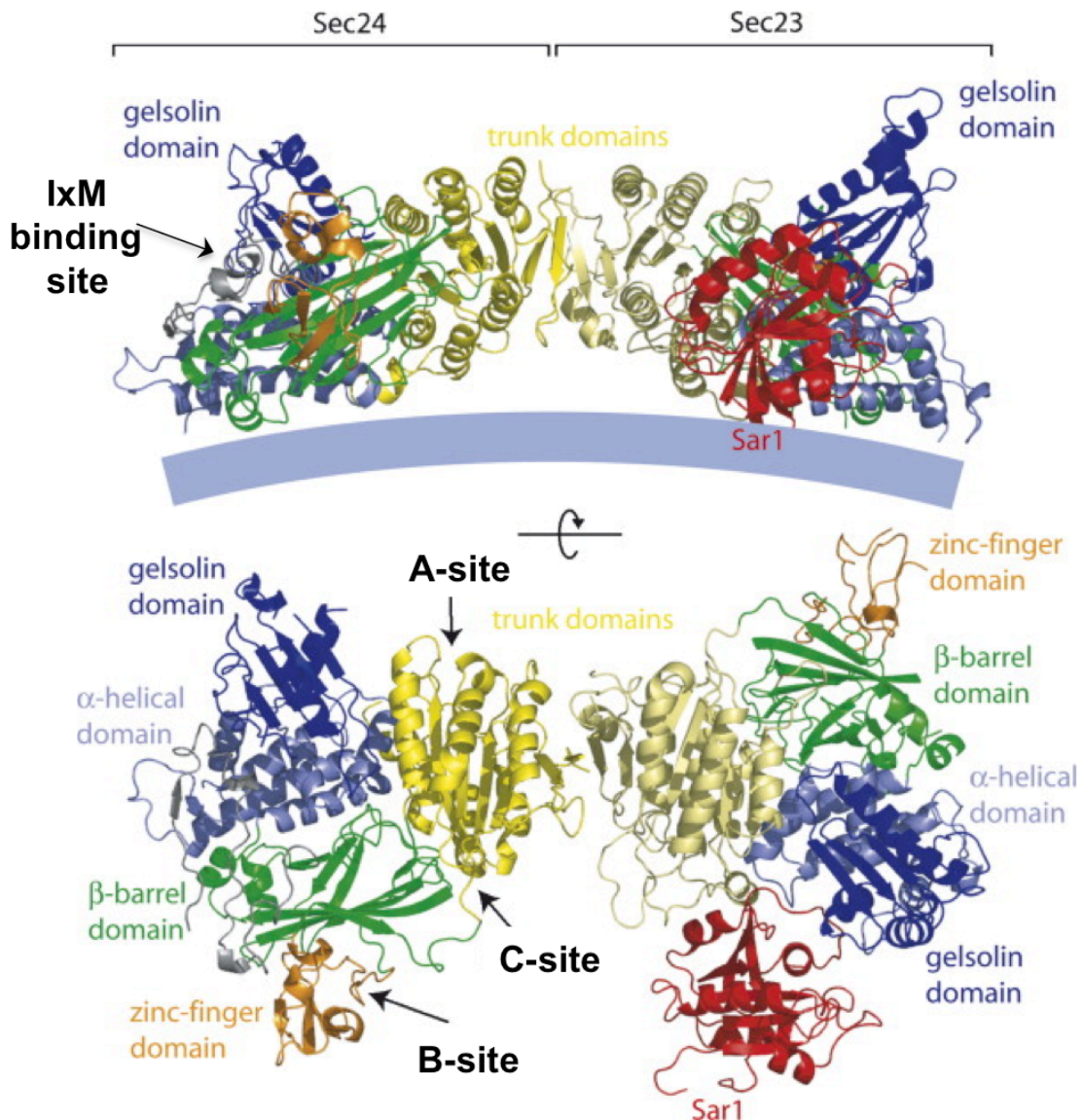
Figure 1-5: Localization of characterized binding sites on the structure of SEC24

The four characterized cargo recognition sites depicted on SEC24 are shown on the ribbon diagram of the SEC23-SEC24 heterodimeric complex. The A-site is located within the trunk domain of SEC24, near the interface with SEC23. The B-site is located on the membrane face near a basic pocket of SEC24. The C-site is also near the trunk domain, and the IxM recognition site is on the opposite face from SEC23, halfway between the A- and B-sites. *(Adapted from Lee and Miller, Seminars in Cell & Developmental Biology, 2007)*

Protein or Motif	Class	SEC24 preference	Species
Pmap1	transmembrane	Lst1p	yeast
PCSK9	soluble	SEC24A	mouse
VANGL2	transmembrane	SEC24B	mouse
SEC22	SNARE	SEC24A/B	human
(IxM) motif	export signal	SEC24C/D	human
(DxE) motif	export signal	SEC24A/B	human
(LxxLE) motif	export signal	SEC24A/B	human
SERT	transmembrane	SEC24C	human
GAT1	transmembrane	SEC24D	human

Table 1-1: Paralog-specific interactions characterized to date.

CHAPTER II: DISRUPTION OF THE *SEC24D* GENE RESULTS IN EARLY EMBRYONIC LETHALITY IN THE MOUSE

ABSTRACT

Transport of newly synthesized proteins from the endoplasmic reticulum (ER) to the Golgi is mediated by the coat protein complex COPII. The inner coat of COPII is assembled from heterodimers of SEC23 and SEC24.   Though mice with mutations in one of the four *Sec24* paralogs, *Sec24b*, exhibit a neural tube closure defect, deficiency in humans or mice has not yet been described for any of the other *Sec24* paralogs.   We now report characterization of mice with targeted disruption of *Sec24d*.   Early embryonic lethality is observed in mice completely deficient in SEC24D, while a hypomorphic *Sec24d* allele permits survival to mid-embryogenesis. Mice haploinsufficient for *Sec24d* exhibit no phenotypic abnormality.   A BAC transgene containing *Sec24d* rescues the embryonic lethality observed in *Sec24d*-null mice.   These results demonstrate an absolute requirement for SEC24D expression in early mammalian development that is not compensated by the other three *Sec24* paralogs. The early embryonic lethality resulting from loss of SEC24D in mice contrasts with the previously reported mild skeletal phenotype of SEC24D deficiency in zebrafish and restricted neural tube phenotype of SEC24B deficiency in mice. Taken together, these observations suggest that the multiple *Sec24* paralogs have developed distinct functions over the course of vertebrate evolution.

INTRODUCTION

Approximately one-third of all vertebrate proteins traverse the intracellular secretory pathway prior to being secreted into the extracellular space or transported to any of a number of intracellular compartments, including the Golgi, endosome, lysosome, or plasma membrane [1,2,35]. Following co-translational translocation into the endoplasmic reticulum (ER) lumen, newly synthesized proteins are folded and undergo initial post-translational modification, followed by exit from the ER at ribosome-free regions called ER exit sites (ERES) [3] via COPII-coated vesicles [1,4,5]. In yeast, the COPII coat is composed of the small GTP-binding protein Sar1p, the heterodimeric Sec23p/Sec24p complex and the heterotetrameric Sec13p/Sec31p complex [72]. Sar1p generates membrane curvature and initiates vesicle formation by inserting an N-terminal amphipathic helix into the ER membrane [10]. The active membrane-bound Sar1p-GTP recruits Sec23p/Sec24p, and Sec24p drives the selective recruitment of cargo proteins into budding vesicles [43,46,73]. Polymerization of the outer Sec13p/Sec31p complex is the final step in vesicle budding [13].

While the components of the COPII coat are highly conserved, and the fundamental interactions appear to be similar from yeast to mammals, most components exhibit multiple paralogs in higher eukaryotes. Studies in yeast suggest that Sec24p is the major cargo binding component of the COPII coat, with three cargo-binding sites in the N-terminal region interacting with either cytoplasmic domains of the cargo itself or cargo adaptors [73]. Deletion of yeast Sec24p is lethal, whereas deletion of either of two non-essential Sec24p paralogs, Lst1p and Iss1p, results in specific cargo-transport defects [24,25,74]. In vertebrates, four *Sec24* paralogs (*SEC24A-D*) have been identified. These

Sec24 paralogs fall into two subfamilies, a SEC24A/B subgroup and a SEC24C/D

subgroup, based on protein sequence similarity, with the A/B subgroup closer to yeast

Sec24p [75]. All four paralogs contain highly conserved C-terminal domains and a more

variable N-terminal region, while the SEC24A/B and SEC24C/D subgroups appear to

differ in their affinity for a subset of known cargo-sorting signals [23,59].

Mice with mutations in *Sec24b* exhibit neural tube closure defects as a result of

decreased VANGL2 trafficking out of the ER [27], though no human disorders resulting

from deficiencies of SEC24B or any of the other *Sec24* paralogs have been reported.

Mutations in the SEC24 binding partner SEC23, which has two paralogs (*Sec23a* and

Sec23b), have been characterized both in humans and in fish. Missense mutations in

human *SEC23A* lead to cranio-lenticulo-sutural dysplasia (CLSD), characterized by the

persistence of wide-open fontanelles into childhood and the development of Y-shaped

cataracts [76]. Mutations in human *SEC23B* cause congenital dyserythropoietic anemia

type II (CDAII), characterized by a specific defect in erythrocyte development [77] while

Sec23b deficient mice have a markedly different phenotype, exhibiting pancreatic

disruption and disintegration [78]. Disruption of either *Sec23a or Sec23b* in zebrafish

both result in defects in extracellular matrix (ECM) protein secretion, producing a

phenotype reminiscent of CLSD in humans [32,76]. Zebrafish lacking SEC24D exhibit

similar craniofacial dysmorphology, presumably due to defects in the trafficking of

extracellular matrix (ECM) proteins including type II collagen and matrilin [31] and

medaka fish with a nonsense mutation in *sec24d* have also have skeletal defects [33]. We

now report the characterization of murine SEC24D deficiency. Mice null for SEC24D

exhibit very early embryonic lethality, suggesting an essential role for SEC24D in the transport of critical protein cargoes from the ER.

MATERIALS AND METHODS

Generation of SEC24D deficient mice

ES cell clones RRT226 and RRR785 were obtained from the International Gene Trap Consortium (IGTC, Bay Genomics, San Francisco, CA), and will be referred to as Sec24d^{gt} and Sec24d^{gt2}, respectively. Both ES cell clones were cultured as described [79] and expanded for microinjection and preparation of total RNA and genomic DNA. ES cell-mouse chimeras were prepared by blastocyst microinjection as described [80] and bred with C57BL/6J mice to obtain germ-line transmission. ES-cell derived F1 agouti offspring were genotyped using primers Neo A and Neo B to amplify a region of the neomycin cassette to determine the presence (Neo+) or absence (Neo-) of the gene trap allele. Sequences for all primers used in this study are listed in Table 2-1. Mice carrying the gene trap allele were maintained by backcrossing to C57BL/6J.

Mapping of the gene trap vector insertion sites

The gene trap vector insertion sites in intron 8 of Sec24d^{gt} and intron 20 of Sec24d^{gt2} were determined by PCR amplification and DNA sequencing. A series of forward primers evenly spaced throughout the intronic sequence (I8F1-21 and I20F1-7, Table 2-1) were combined with a reverse primer (Vector 19 or Vector 20, Table 2-1) specific to the 5'end of the gene trap vector sequence. Amplicons corresponding to a

specific product spanning the insertion site were confirmed by DNA sequencing. Insertion site sequences for both gene trap alleles were deposited into GenBank.

Genotyping mice by PCR and Southern Blot

Mice from Sec24d^{gt} were genotyped using a three-primer competitive PCR assay consisting of a common forward primer, (In8F3) located upstream of the insertion site in intron 8, and two reverse primers, located downstream of the insertion site in intron 8, (In8R4) or within the gene trap vector (V19) (Figure 2-1A). This reaction produces products of different sizes from the wild-type (762bp) and gene trap (666bp) alleles, which are resolved by agarose gel electrophoresis (Figure 1B). Genotypes for four representative Sec24d^{gt} mice were also confirmed by Southern blot analysis using a 371bp probe amplified from C57BL/6J genomic DNA with the primers ApaI A and ApaI B. The probe was hybridized to ApaI-digested genomic DNA, as previously described [81]. Mice from Sec24d^{gt2} were also genotyped using a three-primer competitive PCR, with a common forward primer (In20F1) located upstream of the insertion site in intron 20, and two reverse primers, located downstream of the insertion site (In20R1) in intron 20 or within the gene trap vector (V20) (Figure 2-2A). This reaction produces a 715bp product from the wild-type allele and a 558bp product from the gene trap allele, which are resolved by agarose gel electrophoresis (Figure 2-2B).

Timed mating

Timed matings were performed by intercrossing *Sec24d* heterozygous mice. Embryos were harvested at multiple time points, including day E10.5-11.5 for genotyping and histological analysis and E13.5 for the preparation of mouse embryonic fibroblasts. Genotyping was performed on genomic DNA isolated from embryonic yolk sacs. For

blastocyst collection, female $Sec24d^{+/gt}$ or $Sec24d^{+/gt2}$ mice were superovulated by intraperitoneal injection of 0.5 IU pregnant mares' serum gonadotropin (PMSG) on day 1 and 0.5 IU human chorionic gonadotropin HCG on day 3. Females were then mated with $Sec24d^{+/gt}$ or $Sec24d^{+/gt2}$ males, and copulation plugs were noted on day 4, counted as day E0.5 of embryonic development. Blastocysts were harvested on day 7 (E3.5) as previously described [82], and crude lysates were genotyped by three-primer PCR. Morula (8-cell embryos) were collected from juvenile superovulated $Sec24d^{+/gt}$ female mice mated with $Sec24d^{+/gt}$ male mice and placed in culture as described [83].

Reverse-transcription PCR

Total RNA was isolated from a panel of frozen tissues from wild-type mice, $Sec24d^{+/gt}$ mice and wild-type E10.5 embryos using the RNeasy Mini Kit (Qiagen), as per manufacture's instructions, including the optional DNaseI digestion step. cDNA synthesis and PCR were carried out in one reaction using the SuperScript® III One-Step RT-PCR System with Platinum®Taq (Invitrogen). Primers were designed such that amplicons for each gene were approximately the same size. Primer sequences are listed in Table 2-1.

Quantitative analysis by Southern blot

To determine the limit of sensitivity for our PCR assay, serial dilutions of total RNA from wild type into total RNA from $Sec24d^{gt2/gt2}$ embryos were used as template for RT-PCR. PCR products 190bp in length were amplified from the resulting cDNA using Sec24dExon20-21F and Sec24dExon20-21R (Table 2-1) and analyzed by Southern blotting using a 144bp ^{32}P-labelled DNA probe generated from wild-type cDNA using primers 24dEx20-21ProbeF and 24dEx20-21ProbeR (Table 2-1).

Generation of BAC Transgenic Mice

Two bacterial artificial chromosome (BAC) clones containing the entire *Sec24d* gene, RP23-355K12 (RP23) and RP24-271N12 (RP24) were obtained from the BACPAC Resources Center at Children's Hospital Oakland Research Institute (CHORI, http://bacpac.chori.org/). BAC DNA was purified using a NucleoBond® BAC 100 kit (Machery-Nagel), per manufacturer's instructions. C57BL/6JxSJL F1 female mice were generated by the University of Michigan Transgenic Animal Model Core and crossed to *Sec24d* heterozygous mice. Zygotes from this cross were injected with BAC DNA and transgenic founders for RP23 and RP24 were detected by PCR using pBACe3.6F1 and pBACe3.6R1 (Table 2-1), primers specific for the vector backbone.

Transgenic founders carrying the BAC transgene (*Tg+*) were generated using *Sec24d^{+/gt}* females as the egg donors, resulting in both *Sec24d^{+/+}* and *Sec24d^{+/gt}* founders for RP23 or RP24. *Sec24d^{+/gt} Tg+* founders were immediately crossed with mice heterozygous for the *Sec24d* gene trap allele (*Sec24d^{+/gt}*) to generate potential *Sec24d^{gt/gt} Tg+* rescues. To generate *Sec24d^{+/gt} Tg+* for lines with wild-type founders, an additional cross between *Sec24d^{+/+} Tg+* and mice heterozygous for the *Sec24d* gene trap allele (*Sec24d^{+/gt}*) was required. The resulting *Sec24d^{+/gt} Tg+* mice were crossed with *Sec24d^{+/gt}* mice to generate potential *Sec24d^{gt/gt} Tg+* mice. All progeny were subjected to genotyping for *Sec24d* as well as the presence of the BAC transgene. However, these assays cannot distinguish the endogenous wild-type allele from the copy of *Sec24d* present on the BAC-transgene. Thus, *Sec24d^{gt/gt} Tg+* mice were distinguished from *Sec24d^{+/gt} Tg+* mice by genotyping for microsatellites differing between the *Sec24d* "+"

and "gt" alleles (see below). All *Sec24d*$^{+/gt}$ mice used for this study had been
backcrossed to C57BL/6J ($\geq$N8).

Microsatellite Genotyping

A microsatellite genotyping assay was designed to distinguish the wild-type allele
from the gene trap allele originally targeted on the 129S1/SvImJ background. To ensure
that the correct genotype assignments were given, four independent microsatellites near
Sec24d but outside both BAC transgenes were chosen, two on either side of *Sec24d*
(Figure 2-3). These microsatellites, three tetra-nucleotide repeats and one tri-nucleotide
repeat, were selected for use in the microsatellite genotyping assay using the Tandem
Repeat Database [84] because they differed in allele size among the relevant mouse
strains to distinguish the endogenous *Sec24d* from the copy of the *Sec24d* gene in the
BAC-transgene. Each microsatellite was evaluated for every potentially transgenic
Sec24d$^{gt/gt}$ mouse by PCR on genomic DNA using GoTaq$^®$ Hot Start Green Master Mix
(Promega). A forward primer located upstream and a reverse primer located downstream
of the microsatellite repeat were used for amplification (see primer sequences, Table 2-1).

Primers were designed using Primer3 such that the amplicon size was
approximately 200bp in length, based on the C57BL/6J reference sequence.   PCR was
performed as per manufacturer's instructions, using 29 cycles and scaled up to a 30µl
reaction volume. Amplification annealing temperatures were optimized for each primer
set. PCR products were separated by PAGE using 20% polyacrylamide gels and
ethidium bromide staining. The gene trap allele is expected to be 129/SvImJ within the
congenic interval, in contrast to wild-type alleles, which should be either C57BL/6J,
DBA/2J, or SJL/J, based on the breeding strategy. (SJL/J was introduced with some of

the original transgenic founders, and DBA/2J with some early matings, though all subsequent backcrosses were to C57BL/6J).   Genomic DNA isolated from pure C57BL/6J, DBA/2J, 129S1/SvImJ, and SJL/J mouse strains was also used as templates to determine the amplicon size corresponding to each strain for each microsatellite marker. The genotypes of mice identified as $Sec24d^{gt/gt}$ $Tg+$ were confirmed by progeny testing through crosses with $Sec24d^{+/gt}$ mice ($\geq$ N7 on C57BL6/J). Data shown in the tables excludes 6 mice in which a recombination event occurred between the upstream and downstream sets of markers. This number is consistent with the predicted recombination frequency of ~1:50 within this 4.2Mb interval. We cannot exclude the possibility that we missed a double recombination event, though the chance of that occurring within our sample size is unlikely (predicted frequency for double recombinants, ~1:2500).

Ethics Statement

All animal care and use complied with the Principles of Laboratory and Animal Care established by the National Society for Medical Research. The University of Michigan's University Committee on Use and Care of Animals (UCUCA) approved all animal protocols in this study under protocol number 08571

Statistical Analysis

To determine statistical deviation from the expected Mendelian ratios of genotypes from a given cross, the *p*-value reported is the chi-squared value of observed ratio of genotypes compared to the expected ratio. Complete blood counts parameters were evaluated for significance using Student's T-test comparing levels from wild-type mice to levels from $Sec24d^{+/gt}$ mice. An initial analysis showed no significant difference between males and females for each genotype, therefore data from males and females

were pooled. The wild-type group consisted of 1 male and 2 females, and the $Sec24d^{+/gt}$ group consisted of 2 males and 2 females. Alpha levels were adjusted for multiple observations according to the Bonferonni correction.

RESULTS

SEC24D is required for early embryonic development in the mouse.

Genomic PCR and sequencing identified the Sec24d^{gt} gene trap insertion site at position 3378 of intron 8, numbering from the start of the intron (GenBank accession number KC763189) (Figures 2-1A,B, 2-4A-C). This insertion is consistent with the mRNA RT-PCR data mapping the gene trap to exon 8 [85] and is anticipated to disrupt SEC24D, generating a fusion transcript containing SEC24D exons 1 through 8 (encoding amino acids 1-347 of the total 1032 in SEC24D) fused to the β-geo selection cassette. Germline transmission of the $Sec24d^{gt}$ allele was achieved, as confirmed by PCR and Southern blot analysis (Figure 2-1B-D).

The genotypes of progeny mice generated from $Sec24d^{+/gt}$ intercrosses are shown in Table 2-2. Of 209 pups genotyped at weaning, no $Sec24d^{gt/gt}$ mice were observed ($p < 7 \times 10^{-17}$). Similarly, 0/28 and 0/27 $Sec24d^{gt/gt}$ genotypes were observed at E10.5-11.5 or the blastocyst stage, respectively. Only 1 out of 17 embryos genotyped at the 8-cell stage was $Sec24d^{gt/gt}$ ($p < 0.07$). Though a statistically significant excess of $Sec24d^{+/gt}$ mice was observed ($p < 0.001$) in intercross progeny at all of the above stages, analysis of a much larger number of backcross progeny (634) was consistent with the expected Mendelian ratio of 50:50 for wild-type and heterozygous mice (p>0.8).

No phenotypic abnormalities in *Sec24d^{+/gt}* mice

Sec24d^{+/gt} mice are viable and fertile and exhibit no gross or microscopic

abnormalities on standard autopsy examination. Complete blood count analysis identified

no significant differences between *Sec24d^{+/gt}* and *Sec24d^{+/+}* littermates for most

parameters measured after correction for multiple observations (Table 2-3). Electron

microscopy of pancreas and liver tissues from *Sec24d^{+/gt}* mice, as well as mouse

embryonic fibroblasts derived from *Sec24d^{+/gt}* mice, showed no abnormalities in the

cellular organization or ER structure compared to tissues or fibroblasts derived from

littermate *Sec24d^{+/+}* controls (Figure 2-5A-D).

***Sec24d* is ubiquitously expressed**

Analysis of *Sec24a-d* mRNA expression by real-time RT-PCR detected

expression of all four *Sec24* paralogs at E10.5, E14, and E18.5, consistent with data from

the EMAGE gene expression database (http://www.emouseatlas.org/emage/) [86]. RT-

PCR analysis of *Sec24d* expression in adult mouse tissues demonstrated broad expression

in a wide range of tissues, similar to previous reports of human expression patterns [75]

from the RNA Atlas [87]. The latter identifies SEC24D expression in all measured

human tissues (including colon, heart, hypothalamus, kidney, liver, lung, ovary, skeletal

muscle, spleen testes, and adipose tissue). Taken together, these data indicate that

Sec24d is expressed early and broadly across tissues.

***Sec24d* BAC transgenes rescue the embryonic lethal *Sec24^{gt/gt}* phenotype**

Two independent BAC transgenes, both spanning the full *Sec24d* gene (Figure 2-

3), exhibited rescue of the *Sec24d^{gt/gt}* embryonic lethal phenotype (Table 2-5). *Sec24d^{gt/gt}*

Tg+ mice appeared healthy, and exhibited normal fertility and lifespan with no apparent

abnormalities on gross autopsy. The *Sec24d^{gt/gt}* *Tg+* genotypes were confirmed by progeny testing (Table 2-5). While both BACs were able to rescue *Sec24d^{gt/gt}* mice, all of the transgenes generated less than the expected number of *Sec24d^{gt/gt}* *Tg+* rescues. A range of penetrance was observed, both between the RP23 and RP24 (average of 58% and 37.5%, respectively) and within founder lines of individual BACs (penetrance for RP23 transgene founders range from 40-75% and for RP24 founders, from 23-58%).

A hypomorphic *Sec24^{gt2}* allele supports development to mid-embryogenesis

A second *Sec24d* null mouse line (*Sec24d^{gt2}*) was generated from ES cells derived from an independent gene trap insertion and mapped to position 639 of intron 20 in Sec24d^{gt2} (GenBank accession number KC763190) (Figures 2-2, 2-6). The fusion transcript encoded by Sec24d^{gt2} contains SEC24D exons 1 through 20 (encoding the first 872 amino acids out of 1032) fused to the β-geo selection cassette. *Sec24d^{+/gt2}* mice, like the *Sec24d^{+/gt}* mice, are healthy and exhibit no apparent abnormalities upon standard autopsy examination. Intercrosses also yielded no *Sec24d^{gt2/gt2}* pups at weaning (Table 2-6), confirming the embryonic lethality of SEC24D deficiency. In contrast to the *Sec24d^{gt}* allele, analysis at both the blastocyst-stage and E10.5-11.5 identified *Sec24d^{gt2/gt2}* embryos in the expected Mendelian ratio, though no *Sec24d^{gt2/gt2}* embryos were observed beyond E13.5. The *Sec24d^{gt2/gt2}* E10.5 embryos appeared grossly and histologically indistinguishable from their wild type and heterozygous littermates, with visible heartbeats just prior to dissection. RT-PCR of total RNA prepared from *Sec24d^{gt2/gt2}* embryos at E10.5 detected a low level of residual normal splicing around the gene trap, though quantitative analysis by PCR and Southern blotting suggests that the level of this

residual full-length transcript in *Sec24d*^{gt2/gt2} mice is less than 0.1% of the wild type

allele.

DISCUSSION

Our data demonstrate that SEC24D is absolutely required for early embryonic

development in the mouse, with complete deficiency resulting in uniform loss prior to the

blastocyst stage. Low levels of SEC24D expression (< 0.1% of wild-type) are sufficient

to support development through mid-embryogenesis, though incompatible with survival

to term, whereas SEC24D haploinsufficiency results in no apparent phenotypic

abnormalities. The lack of a heterozygous phenotype, together with transgenic rescue of

the homozygous null phenotype, excludes a contribution from a dominant negative effect

of the truncated SEC24D fusion protein to the embryonic lethality. The transgenic rescue

also excludes a passenger gene effect at a nearby locus related to the gene targeting [88].

The transgenic rescue also demonstrates that the critical *cis*-regulatory sequences

required for SEC24D expression are contained within the ~140 Kb genomic segment

shared by the 2 BACs used in these experiments.

The reduction in null embryos as early as the 8-cell stage (Table 2-2) suggests that

residual maternal SEC24D is insufficient to maintain normal cellular function beyond the

first few cell divisions. These data suggest a role for an essential secretory cargo in the

early embryo that is specifically dependent on SEC24D for transport from the ER.

Alternatively, SEC24D could be the major or only *Sec24* paralog expressed at this early

developmental stage. In either case, the low level of normal *Sec24d* mRNA (<0.1%)

resulting from residual splicing around the hypomorphic *Sec24d*^{gt2} gene trap allele

appears to be sufficient to support development past this stage, though the corresponding

level of wild-type SEC24D protein was not directly measured. Also, a higher level of residual splicing in the early embryo, or partial function of the SEC24D β-gal fusion protein cannot be excluded.

The early essential role for SEC24D contrasts with the isolated neural tube developmental defect resulting from SEC24B deficiency in the mouse [27]. These results are surprising, in light of the observation of a higher extent of sequence identity for SEC24A and B to the essential yeast Sec24p protein than the SEC24C and D vertebrate paralogs, which are closer to the nonessential yeast genes Lst1p and Iss1p. The profound dependence of the early mammalian embryo on SEC24D was also unexpected, given the much milder phenotype observed in SEC24D deficient zebrafish. The latter animals exhibit craniofacial dysmorphology, thought to result from a specific defect in collagen secretion from chondrocytes [31], but otherwise develop normally. Variances in the level of maternal mRNA deposition between mice and zebrafish is not a likely explanation for these differences given the lengthy embryonic survival of the zebrafish compared to the mouse [31]. Rather, these observations suggest that the specific functions of the vertebrate SEC24s, mediated either through unique cargo selectivity or tissue-specific expression programs, may have shifted over evolutionary time. Consistent with this notion, the phenotypes of SEC23B deficiency differ markedly between humans, mice and zebrafish [32,77,78,89].

The initial F_1 and N_5 intercrosses of $Sec24^{gt/+}$ heterozygous mice revealed a puzzling excess of heterozygotes compared to wild-type offspring, significantly exceeding the expected 2:1 ratio (Table 2-2, $p < 0.0094$). This apparent selective advantage to the $Sec24^{gt/+}$ heterozygote was no longer evident after further backcrossing

into C57BL/6J, with genotyping of 634 backcross animals no longer showing an imbalance between the $Sec24^{gt/+}$ and $Sec24^{+/+}$ genotypes. Intercrosses of the second gene trap allele ($Sec24^{gt2}$) also failed to confirm an excess of heterozygous offspring (Table 2-6). Taken together, these data suggest the presence of an incidental "passenger" gene mutation at a locus tightly linked to the initial $Sec24^{gt}$ allele [88], which was eventually lost as a result of backcrossing to C57BL/6J.

The early embryonic lethality observed in SEC24D-deficient mice is consistent with the absence of a previously identified human SEC24D phenotype, though human deficiencies have also not yet been identified for any of the other $Sec24$ paralogs. However, the specific phenotypes of human mutations at $SEC23A$ and $SEC23B$ suggest the possibility of unique disorders associated with more subtle SEC24D mutations. Only 2 SEC23A-deficient pedigrees have been identified, each carrying unique missense mutations, likely associated with a partial change/loss of function. Similarly, though many different $SEC23B$ mutations have been identified in patients with CDAII, no patients have yet been identified who are homozygous or compound heterozygous for 2 null mutations, suggesting that complete SEC23B deficiency might also result in early lethality. The diverse phenotypes of humans, mice, and zebrafish with mutations in genes encoding components of the COPII coat suggests a complex balance of function among the multiple paralogous genes in each family. The availability of genetic models for deficiency in COPII component genes should enable future studies of COPII function and cargo selection *in vivo*.
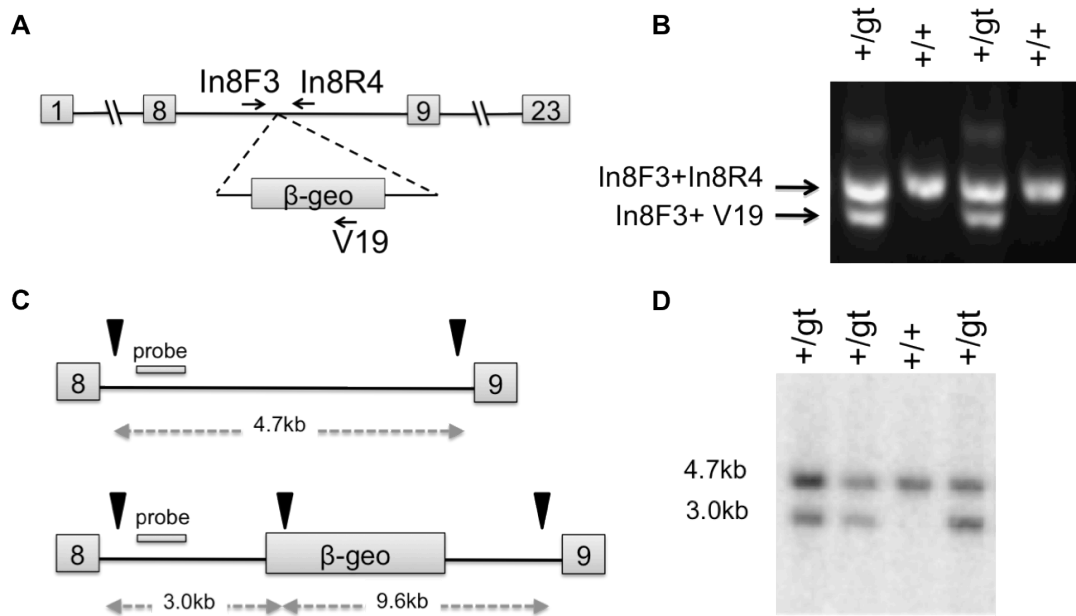
Figure 2-1: Generation of SEC24D-deficient mice

(A) Schematic representation of the *Sec24d*^{gt} gene trap allele in intron 8 at the mouse *Sec24d* locus. (B) Genotype results using a three-primer competitive PCR assay with the primers indicated in A. (C-D) Confirmation of the gene trap insertion by Southern blot analysis of ApaI-digested *Sec24d*^{gt} genomic DNA from *Sec24d*$^{+/+}$ and *Sec24d*$^{+/gt}$ mice; ApaI restriction enzyme sites (arrows) and probe location are depicted in C. Hybridization to the wild-type allele detected a 4.7kb ApaI restriction fragment, whereas hybridization to the gene trap allele detected a 3.0kb ApaI restriction fragment. In mice heterozygous for the *Sec24d*^{gt} gene trap allele, both restriction fragments are present.
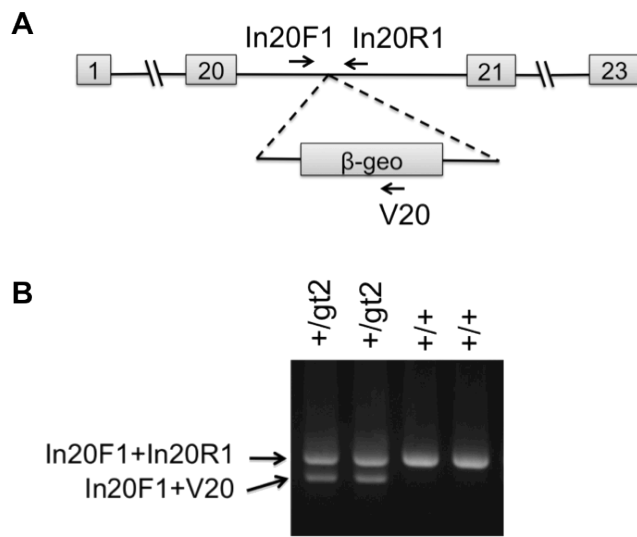
Figure 2-2: Generation of a second allele for SEC24D-deficient mice

(A) Schematic representation of the *Sec24d^{gt2}* gene trap allele in intron 20 at the mouse *Sec24d* locus.   (B) A three-primer competitive PCR genotyping assay to identify the *Sec24d^{gt2}* allele. Primer locations are depicted in (A).
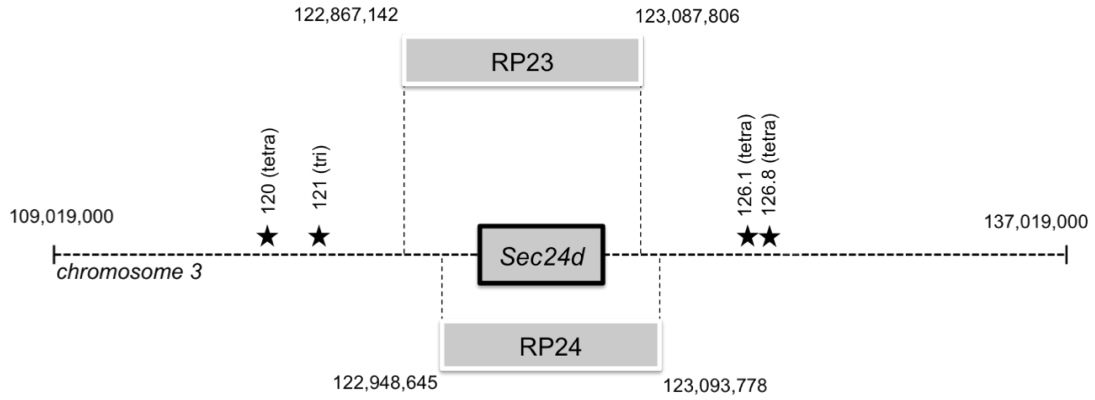
Figure 2-3: Locations of BACs and microsatellite markers on mouse chromosome 3

BAC transgenes RP23-355K12 (RP23) and RP24-271N12 (RP24) contain the entire *Sec24d* gene as well as upstream and downstream sequences as indicated. Stars depict relative locations of the four microsatellite markers used for microsatellite genotyping. Exact genomic locations are listed for genotyping primers are listed in Table 2-1.

A *Sec24d* **intron 8 sequence around the vector insertion site**

GAATGAACACATCCTAGGTATATGTAATTTCCACTTTCCTGAAACCCATTCGTTGCTCCTCCTCTTTTCTCTAGG
<u>In8F3</u> →

GACAAACATGGGTACACTTGATGTTTGTCCTCACCTTAGGACACTATGCAGCATTATTCTTCACGTAGCTGAGAT

GATGACATTTGAATTATAAAGTGTGAATGAGTGGTCACAGGGAAGGAGCCTCTTCGCTTCCCTTGCTGCCGTCCC

ACCACCGCTTTTCAGACAGCTCTTACCTATTTTATTTTTTTGAGATTCCTTTTACACATACAATATACATACTTC

TAAAATATACATATAAGTATGTGAATTCACCCCTTGACTAAGTCAGTATGTAAATTACTATATTAGTGCTGGCAC

AGCTGTGCCCTGCTCTCAGCATTTCAGTTTTTTTCATTAGTGTGTTGTGCCCTGGTCCCATGCTGCTGTTTGCAG

AGCTGCCTCACCCTTCACAGC**AGGGA**[▽]**GGCG**CACAGGTAGGAAGAGAAACATTGTTACCACCGTTAAGGCTTA

GGAAGAGAAACATCCTAGCCGCTGTGTGTGCCAGTTTGAGACTAGGGCTGATTCTAAAACCCGCTGAACTTTTGC

AGCTCGAGTCCTGCTGTTGAGCAGCTGTGGTCTCATAAATGAGCCCGGAAGAAGTGAATTATCTCCTCCTTGGGG

TGGGGAGATGAAGATGGACTTTTCCTGCCTTGCCTTTGATACCTATAACTTTCATAATCAAGAGAAGATGAGTTT

ATATCTTAAAAGCTACCCTAAGAATATGCAAAATAAGAAGT<u>TCTGTCGTGTTTCCTGGACA</u>TGGGATTTCGGCCC
← <u>In8R4</u>

CTGACAGTGTCCTTGTTGGGTATCCTGTGCCCACTGCAGTGAATTTGCCTGATTTTTAACAAGGCATGAGTAAA

B **5' end of pGT0Lxf vector sequence inserted into** *Sec24d* **intron 8**

AGGGAtgtgaaataccatgtgattgagaacgggtcaccatgtgcggtatttcacagcagggatgtgatttccgctagaagc

gGAAGAGGCCTTGATGGAAGGGCCCGCATGTCTCCAAAGTTGATTCATGCTTCTTGCACAGAGAAAGACCAGAAAGAAGGT

CTCAAGTTTTAGCCGGTAGCCCGGATGGCCTTTTCCTGCACGGCACCATATGAACCTTG<u>TGACCCTGACTTTGAGACCC</u>
← vector 19 primer

C **3' end of pGT0Lxf vector sequence inserted into** *Sec24d* **intron 8**

AACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGC

GATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGCCAAGCTT**GGCG**

Figure 2-4: *Sec24d*^gt^ allele sequence

(A) Sequence of the *Sec24d*^gt^ gene trap insertion site in intron 8 of the *Sec24d* gene. The arrowhead indicates the site of the gene trap vector pGT0LxF insertion, with the flanking *Sec24d* intron 8 sequence in bold. The locations of genotyping primer sequences are underlined. (B) 5' end of the vector sequence inserted into intron 8. Sequence aligning to intron 8 is in bold, while the lowercase sequence represents a 77bp insertion that is not present the pGT0LxF vector and cannot be identified in the mouse genome. The position of primer V19 within the 5' vector sequence is indicated. (C) 3' end of the vector sequence inserted into intron 8. Intron 8 sequence is in bold.
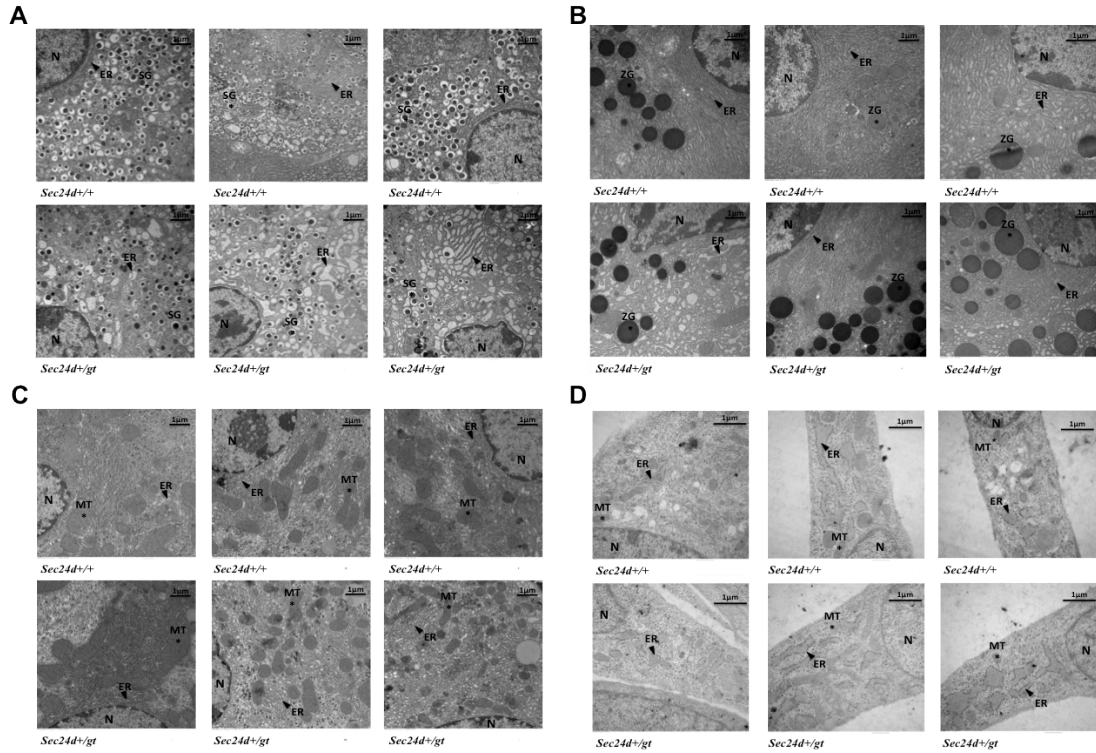
Figure 2-5: Transmission electron micrograph of *Sec24d*$^{+/+}$ and *Sec24d*$^{+/gt}$ cells

(A) pancreatic islet cells, (B) pancreatic acinar cells, (C) hepatocytes, and (D) mouse embryonic fibroblasts. Islet cells are identified by the presence of specialized secretory granules and acinar cells by the presence of zymogen granules. Samples were viewed at 10,500-13,500x direct magnification. A scale bar at the top right of each image corresponds to 1μm. Abbreviations: N = nucleus, ER = endoplasmic reticulum (black arrowheads), SG = secretory granules, ZG = zymogen granules, MT = mitochondria.

A *Sec24d* intron 20 sequence around the vector insertion site

GCATGCTCATGTGCTCAGCGT<u>GGCAGTGGAAGGTGTAAGGA</u>GTCCTGCTTCACTGGCTTGTTCAGGTCGCTTTC
In20F1 →

CCTTCTGCATTACTCTCAAGGTGGTAGTAACCCATTAAGTAGGCCCCTTTAGAAGTCTATAAACAAGTTGGTGC

CAGAAGACTTCAATCGAATTCCACATCTTAGAAGACCACTCACTTCTACATATGCTGGGGGAATCTTGTGTGGG

CGTGTGTGGTCTGTCTTCCTATAATGTGTGACTAAGAGTCACCATGTGCAGTCAGAGGCTAACAGATGAACTCG

TACTTGAGGGTACTCATGCCATGCAAGAGTCCTCAGTGTCTCCCATGGCAACACATTTCTCTCTGTTGAGGTGT

TGTTAACGTCTTGAAAACCCTGCACTCTGCCTGGTGGGATGTGTACAACTCCTCCTTGATCATAGTGCTAACAG

ATAGTAGGCACTTGTTGCTGGAAATATTAAAAAGAACAGCAAGTTCCCGCGCTACACCCAGCTTCTCTCTGCCC

CTGTCTCTAAGCCTCTGCCTC**TAAG[▽]CTAG**TTCCCAGCTCCAGTTCGCTTGTCTCAGTACTGCCTGTACCTT

CTCAACCAAAAACCCCCTGTGGTCACCAGTCCCACATTCCACTAACCCAGTAAAACAAAACTCTAGAAGCTTAT

AATTAATCAGTCAGATTTATATATCAATAAATTCTCAATATACAAGATGCC<u>CACACAATAAATTCAGGGCCAG</u>T
← In20R1

TGATAATGGTACATACAA

B 5' end of pGT0Lxf vector sequence inserted into *Sec24d* intron 20

TAAGGTCAGCCCCCCCCCCC<u>TATCCCATAGGAGCCAGGTC</u>CCTCTCCTGGACAGGAGGTGGTCCCAAGGTCTGG
← vector 20 primer
GGTAGAAGGTGAGAGGGACAGGCCACCAAGGAAGACTGAAGGGGAGATGCCAGAGACTCAGTGAAGCCTG

C 3' end of pGT0Lxf vector sequence inserted into *Sec24d* intron 20

GCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCA

AGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGCCAAGCTT**CTAG**

Figure 2-6: *Sec24d*^{gt2} allele sequence

(A) Sequence of the *Sec24d*^{gt2} gene trap insertion site in intron 20 of the *Sec24d* gene. The arrowhead indicates the site of the gene trap vector pGT0LxF insertion. Intron 20 sequences flanking the insertion site are in bold. The locations of genotyping primer sequences are underlined. (B) 5' and (C) 3' ends of the vector sequence inserted into intron 20. The position of primer V20 within the 5' vector sequence is indicated, and flanking intron 20 sequences are in bold.

Mapping and Genotyping Primers		Primers for BAC Transgene Rescue	
In8F1	CCATGCAGTGCTACACAAGC	pBACe3.6F1	GCTGCAGATCCCTAAACAGC
In8F2	CTGCTGCCTGAAGATCAAGA	pBACe3.6R1	TTCCGTCTCCGTCAAAAATC
In8F3	CATTCGTTGCTCCTCCTCTT	MS-1F	TGGTAGCAGGACACAGCTGATA
In8F4	AATCTGACCTGGGGAGAAGC	MS-1R	GGTCTAACACACGAGAATTTGAA
In8F5	CCCAACCCTCACGACAATAA	MS-2F	GCATGGAAAAACCCTGTCTC
In8F6	GCAGAGGCTGCTATTCCATC	MS-2R	CACCATTCAGCAATGATTCTC
In8F7	ATGGATGCTGCTGGAACTCT	MS-3F	TTCGGCTATTGTCTTCCACA
In8F8	CACAGGGAAAACGTGGAAAG	MS-3R	ACGGGGTTAGGTAGCCAGAT
In8F9	CGTGTCCTTCCCTAAACAGC	MS-4F	TGAGTCTGGCTAATTGCACTG
In8F10	CAGGTGGGGGATCTTATGAG	MS-4R	GATGGGAGGAGCATTCTGAG
In8F11	GGTGCTTTCAAATTGGTCAC		
In8F12	GATGGCGTGTAAGCTGTTGA	**Southern Blot Probe Primers**	
In8F13	GGGACAAAACAGCAGCCTAC	ApaI A	AATCCGTGGTTGTAGGTTGC
In8F14	CACTGGGGATATGGAACCTG	ApaI B	CAAAGGATCTCCCCACTCTG
In8F15	GGTGGGGAAGAGAACTTGTG	24dEx20-21ProbeF	ACAGTTTGTTGAAAAACTG
In8F16	CTGGCCTCTTTACACCCTTG	24dEx20-21ProbeR	ACGTGTGGATTGGCAGGAGCAG
In8F17	AAAGAGCGAGACCAACCTGA		
In8F18	TTTTCCTGTAGGCCCATGAC	**RT-PCR Primers**	
In8F19	AGCACAGGGAAGCCTAAGTG	GAPDH-F	TGTGATGGGTGTGAACCACGAGAA
In8F20	CCCTTTCCTCTTCCTCCACT	GAPDH-R	ACCAGTGGATGCAGGGATGATGTT
In8F21	GAGGTCAGAAGAGGGATCA	Sec24dEx20-21F	TGAAGGTGCTGCCTGTGTACATGA
In20F1	GGCAGTGGAAGGTGTAAGGA	Sec24dEx20-21R	ACATGTTCAGACCGTTAGCCAGCA
In20F2	GCCATGCAAGAGTCCCTCAGT		
In20F3	GCCCCTGTCTCTAAGCCTCT		
In20F4	CATCCTGTTCGTCCTCCATC		
In20F5	TGATCGGTTGCCACATAAAA		
In20F6	CCCTAGTCGGGCTCTTACCT		
In20F7	GGCCTTTCTCCCTCAAAAAG		
In8R4	TGTCCAGGAAACACGACAGA		
In20R1	CTGGCCCTGAATTTATTGTGTG		
Vector 19	GGGTCTCAAAGTCAGGGTCA		
Vector 20	GACCTGGCTCCTATGGGATA		
NeoA	CTTGCGCAGCTGTGCTCGACGTTG		
NeoB	TCTTCGTCCAGATCATCCTGATCG		

Table 2-1: List of primers used in Chapter 2

All primers are listed 5' to 3'.

Sec24d genotype:	+/+	+/gt	gt/gt	*p*-value
Sec24d$^{+/gt}$ ×*Sec24d*$^{+/gt}$ expected:	25%	50%	25%	
Progeny at weaning (n=209)	21.1% (44)	78.9% (165)	0	< 7.1 x 10^{-17}
E10.5 to E11.5 (n=28)	21.4% (6)	78.6% (22)	0	< 2.3 x 10^{-3}
Blastocyst (n=27)	29.6% (8)	70.4% (19)	0	< 2.7 x 10^{-3}
8-cell embryo (n=17)	11.8% (2)	82.4% (14)	5.8% (1)	< 6.0 x 10^{-2}
Sec24d$^{+/gt}$ × *Sec24d*$^{+/+}$ expected:	50%	50%	---	
N2 to N17 Progeny at weaning (n=634)	49.5% (314)	50.5% (320)	---	> 0.81

Table 2-3: Results of *Sec24d*$^{+/gt}$ intercrosses, and backcrosses to *Sec24d*$^{+/+}$ mice.

	$Sec24d^{+/+}$	$Sec24d^{+/gt}$	p-value*
WBC ($\times 10^3$ cells/μl)	5.50 ± 0.74	7.25 ± 0.63	> 0.12
RBC ($\times 10^6$ cells/μl)	9.57 ± 0.22	9.93 ± 0.13	> 0.19
HGB (g/dL)	15.67 ± 0.88	16.0 ± 0.0	> 0.67
HCT (%)	45.0 ± 1.0	46.25 ± 0.75	> 0.35
MCV (fL)	46.93 ± 0.35	47.35 ± 0.23	> 0.34
MCH (pg)	16.10 ± 0.74	16.15 ± 0.06	> 0.93
MCHC (g/dL)	34.37 ± 1.68	34.08 ± 0.13	>0.84
RDW (%)	12.07 ± 0.09	11.75 ± 0.09	> 0.04
PLT ($\times 10^3$ cells/μl)	950.0 ± 120.1	1035.0 ± 46.3	> 0.49
MPV (fL)	6.13 ± 0.52	6.50 ± 0.17	>0.48

Table 2-4: Complete blood count analysis of $Sec24d^{+/+}$ and $Sec24d^{+/gt}$ mice

Whole blood was drawn by retro-orbital puncture and analyzed on an Advia120 whole blood analyzer. All values are + or – standard error of the mean. *Based on the Bonferonni correction for multiple observations, the level of significance corresponding to $p < 0.05$ for a single observation would be adjusted to $p < 3.8 \times 10^{-3}$.

Genotype:	$Sec24d^{gt/gt}$ $Tg+$	other	p-value
Expected Ratio with Rescue	**14.3% (1/7)**	**85.7% (6/7)**	
Total for RP23 BAC (n=109)	8.3% (9)	91.7% (100)	$< 7.3 \times 10^{-02}$
RP23-677 (n= 53)	5.7% (3)	94.3% (50)	$< 7.3 \times 10^{-02}$
RP23-686 (n=56)	10.7% (6)	89.3% (50)	$< 4.5 \times 10^{-1}$
Total for RP24 BAC (n=243)	5.3% (13)	94.7% (230)	$< 6.9 \times 10^{-05}$
RP24-122 (n=61)	3.3% (2)	96.7% (59)	$< 1.5 \times 10^{-02}$
RP24-139 (n=85)	3.5% (3)	96.5% (82)	$< 4.6 \times 10^{-03}$
RP24-157 (n=97)	8.3% (8)	91.8% (89)	$< 9.0 \times 10^{-02}$
Total (n=352)	**6.25% (22)**	**93.75% (330)**	$< 1.7 \times 10^{-05}$

Table 2-5: Rescue of $Sec24d^{gt/gt}$ mice by BAC transgenes

Genotypes indicated are of 2-week old progeny resulting from a cross between $Sec24d^{+/gt}$ $Tg+$ mice and $Sec24d^{+/gt}$ mice.

Genotype:	$Sec24d^{+/+}$ $Sec24d^{+/+}$ Tg+	other	
Expected Ratio if parent was Sec24d$^{+/gt}$, BAC-Tg+	**28.6% (2/7)**	**71.4% (5/7)**	***p*-value**
Total for RP23 BAC (n=47)	0%	100% (47)	< 1.5 x 10^{-05}
RP23-677 (n= 53)	0%	100% (47)	< 1.5 x 10^{-05}
Total for RP24 BAC (n=83)	0%	100% (83)	< 8.4 x 10^{-09}
RP24-139 (n=53)	0%	100% (53)	< 4.2 x 10^{-06}
RP24-157 (n=30)	0%	100% (30)	< 5.4 x 10^{-04}
Total (n=130)	**0%**	**100% (130)**	< 5.6 x 10^{-13}

Table 2-6: Progeny testing of *Sec24d$^{gt/gt}$ Tg+* mice

Tested mice were crossed with *Sec24d$^{+/gt}$* mice, and progeny were genotyped. The presence of any *Sec24d$^{+/+}$* mice would indicate that the test parent was heterozygous for the gene trap allele. P-values are calculated for the observed genotypes compared to the expected if the tested transgenic mouse was *Sec24d$^{+/gt}$* rather than *Sec24d$^{gt/gt}$*.

Sec24d genotype:	+/+	+/gt2	gt2/gt2	p-value
Sec24d$^{+/gt2}$ × Sec24d$^{+/gt2}$ expected:	25%	50%	25%	
Progeny at weaning (n=88)	34.1% (30)	65.9% (58)	0	$< 6.1 \times 10^{-08}$
E13.5 to E18.5 (n=29)	37.9% (11)	62.1% (18)	0	$< 1.9 \times 10^{-03}$
E10.5 to E11.5 (n=105)	24.8% (26)	56.2% (59)	19% (20)	$< 1.6 \times 10^{-01}$
Blastocyst (n=99)	26.3% (26)	47.4% (47)	26.3% (26)	$< 7.8 \times 10^{-01}$
Sec24d$^{+/gt2}$ × Sec24d$^{+/+}$ expected:	50%	50%	---	
Backcrosses at N2 to N12 (n=367)	56% (207)	44% (160)	---	$> 1.0 \times 10^{-02}$

Table 2-7: Results of Sec24d$^{+/gt2}$ intercrosses, and backcrosses to Sec24d$^{+/+}$ mice

Notes

This chapter was published in the journal PLoS One in 2013 under the title "Disruption of the *Sec24d* gene results in early embryonic lethality in the mouse" by Andrea C. Baines*, Elizabeth J. Adams*, Bin Zhang, and David Ginsburg.

*Authors contributed equally to this work.

CHAPTER III: MAMMALIAN COPII COMPONENT SEC24C IS REQUIRED FOR EMBRYONIC DEVELOPMENT IN MICE

ABSTRACT

COPII coated vesicles mediate the transport of newly synthesized proteins from the endoplasmic reticulum (ER) to the Golgi. SEC24 is the COPII component primarily responsible for recruitment of protein cargoes into nascent vesicles. There are four *Sec24* paralogs in mammals, with mice deficient in SEC24A, B, and D exhibiting a wide range of phenotypes. We now report characterization of mice with deficiency in the fourth *Sec24* paralog, *Sec24c*. Although mice haploinsufficient for *Sec24c* exhibit no apparent abnormality, homozygous deficiency results in embryonic lethality around embryonic day 7. Tissue-specific knockouts of *Sec24c* in hepatocytes, pancreatic acinar cells, smooth muscle cells, and intestinal epithelium are phenotypically normal. Thus, SEC24C is required in early mammalian development, but is dispensable in a number of tissues, likely as a result of compensation by other *Sec24* paralogs. The embryonic lethality resulting from loss of SEC24C occurs considerably later than the lethality previously observed in SEC24D deficiency, and is quite distinct from the restricted neural tube phenotype of *Sec24b* null embryos, and the mild hypocholesterolemic phenotype of adult *Sec24a* null mice. Taken together, these results demonstrate that the four *Sec24* paralogs have developed unique functions over the course of vertebrate evolution.

INTRODUCTION

Approximately one-third of all mammalian proteins traverse the secretory pathway en route to their final destinations, including localization to cell surface membranes, the lysosomes and other intracellular compartments, or secretion into the extracellular space [1,2]. These proteins begin their journey in the endoplasmic reticulum (ER), where they are recruited into newly forming COPII vesicles located at ribosome-free ER exit sites [1,3,4,5]. The yeast COPII coat is composed of the small GTP-binding protein Sar1p, a heterodimeric complex of Sec23p/Sec24p, and a heterotetrameric of Sec13p/Sec31p [72]. The COPII vesicles containing these cargoes are then trafficked from the ER to the ER-Golgi intermediate compartment (ERGIC), where they are further modified and directed to their final destination.

The mechanism of COPII coat assembly has been elegantly dissected in *S. cerevesiae* and the fundamental mechanisms appear to be conserved from yeast to humans [90]. Vesicle formation is initiated with the activation of the small GTPase Sar1p by the GEF Sec12p, which is localized on the ER membrane. GTP binding to Sar1p induces a conformation change and insertion of a small amphipathic helix of Sar1p into the ER membrane and begins the process of vesicle formation [9,10]. Activated Sar1p recruits the Sec23p-Sec24p heterodimer to the ER forming the inner layer of the COPII coat [11,12]. Polymerization of Sec13p and Sec31p heterotetramers to form the outer layer, promoting further curvature and budding of the nascent COPII vesicle from the ER [13].

SEC24 is the COPII component thought to be primarily responsible for cargo recruitment, with ER exit motifs on protein cargos interacting with specific binding sites

on SEC24 [46,73]. SEC24-cargo interactions are either direct, in the case of transmembrane proteins, or require a transmembrane adaptor to mediate interaction of ER luminal cargo with the COPII coat located on the cytoplasmic face of the ER membrane. Examples include the well-characterized interaction between the adaptor component LMAN1/MCFD2 and its cargos Factor V and Factor VIII [39,91,92].

In *S. cerevisiae*, loss of Sec24p is lethal, while deficiency of either of two non-essential Sec24p paralogs, Iss1p or Lst1p, results in specific cargo-transport defects [24,25,74]. The mammalian genome encodes four paralogs of *Sec24* (*Sec24a-d*), which can be classified into two subgroups, SEC24A/B and SEC24C/D, based on protein sequence identity, with the A/B subgroup more closely related to the ancestral yeast paralog Sec24p [75,93]. All four paralogs contain a highly conserved C-terminus and a variable N-terminal region, with previous reports suggesting differences between the two SEC24 subgroups in their affinity for cargo-sorting signals [23,59]. All four *Sec24* paralogs appears to be ubiquitously expressed in the adult and in the developing embryo mouse [28,86].

Mutations in several *Sec24* paralogs in Zebrafish suggest a critical role in proper the secretion of extracellular matrix collagens [29,30,31,32]. While there have been no diseases attributed to mutations in any of the human *SEC24* paralogs, mutations in *SEC23A* result in the human disorder cranio-lenticulo-sutral dysplasia [76] and mutations in human *SEC23B* cause congenital dyserythropoietic anemia type II (CDAII) [77]. In contrast to humans, *Sec23b* deficient mice exhibit perinatal lethality due to the destruction of the pancreas [78]. Mice deficient in SEC24A exhibit low plasma cholesterol levels, attributed to a block in the secretion of the regulatory protein PCSK9

[26]. SEC24B-deficient mice have a specific defect in the secretion of VANGL2 leading to a neural tube closure defect[27]. SEC24D-deficient mice die very early in embryonic development (~E3.5)[28]. We now report that the loss of murine SEC24C also results in embryonic lethality, though at a later time point (~E7.5). However, ablation of SEC24C in multiple tissues is surprisingly well tolerated.

MATERIALS AND METHODS

Generation of SEC24C deficient mice

ES cell clone EPD0241-2-A11 for $Sec24c^{tm1a(EUCOMM)Wtsi}$, an allele for $Sec24c$ with conditional potential, was obtained from the European Mouse Mutagenesis program (EUCOMM) and will be referred to as $Sec24c^{GT}$. ES cells were cultured and expanded for microinjection and genomic DNA isolation as previously described [80]. ES cell-mouse chimeras were generated as described [81] and subsequently bred to B6(Cg)-Tyr^{c-2J}/J (JAX stock #000058) to achieve germ-line transmission. ES-cell derived F1 black progeny were genotyped using primers to detect the presence ($Sec24c^{GT}$) or absence ($Sec24c^{wt}$) of the targeted allele (primers Ex and E, Figure 3-2, Table 3-7). The $Sec24c^{GT}$ allele was maintained by continuous backcrosses to C57BL/6J mice. Genotyping was performed with mouse tail clip DNA using Go-Taq Green Master Mix (Promega, Madison, WI), and the resulting PCR products resolved by 2% agarose gel electrophoresis.

Long-range PCR to confirm insertion site of *Sec24c* gene trap allele

Genomic DNA isolated from the ES cell clone EPD0241-2-A11 and genomic DNA from a wild type C57BL/6J mouse tail clip were amplified by long-range PCR to

confirm correct targeting of the $Sec24c^{GT}$ allele with primers 24c-GF3 and primer B (5'

end) and 24c-GR4 and RAF5 Fwd using Phusion Hot Start II DNA Polymerase (Thermo

Scientific), with products resolved on a 0.8% agarose gel.

Generation of a *Sec24c* conditional allele

To generate mice carrying the conditional $Sec24c^{tm1c(EUCOMM)Wtsi}$ allele, referred to

from now on as $Sec24c^{FL}$ allele, $Sec24c^{+/GT}$ were crossed to mice transgenic for Flpe

recombinase driven by an actin promoter (C57BL/6J background, JAX stock no.

005703). Mice were genotyped with primers Ex and E to detect both the wild type

(308bp) allele and the presence of the LoxP site (278bp), primers A,B and C to

distinguish between the $Sec24c^{GT}$ (204bp) and $Sec24c^{FL}$ (534bp) alleles, and with primers

FLP1 and FLP2 to detect the FLPE transgene (750bp). $Sec24c^{+/FL}$ mice were backcrossed

to C57BL/6J mice to remove the FLPe transgene.

Generation of a *Sec24c* null allele

$Sec24c^{+/FL}$ were crossed to EIIA driven Cre recombinase transgenic mice

(C57/BL6 background, JAX stock no. 003724), and offspring were genotyped with

primers A, Ex, and E to distinguish $Sec24c^{+}$, $Sec24c^{FL}$, and the $Sec24c^{tm1d(EUCOMM)Wtsi}$

allele, referred to as the $Sec24c^{-}$ allele (Figure 3-2). The Cre transgene was detected with

primers Cre fwd and Cre rev. $Sec24c^{+/-}$ mice were backcrossed to C57BL/6J mice to

remove the Cre transgene.

Generation of tissue-specific knockout mice

$Sec24c^{FL/FL}$ mice were crossed to $Sec24c^{+/-}$ $Cre+$ using the following tissue-

specific transgenes: P48-Cre (pancreas-specific,[94]), Villin-Cre (intestinal-epithelial-

specific, JAX #004586) [95], Albumin-Cre (hepatocyte-specific, JAX #003574) [96] SM22-Cre mice (smooth-muscle-specific, JAX #004746) [97], and Meox2-Cre transgene (ubiquitous expression beginning at embryonic day 5, JAX #003755) [98]. Progeny were genotyped at the *Sec24c* locus with primers A, Ex, and E and with primers AnyCreF and AnyCreR to detect the presence of the appropriate Cre transgene. The level of Cre-mediate excision was assessed by isolation of genomic DNA from tissue samples of either $Sec24c^{+/FL}$ $Cre+$ or $Sec24c^{FL/-}$ $Cre+$ mice and PCR using primers A, Ex, and E (Table 3-7) at the *Sec24c* locus.

Phenotypic analysis of $Sec24c^{+/-}$ and $Sec24c^{+/-}Sec24d^{+/GT}$, and tissue-specific knockout mice

Whole blood was collected and complete blood counts were carried out as previously described [28]. For histology, tissues were fixed in 4% paraformaldehyde in PBS overnight at 4°C, then transferred to 30% EtOH, 50% EtOH, and 70% EtOH, each for thee times ten minutes. Processing, embedding, sectioning and H&E staining were performed at the University of Michigan's Microscopy and Image Analysis Laboratory. Body weights were measured weekly from weaning up to 12 weeks of age. A high fat diet (45% of calories from fat) was purchased from Research Diets (New Brunswick, NJ) and fed ad libitum. For insulin and blood lipid analysis, mice were fasted 16 hours before blood collection. Blood was collected using heparin-coated collection tubes (Fisher, Pittsburgh, PA) by retro-oribtal bleeding from mice anaesthetized with isoflurane. Plasma samples were then collected by centrifugation of heparinized blood samples at 3,000 g for 5 min at 4°C. Plasma cholesterol and triglyceride levels were measured with 5ul of plasma samples colorimetric assays using the LiquiColor Cholesterol test kit (Stanbio,

Boerne, TX), or Serum Triglyceride Determination Kit (Sigma-Aldrich, St. Louis, MO)

according to the manufacturer's instructions. Insulin levels were measured using the Ultra

Sensitive Mouse Insulin ELISA Kit (Chrystal Chem), per manufacturers instructions.

Timed Matings

Timed matings were carried out by intercrossing *Sec24c* heterozygous mice.

Embryos were harvested at multiple time points, including day E7.5–11.5 for genotyping

and histological analysis. Genotyping was performed on genomic DNA isolated from

embryonic yolk sacs. Blastocyst collection and genotyping was performed as previously

described [82] using super-ovulated *Sec24c*$^{+/-}$ females and *Sec24c*$^{+/-}$ males.

Construction of *Sec24c* and *Sec24d* transgenes and generation of transgenic mice

Transgenes carrying either *Sec24c* or *Sec24d* were designed using the previously

reported pCAG3Z vector [99]. The entire transgenic construct is shown in Figure 6A.

The CAG promoter contains the CMV intermediate-early enhancer and the chicken β-

actin promoter to drive early, ubiquitous expression [100,101] of either *Sec24c* or

Sec24d.   AgeI and NotI sites were added upstream of an SV40 early termination signal

and a poly A signal isolated from Tg2.33 [99] and this was placed downstream of the

CAG promoter in the pCAG3Z vector to create pCAG3zS (Figure 6A).   *Sec24c* (cDNA -

9 to 3290) or *Sec24d* (cDNA -9 to 3098) was inserted between the CAG promoter and the

SV40 signal using AgeI and NotI. DNA sequencing was performed on all constructs

prior to microinjection to verify the integrity of the transgenes. All primers used to create

these constructs are listed in Table 3-7.   pCAG3zS-Sec24c and pCAGzS-Sec24d were

liberated from the vector backbone with a double digest with BamHI-HF/SphI-HF or

SacI/HindIII, respectively (New England Biolabs), and transgenic mice generated by the

University of Michigan Transgenic Animal Model Core as previously described [102]. Transgenic founders (C57BL/6J X SJL F2) for both lines were detected by PCR using Primer-7F and Primer-7R located in the promoter region of the transgenes. Mice transgenic for either pCAG3zS-Sec24c or pCAG3zS-Sec24d were crossed with $Sec24c^{+/-}$ mice to generate $Sec24c^{+/-}$ Tg+ mice, which were then crossed to $Sec24c^{+/-}$ mice to generate potential $Sec24c^{-/-}$ Tg+ mice. All progeny were genotyped at the $Sec24c$ locus and for the presence of the transgene. Similar crosses were also performed with $Sec24d^{+/GT}$ [28] mice. Five founders for pCAG3zS-Sec24c were used to generate transgenic lines to test the ability of pCAG3zS-Sec24c to rescue the loss of $Sec24c$. Four pCAG3zS-Sec24d founders were used to generate transgenic lines.

Reverse-transcription PCR

Total RNA was isolated from a panel of frozen tissues from transgenic mice from each of the founder lines using RNAeasy kit (Qiagen) as per manufacture's instructions. cDNA synthesis and PCR were carried out in one reaction using SuperScript® III One-Step RT-PCR System with Platinum®*Taq* (Invitrogen) per manufacturer's instructions. Primers V, W, X, Y located in *Sec24c* exons 4 through 7 were used to amplify signal from *Sec24c* cDNA (Figure 3-1). Primers 24c-F or 24d-F and primers 24c-R or 24d-R were used to detect cDNA specific for the *Sec24c* or *Sec24d* transgene (Figure 3-6). GAPDH amplification with primers GAPDH-QFwd3 and GAPDH-QRev3 was carried out in parallel for each sample. Primer sequences are listed in Table 3-7.

Statistical Analysis

P-value for progeny genotypes were calculated by a chi-squared test comparing expected ratio of genotypes compared to those observed. Complete blood counts

parameters were evaluated for significance using Student's T-test. Alpha levels were adjusted for multiple observations according to the Bonferonni correction.

RESULTS

$Sec24c$ is ubiquitously expressed, although two alternative splice forms are tissue-specific

$Sec24c$ mRNA expression was detected by RT-PCR in all 15 adult mouse tissues tested, consistent with previously reported human expression patterns [75,87]. RT-PCR analysis also detected an alternative, in-frame exon (exon 6*, encoding 23 amino acids), present in a subset of $Sec24c$ mRNAs (Figure 3-1A, B). The transcript containing this additional exon ($Sec24c-2$) is expressed only in the heart, brown adipose tissue, skeletal muscle, and the brain (Figure 3-1C). The transcript skipping exon 6* ($Sec24c-1$) is the only splice form detected in the remaining tissues (adrenal gland, liver, lung, kidney, spleen, stomach, large intestine, small intestine, white adipose tissue, salivary gland, testis), with only low levels of $Sec24c-1$ detected in $Sec24c$-2 expressing tissues (Figure 3-1C).

$Sec24c^{GT/GT}$ and $Sec24c^{-/-}$ mice exhibit lethality

Correct targeting of the $Sec24c^{GT}$ allele (Figure 3-2A) was confirmed by long-range PCR (Figure 3-2B). Genotypes of 2-week-old progeny mice generated from $Sec24c^{+/GT}$ intercrosses revealed the expected number of wild type and heterozygous offspring but none of the expected ¼ $Sec24c^{GT/GT}$ ($p<2X10^{-06}$, Table 3-1). Intercrosses of $Sec24c^{+/FL}$ (obtained by excision of the β-geo cassette from $Sec24c^{GT}$) generated the

expected number of $Sec24c^{FL/FL}$ mice (Table 3-2) demonstrating that death of the

$Sec24c^{GT/GT}$ mice results from the presence of the gene trap cassette, rather than a

passenger gene effect [88]. EIIA-Cre mediated excision of the $Sec24c^{FL}$ generated the

$Sec24c^{-}$ allele, which lacks exon 3 and results in a frame-shift and early termination

codon. Intercrosses of $Sec24c^{+/-}$ mice confirm uniform loss of $Sec24c^{-/-}$ mice by 2 weeks

of age (Table 3-3). At embryonic day 10.5-11.5, no intact $Sec24c^{-/-}$ embryos were seen,

though their yolk sacs were still present and were able to be separated from maternal

tissue and genotyped and found to be present at the expected ratios ($p > 0.25$). Genotypes

were assessed at E9.5, yet still only the yolk sacs of null embryos were found ($p < 0.05$),

while at E8.5, the $Sec24c^{-/-}$ embryo was noted to have been dead for approximately 12

hours (blinded observation prior to genotyping). Histological observations of E7.5

embryos revealed that they fell into 3 groups: those that are just remnants of the egg

cylinder (#1, Figure 3-3A), those that are not yet gastrulating (#2, Figure 3-3A) and those

that are gastrulating and developing normally (#3, Figure 3-3A), though no genotypes

were assigned to any of these embryos. Taken together, these data suggest that SEC24C

deficient embryos are dying between E7.5 and E8.5, though the cause of death is

unknown. Analysis of progeny from $Sec24c^{+/-}$ intercrosses indicates that heterozygous

mice are present in the expected ratios ($p > 0.09$) (Table 3-3).

No phenotypic abnormalities in $Sec24c^{+/-}$ mice

As noted above, the expected numbers of $Sec24c^{+/GT}$ and $Sec24c^{+/-}$ mice are

observed at 2 weeks of age. Growth and complete blood counts of heterozygous mice

were also indistinguishable from their wild type littermates (Figure 3-4, Table 3-4). Adult

$Sec24c^{+/-}$ mice are fertile and exhibit no gross or microscopic abnormalities on standard

autopsy examination (Figure 3-4C). While only a small number of wild type littermates were followed beyond 100 days, no significant difference (*p>0.05*) was observed in the lifespan of *Sec24c*$^{+/-}$ mice compared to controls.

Combined haploinsufficiency for SEC24C and SEC24D is also well tolerated

Given the high level of sequence similarity between SEC24C and SEC24D [75], we performed intercrosses between *Sec24c*$^{+/-}$ and *Sec24d*$^{+/GT}$ mice [28] (Table 3-5). At two weeks of age, double heterozygous mice are present at the expected Mendelian ratio (p>0.12). No differences were observed in growth to 12 weeks (Figure 3-5A,B) or complete blood count analyses (Table 3-4), when compared to WT or single heterozygous littermates. While only a small number of wild type littermates were followed beyond 100 days, no significant difference (*p>0.05*) was observed in the lifespan of *Sec24c*$^{+/-}$*Sec24d*$^{+/gt}$ mice compared to controls. Routine autopsy and histological survey were also unremarkable (Figure 3-5C).

The embryonic lethality of *Sec24c*$^{-/-}$ mice is not rescued by an early, ubiquitous *Sec24c* or *Sec24d* transgene

Transgenes designed to express SEC24C or SEC24D from the ubiquitous chick β-actin promoter (Figure 3-6A) were generated and tested for their ability to rescue the lethal *Sec24c*$^{-/-}$ phenotype. No *Sec24c*$^{-/-}$ *Tg+* mice survived to weaning for either the SEC24C (1 line, n=34) or SEC24D (1 line, n=49) transgene.   Similarly, crosses to the *Sec24d*^{GT} mice failed to demonstrate rescue of the *Sec24c*$^{GT/GT}$ lethal phenotype by either transgene; for the SEC24D transgene, no rescues were observed out of 149 transgenic mice from 5 different founder lines, with n for each line ranging from 9 to 43 transgenic

progeny. For the SEC24C transgene, no rescues were observed out of 130 transgenic mice from 4 founder lines, with n of each line ranging from 11 to 57 transgenic progeny.

SEC24C is dispensable in pancreatic acinar cells, hepatocytes, smooth muscle cells, and intestinal epithelium

To test the requirement for SEC24C in specific adult tissues, $Sec24c^{FL/FL}$ mice were crossed with mice carrying Cre transgenes specific for pancreatic acinar cells, hepatocytes, smooth muscle cells, and intestinal epithelial cells (Table 3-6). $Sec24c^{FL/-}$ $Cre+$ mice were observed in the expected numbers for p48-Cre ($p>0.65$), SM22-Cre ($p>0.46$), Villin-Cre ($p>0.46$), and Albumin-Cre ($p>0.10$) and SM22-Cre ($p>0.46$). Though Cre mediated excision of the $Sec24c^{FL}$ allele is nearly complete in the pancreas of p48Cre+ mice (Figure 3-7A), pancreatic acinar and islets appear entirely normal by routine histology (Figure 3-7B). $Sec24c^{FL/-}$p48-Cre+ also exhibit normal weight gain though 12 weeks of age compared to their littermate controls on either normal chow (Figure 3-8A,B) or high fat diet (Figure 3-9A,B). Similar results were observed for the Albumin-Cre and Villin-Cre transgenes, though the levels of excision are not as complete for p48-Cre (Figure 3-7).   There were also no differences in plasma cholesterol or triglyceride levels in tissue-specific knockouts versus littermate controls on HFD for p48-Cre, Albumin-Cre, and Villin-Cre transgenes (Figures 3-9, 3-10, 3-11) and insulin levels between $Sec24c^{FL/-}$p48-Cre+ and corresponding littermate controls on HFD were also the same (Figure 3-9E). $Sec24c^{FL/-}$ SM22-Cre+ mice did not have any grossly apparent abnormalities were able to carry litters to term, and while excision of $Sec24c$ in smooth muscle cells in $SM22$-Cre+ mice were not measured directly, it is expected to be

complete, based on previous reports [103]. *Sec24c$^{FL/-}$ Meox2-Cre+* mice were not observed at 2-weeks of age (*p<1.7x10^{-6},* Table 3-6).

DISCUSSION

Our results demonstrate that SEC24C is required for post-implantation embryonic development in the mouse, with SEC24C deficient embryos being lost in the developmental time points between implantation and E7.5. The loss of *Sec24c$^{FL/-}$ Meox2-Cre+* mice is consistent with these data, as Meox2-Cre is expressed ubiquitously beginning around embryonic day 5 [98]. Data from tissue-specific knockouts reveal that SEC24C is dispensable in a number cell types, including pancreatic cells, hepatocytes, intestinal epithelium, and smooth muscle cells, suggesting that the three remaining *Sec24* paralogs can compensate for the loss of SEC24C in these tissues. These observations, along with the requirement for SEC23B in the mouse pancreas [78], the need for SEC24A in the liver to maintain normal plasma cholesterol levels [26], and SAR1B's involvement in the efficient trafficking of chylomicrons in the gut [104], suggest that there is likely a partial overlap in function between the COPII paralogs, but that some cargoes require a specific COPII protein for exit out of the ER.

The normal phenotype of the *Sec24c$^{+/-}$* mice, as well that of the *Sec24c$^{+/-}$ Sec24d$^{+/GT}$* mice, demonstrates a tolerance for moderate quantitative changes in the levels of these two COPII components, consistent with previous reports of *Sec24d$^{+/GT}$* mice [28] as well as *Sec24a$^{+/-}$Sec24d$^{+/GT}$* and *Sec24a$^{+/-}$Sec24b$^{+/-}$* mice exhibiting no abnormalities [26]. The milder phenotype observed for deficiencies in SEC24A/B compared to SEC24C/D is surprising, in light of a higher of sequence identity between the ancestral yeast paralog Sec24p and the SEC24A/B subfamily [75]. In contrast, the SEC24C/D

subfamily exhibits greater similarity to the non-essential yeast homologue Lst1p. Yet, as

in the case of SEC24D-defecient mice [28], the loss of *Sec24c* null embryos at early

developmental stage suggests that there are critical cargoes that require SEC24C for exit

from the ER that are no longer being secreted at sufficient levels to support embryonic

survival. Interestingly, SEC24A and SEC24B demonstrate some partial overlap in

function, where in the absence of SEC24A, SEC24B, but not SEC24C or D, is able to

recruit PSCK9 into COPII vesicles at low levels [26], likely due to the high level of

sequence similarity within the SEC24A/B subfamily. The later survival and specific

phenotypes of *Sec24a* and *Sec24b* mice allowed the identification of specific cargoes

(PCSK9 and VANGL2, respectively) whose trafficking defects explain the mouse

phenotypes [26,27]. Given the much earlier loss of mice deficient in either SEC24C or

SEC24D, identification of precise cause of death and the specific cargoes responsible is

far more difficult. Much like SEC24D, SEC24C is essential for the recruitment of

cargoes at the early stages of development, but our data demonstrate that some essential

cargos reliant on SEC24D must not require SEC24C, as the SEC24C-deficient embryos

are able to survive well beyond the time point at which *Sec24d* null embryos are lost.

The frequency of damaging *SEC24C* alleles in the African American and

European American population is ~0.04%. Given this, there are several explanations for

the absence of any previously identified human SEC24C phenotypes. Humans deficient

for SEC24C may also exhibit embryonic lethality, consistent with the phenotype found in

mice, in which case they would never be observed. SEC24C-deficiency in humans can

also be the cause of a human disorder for which no genetic cause has been found.

Alternatively, the humans without SEC24C may have a subclinical or no phenotype,

which would suggest differential functions of SEC24C between mouse and humans. Though the phenotype of zebrafish lacking SEC24C has yet to be reported, this phenomenon of varied phenotypes between species is also observed in the stark differences in phenotypes between SEC24D deficient mice and the *bulldog* mutant in zebrafish [28,31] and between humans with mutations in *SEC23B* and mice deficient for the same paralog [78,105].

In assessing the expression pattern of *Sec24c* in the mouse using RT-PCR, it became apparent that there were in fact two species of *Sec24c* mRNA, *Sec24c-1* and *Sec24c-2*. *Sec24c-2* is not annotated in the genome browser, and contains an additional, in-frame exon between exons 6 and 7, which encodes 23 amino acids (Figure 3-1A). We were unable to detect a third splice variant with a truncated exon 4 and exon 5 currently found in the genome browser, suggesting it is an extremely rare variant. *Sec24c-2* is restricted to the heart, brown fat, brain, and skeletal muscle (Figure 3-1B,C), and expression of the two splice forms is nearly mutually exclusive, suggesting tissue-specific splicing factors may be involved regulating this exon skipping event. Tissue-specific alternative splicing is common, with one report finding that 10-30% of alternatively-spliced human transcripts to have some tissue-specific forms[106]. Brain is known to have a high level of alternative splicing, with many tissue-specific splice factors found in the brain as well as skeletal muscle and the heart [107]. Also of note are reports that brown fat and skeletal muscle arise from a common progenitor [108], raising the possibility that they may share some common splice factors that may contribute to this alternative splicing event in *Sec24c.* The additional exon provides an additional 23 AA to the N-terminus of SEC24C, between Pro330 and Ile331, just upstream of the conserved

zinc-finger domain, which appears to be a repeat of an upstream motif containing I/L-DPD-A/S-IPSP. It is possible that these residues may provide a binding site for a subset of cargoes expressed in these tissues. Alternatively, given the proximity of these additional residues to the binding site of IxM (the t-SNARE ER exit signal for syntaxin 5 and membrin) on SEC24C, it is possible that this peptide sequence could compete for syntaxin 5 membrin binding at this site, as seen in the case of SEC24B's hypervariable region competing with t-SNARE binding [23].   Though it remains to be tested, it is feasible that SEC24C-2 may have the capacity to bind to and recruit a different subset of clients into the budding COPII vesicle than SEC24C-1.

We sought to test whether exogenous SEC24C would be able to rescue the lethal phenotype observed in the *Sec24c*$^{-/-}$ mice. Though transgene mRNA expression could be detected in all or most adult tissues examined (Figure 3-6B), the level and/or developmental timing of this expression appears to be insufficient to replace that of endogenous *Sec24c* or *Sec24d* genes. These results are in contrast to the previously reported rescue of SEC24D deficient mice by BAC transgenes spanning the *Sec24d* locus [28], suggesting that there is some critical regulatory element that is present in the BAC transgenes that is absent in our chicken β-actin promoter driven constructs. These data highlight a critical advantage that BAC transgenes have over standard transgenes – a properly chosen BAC will likely include the promoter and regulatory elements of a particular gene and will more faithfully recapitulate the endogenous expression pattern of the gene of interest [109].

Together, the ubiquitous expression, markedly different phenotypes, and the lack of dosage defects in heterozygous and double heterozygous mice, suggest that each

SEC24 paralog has a unique function, and that while there may be overlap in cargo specificity, it is likely that many cargos require a particular mammalian SEC24 to mediate their exit from the ER.   This specificity may be the result of an intrinsic cargo selectivity of the SEC24 proteins or it may be a function of a particular spatio- or temporal expression pattern of a given SEC24. While a *Sec24c* mutant zebrafish has yet to be reported, a growing body of evidence suggests that the multiple paralogs of the COPII components each have carved out a distinct role for themselves over evolutionary time. These paralogs strike a fine balance between functional redundancy and having completely unique functions, allowing for mammals to efficiently meet the various demands secretory pathway in so many different cellular contexts, as evidenced by the diversity of the phenotypes observed in humans, mice, and zebrafish with mutations in genes encoding components of the COPII coat. While there are still many fundamental questions remaining about the specific functions of the COPII paralogs, the availability of genetic models for deficiency in COPII component genes should open the door for many future studies of COPII function *in vivo*.
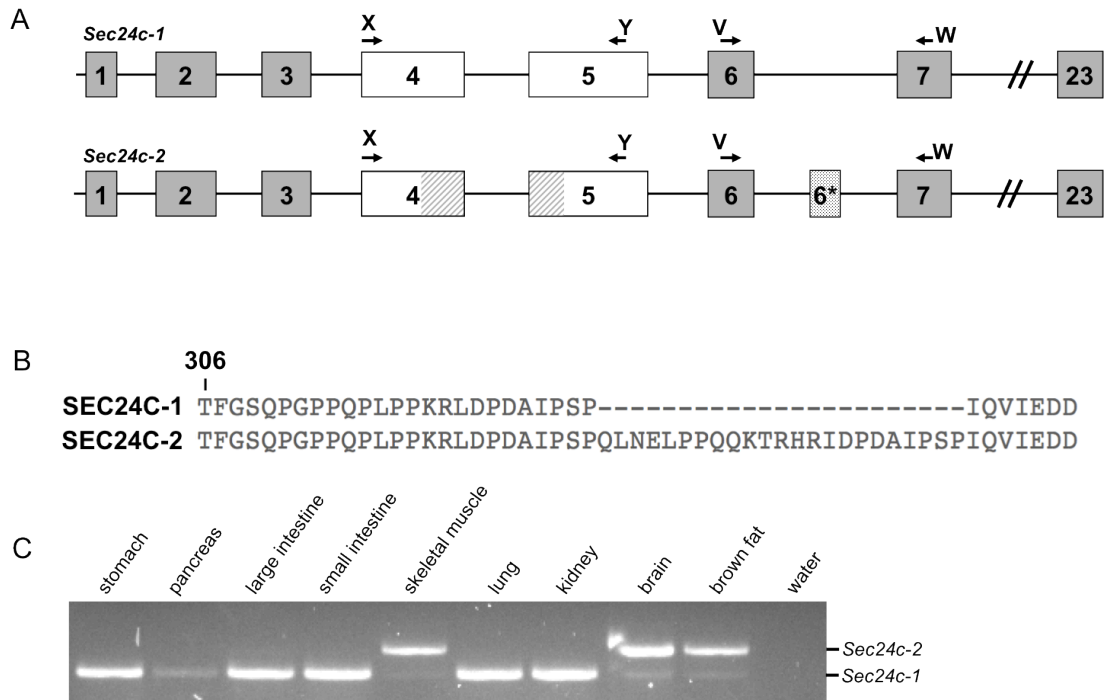
Figure 3-1: Identification of an alternative splice form of *Sec24c*

(A) Schematic of the two *Sec24c* splice forms observed in mice. Arrows indicate primers used to detect the presence of the alternative exon (6*).   Shaded region in exons 4 and 5 represents exonic sequence present in *Sec24c-2* but missing from the second splice form annotated in genome browser. (B) Protein sequence alignment beginning at Thr306 of SEC24C-1 and SEC24C-2 showing additional 23 amino acids encoded in exon 6* found in SEC24C-2.   (C) RT-PCR data showing presence of two forms of *Sec24c* mRNA using primers V and W to generate a 249bp product from *Sec24c-2* and a 180bp product from *Sec24c-1*.
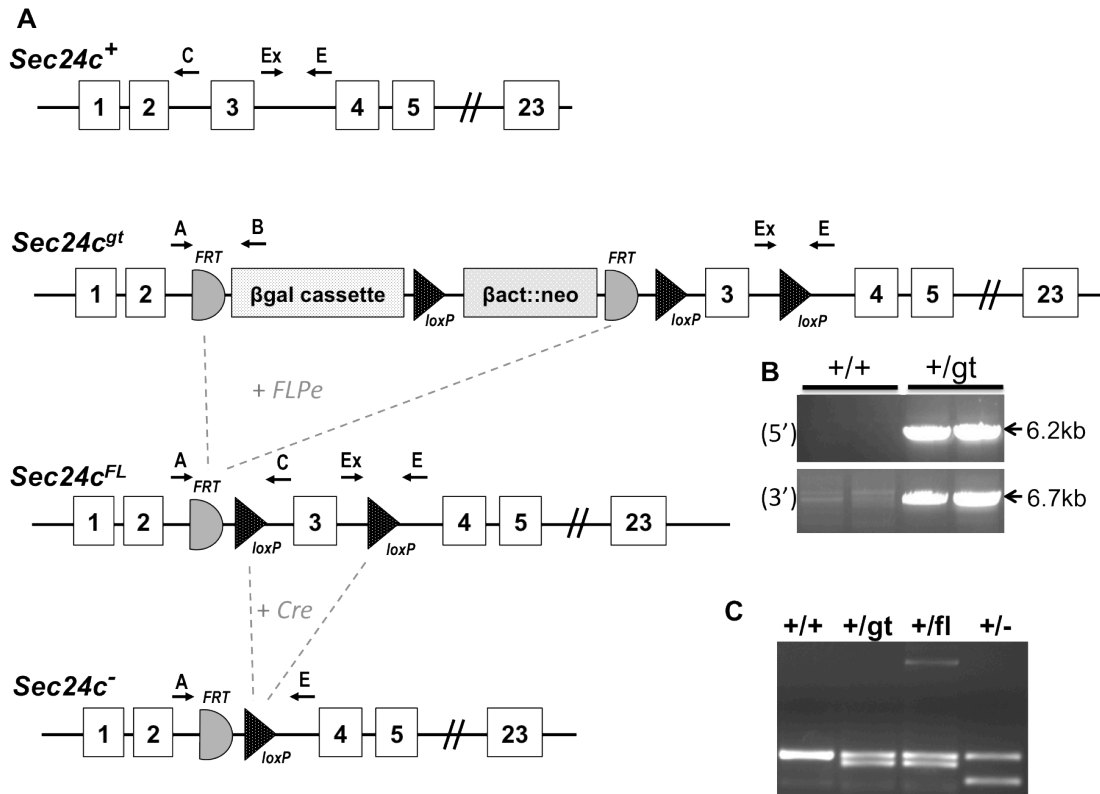
Figure 3-2: Generation of *Sec24c* conditional and null alleles

(A) Schematic of the original $Sec24c^{GT}$ allele.   FLPe mediated excision yields the conditional $Sec24c^{FL}$ allele and subsequent Cre mediated excision gives rise to the null $Sec24c^-$ allele. Primers used for genotyping are indicated. (B) Long range PCR confirms the targeting of the original $Sec24c^{GT}$ allele. Primers are located on either side of the homology arms, so PCR product is only expected in the $Sec24c^{+/GT}$ mice and not in wild type samples. Primers were used to amplify a 6234bp product around the 5' homology arm in the $Sec24c^{GT}$ allele, and a 6648bp product around the 3' homology arm within the $Sec24c^{GT}$ allele. Neither set of primers will yield a band in the $Sec24c^{wt}$ allele. (C) Genotyping PCR assay is able to distinguish between the wild type (308bp), gene-trap or floxed (278bp), and null allele (225bp) using primers A, Ex, and E.
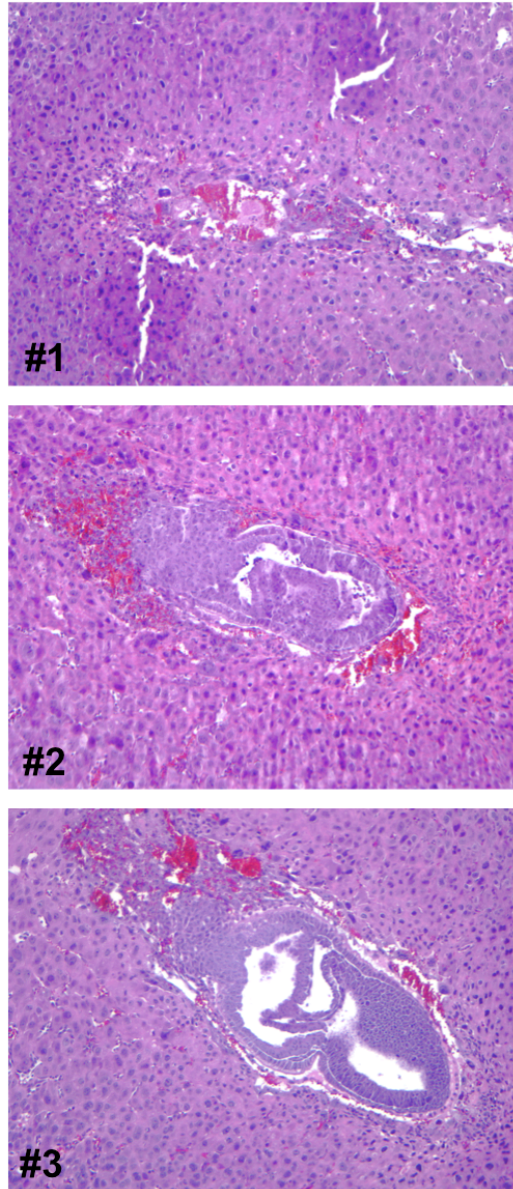
Figure 3-3: Histological analysis of E7.5 *Sec24c$^{-/-}$* embryos

H&E staining of E7.5 embryos resulting from a *Sec24c$^{+/-}$* intercross. E7.5 embryos were fixed, sectioned and stained in the context of the decidual swelling. Images are representative of the 3 classes of developmental states observed in 10 embryos (1= egg cylinder remnants, 2=delayed gastrulation, 3=normal gastrulation).
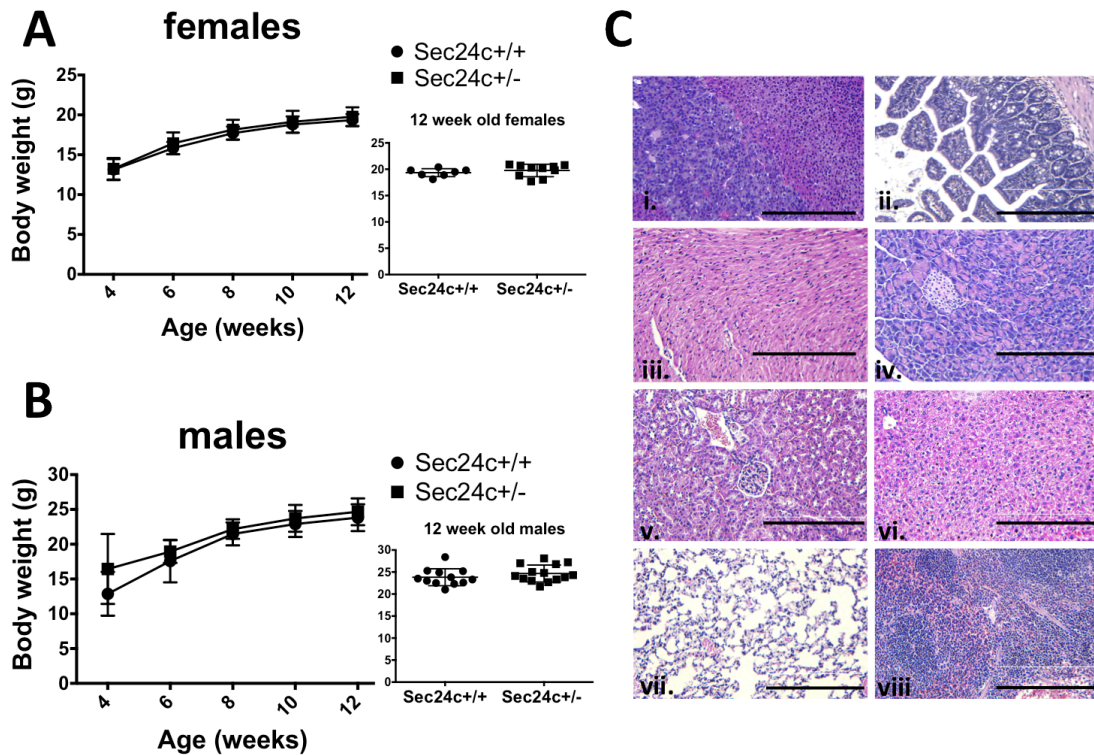
Figure 3-4: *Sec24c*[+/-] **mice are phenotypically normal**

Growth curves of female (A) and male (B) mice indicate no difference in body weights between heterozygous mice and wild type littermate controls, with $p > 0.05$ at all time points. Error bars represent standard deviation. (C) H&E staining of *Sec24c*[+/-] tissues does not reveal any abnormalities in a panel of tissues including adrenal gland (i), intestine (ii), heart (iii), pancreas (iv), kidney (v), liver (vi), lung (vii), and spleen (viii). Scale bar = 0.25μm.
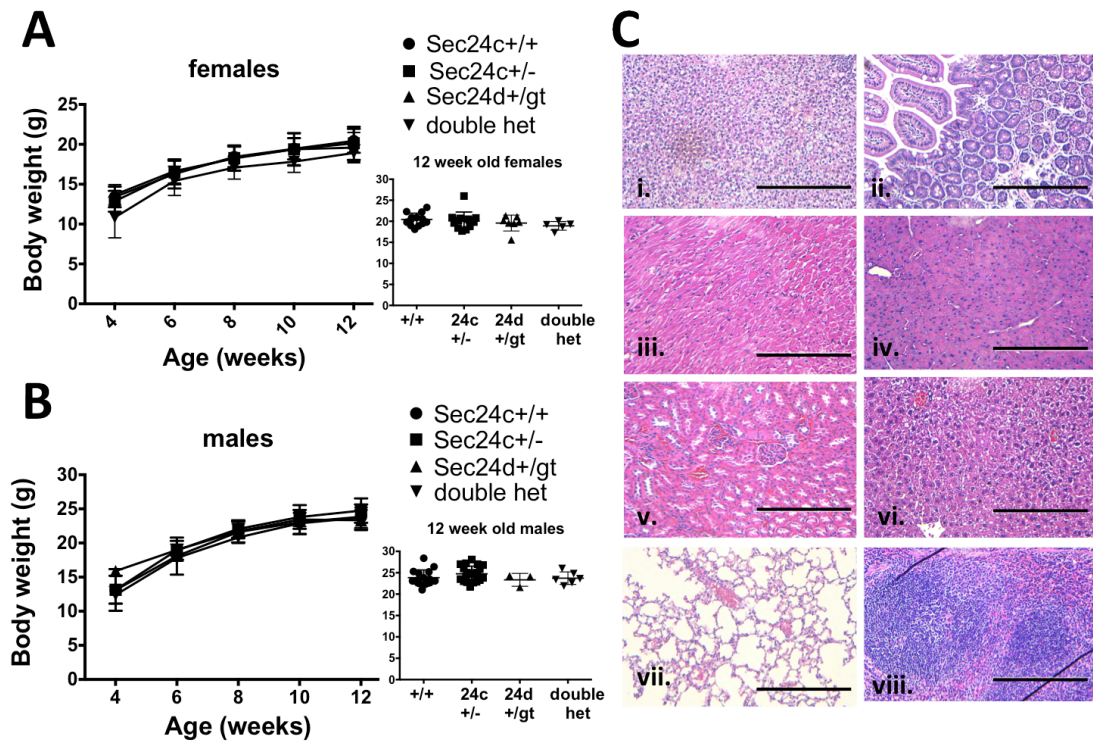
Figure 3-5: *Sec24c$^{+/-}$ Sec24d$^{+/GT}$* **mice are phenotypically normal**

Growth curves of female (A) and male (B) mice indicate no difference in body weights between heterozygous mice and littermate controls, with *p* >0.05 at all time points. Error bars represent standard deviation. "Double het" = *Sec24c$^{+/-}$ Sec24d$^{+/GT}$*. (C) H&E staining of *Sec24c$^{+/-}$ Sec24d$^{+/GT}$* tissues does not reveal any abnormalities in a panel of tissues, including adrenal gland (i), intestine (ii), heart (iii), pancreas (iv), kidney (v), liver (vi), lung (vii), and spleen (viii). Scale bar = 0.25μm.
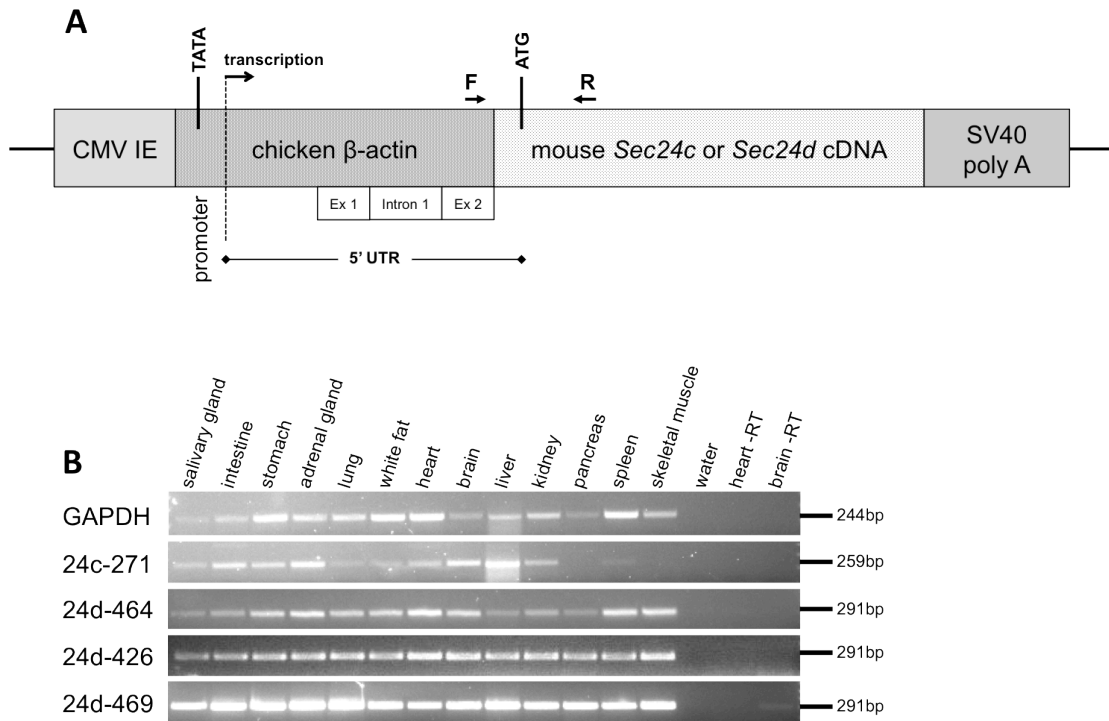
Figure 3-6: Early ubiquitous transgenic expression of *Sec24c* and *Sec24d*

(A) Diagram of the pCAG3zS-Sec24c or -Sec24d transgenic construct, adapted from [99], including transcription start site, initial ATG, and polyadenylation signal.   CMV IE = Cytomegalovirus Immediate-Early Enhancer, UTR = Untranslated Region, SV40 poly A = Simian Virus 40 Polyadenlyation signal. Primer F represents the forward primer present in exon 2 of the promoter used for RT-PCR and primer R represents the reverse primer present in either *Sec24c* or *Sec24d* cDNA used for RT-PCR (B) Transgene expression in a panel of tissues for founder lines 24c-271, 24d-264, 24d-426, 24d-469. Heart and brain –RT samples were carried out in the absence of reverse transcriptase to account for any DNA contamination.

A

p48-Cre+ | Albumin-Cre+ | Villin-Cre+

$Sec24c^{+/FL}$ | $Sec24c^{FL/-}$ | $Sec24c^{+/FL}$ | $Sec24c^{FL/-}$ | $Sec24c^{+/FL}$ | $Sec24c^{FL/-}$
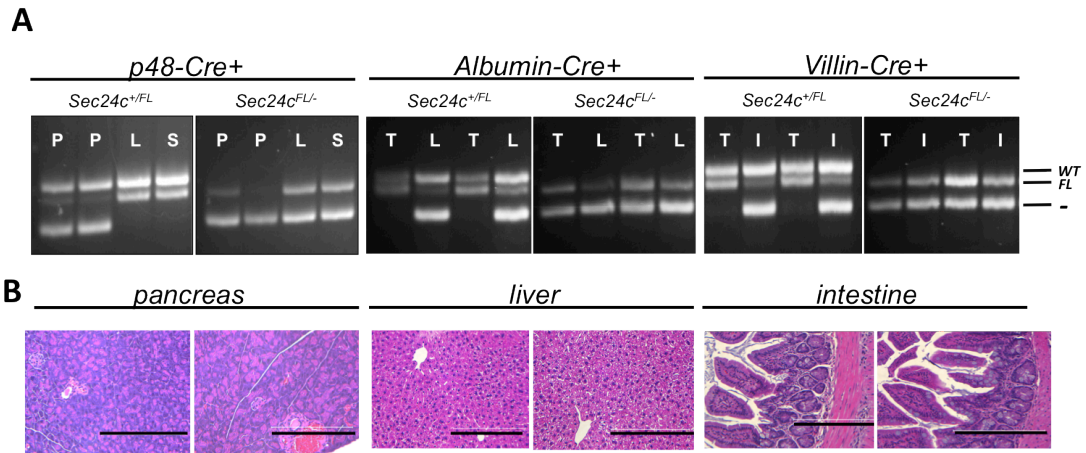
Figure 3-7: Tissue-specific Cre mediated excision of $Sec24c^{FL}$ allele

(A) Genotyping for Cre mediated excision in the pancreas (P), liver (L), spleen (S), intestine (I) and tail (T) in $Sec24c^{+/FL}$ and $Sec24c^{FL/-}$ mice carrying the p48-Cre, Albumin-Cre, or Villin-Cre transgene. The floxed allele is excised to generate the null allele in the pancreas, liver, or intestine, respectively, but is preserved in other tissues. (B) H&E staining of pancreatic tissue from *p48-Cre+* mice, liver tissue from *Albumin-Cre+* mice, and intestine from *Villin-Cre+* mice show no abnormalities in cells lacking SEC24C. Scale bar = 0.25μm.
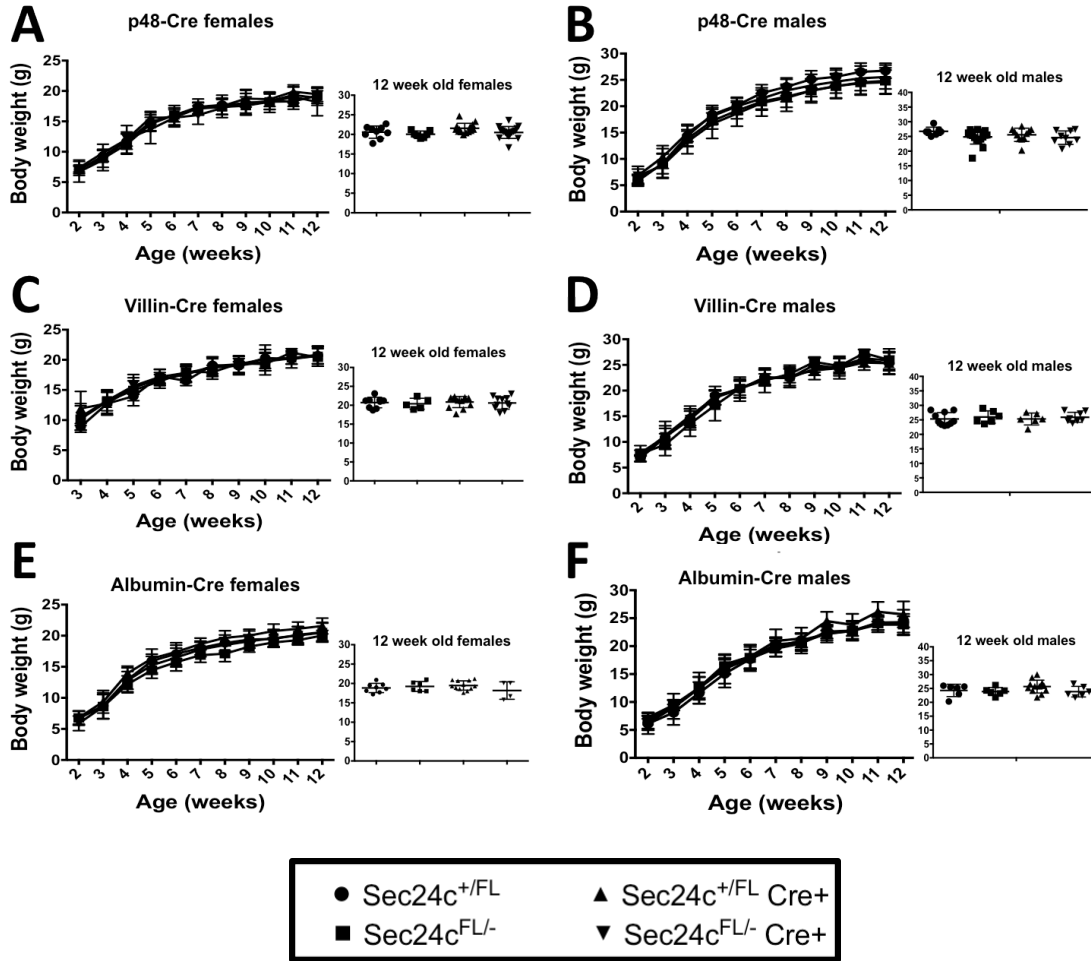
Figure 3-8: Loss of SEC24C in the pancreas, liver, or intestine has no effect on growth

Body weight analysis over 12 weeks of *Sec24c^{FL/-}p48-Cre+* mice (A,B), *Sec24c^{FL/-}Albumin-Cre+* mice (C,D), and *Sec24c^{FL/-}Villin-Cre+* mice (E,F) with corresponding litter mate controls in females and males. Insets show individual body weights at 12 weeks. *p* >0.05 at all time points. Error bars represent standard deviation.
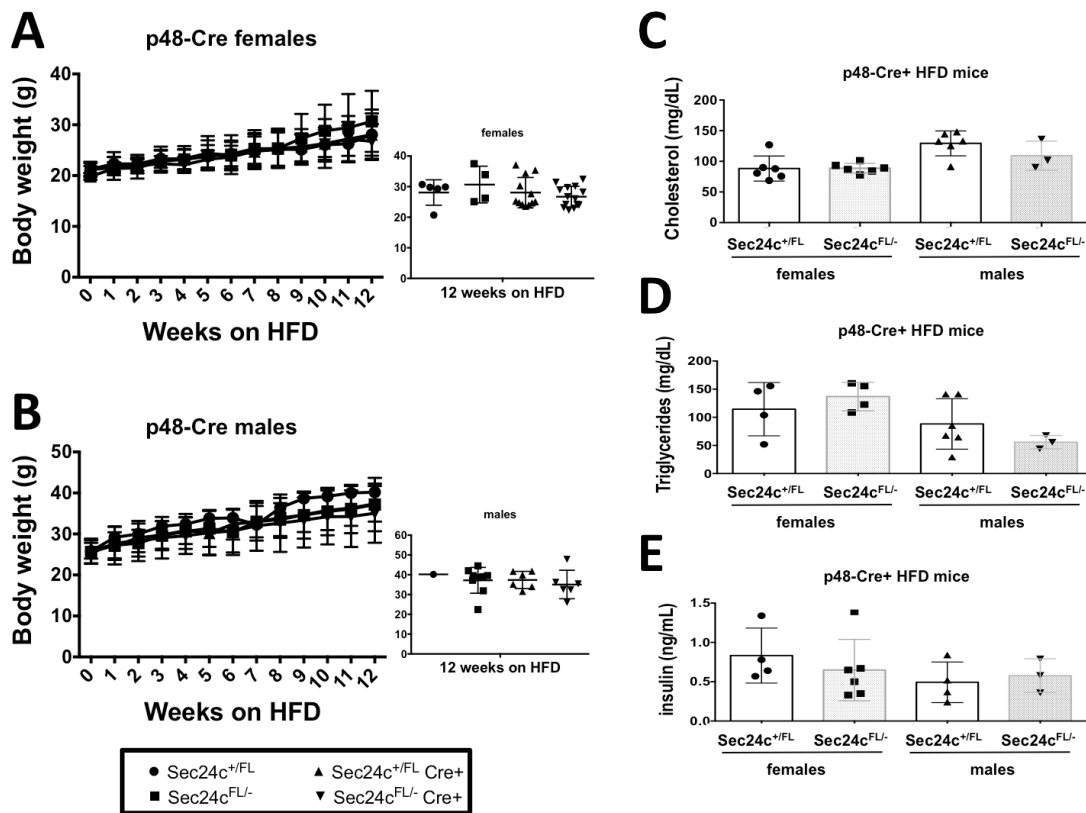
Figure 3-9: *Sec24c^{FL/-}p48-Cre+* **mice display normal diet-induced obesity**

Sec24c^{FL/-}p48-Cre+ and littermate controls show no difference in weight gain over 12 weeks on high fat diet (HFD) for females (A) and males (B), with $p > 0.05$ at all time points. Error bars represent standard deviation. Cholesterol levels (C), triglyceride levels (D) and insulin levels (E) are normal when compared to littermate controls after 12 weeks on HFD. All p-values > 0.05 and error bars represent standard deviation.
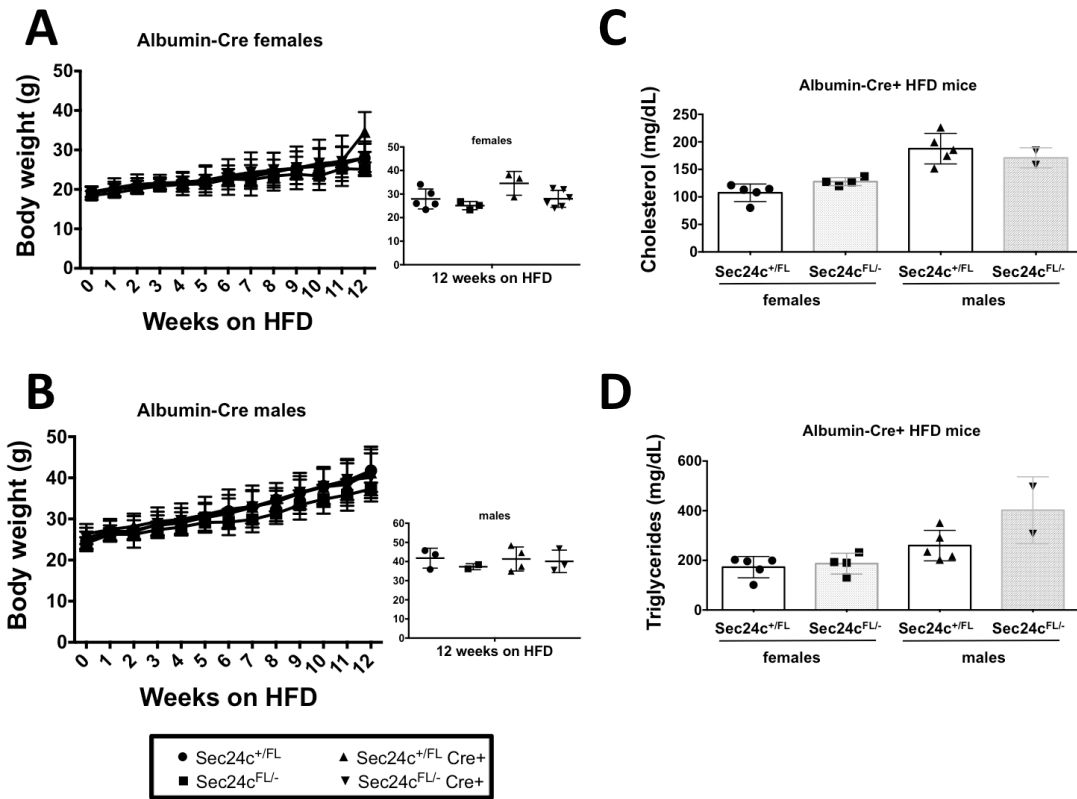
Figure 3-10: *Sec24c^{FL/-}Albumin-Cre+* **mice display normal diet-induced obesity**

Sec24c^{FL/-}Albumin-Cre+ and littermate controls show no difference in weight gain over 12 weeks on high fat diet (HFD) for females (A) and males (B), with $p > 0.05$ at all time points. Error bars represent standard deviation. Cholesterol levels (C), triglyceride levels (D) and insulin levels (E) are normal when compared to littermate controls after 12 weeks on HFD. All p-values > 0.05 and error bars represent standard deviation.
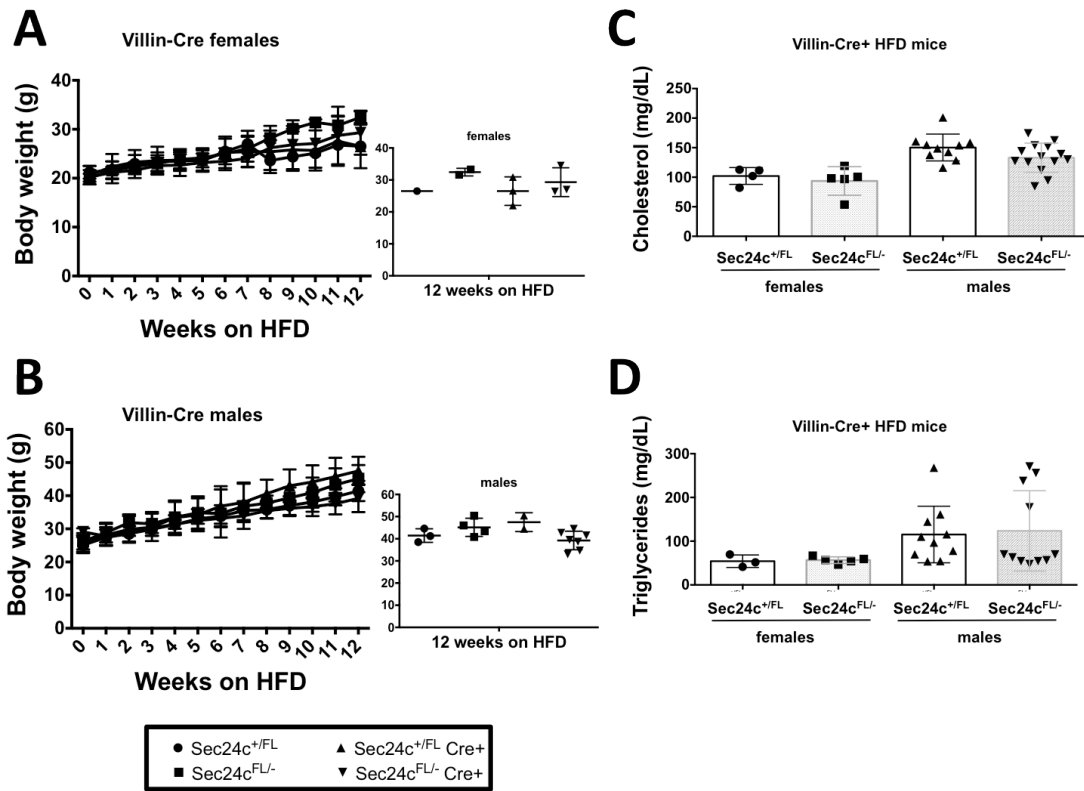
Figure 3-11: *Sec24c^{FL/-}Villin-Cre+* **mice display normal diet-induced obesity**

Sec24c^{FL/-}Villin-Cre+ and littermate controls show no difference in weight gain over 12 weeks on high fat diet (HFD) for females (A) and males (B), with $p > 0.05$ at all time points. Error bars represent standard deviation. Cholesterol levels (C), triglyceride levels (D) and insulin levels (E) are normal when compared to littermate controls after 12 weeks on HFD. All p-values > 0.05 and error bars represent standard deviation.

Genotype:	$Sec24c^{+/+}$	$Sec24c^{+/GT}$	$Sec24c^{GT/GT}$	p-value
$Sec24c^{+/GT}$ x $Sec24c^{+/GT}$ expected:	25%	50%	25%	
2-weeks of age (n=67)	36% (24)	64% (43)	0%	$p < 2.3 \times 10^{-06}$
$Sec24c^{+/GT}$ x $Sec24c^{+/+}$ expected:	50%	50%	--	
2 weeks of age (n=100)	53% (53)	47% (47)	--	$p > 0.54$

Table 3-1: Results of *Sec24c*^{+/GT} intercrosses and backcrosses

Genotype:	$Sec24c^{+/+}$	$Sec24c^{+/FL}$	$Sec24c^{FL/FL}$	p-value
$Sec24c^{+/FL}$ x $Sec24c^{+/FL}$ expected:	25%	50%	25%	
2-weeks of age (n= 102)	26% (26)	42% (43)	32% (33)	$p > 0.08$

Table 3-2: Results of $Sec24c^{+/FL}$ intercross

Genotype:	$Sec24c^{+/+}$	$Sec24c^{+/-}$	$Sec24c^{-/-}$	p-value
$Sec24c^{+/-}$ x $Sec24c^{+/-}$ expected:	25%	50%	25%	
2-weeks of age (n=184)	40% (73)	60% (111)	0%	$p < 4.9x10^{-15}$
E10.5-E11 (n=26)	12% (3)	73% (19)	15% (4*)	$p > 0.25$
E9.5 (n=21)	38% (8)	48% (10)	14% (3*)	$p > 0.25$
E8.5 (n=7)	29% (2)	57% (4)	14% (1)	$p > 0.51$
E7.5 (n=11)	9%(1)	73% (8)	18% (2)	$p > 0.60$
Blastocyst (n=47)	30% (14)	40% (19)	30% (14)	$p > 0.42$
Total (n=170)	**29% (50)**	**58% (99)**	**13% (21)**	$p < 1.4x10^{-04}$
$Sec24c^{+/-}$ x $Sec24c^{+/+}$ expected:	50%	50%	--	
2 weeks of age (n=146)	57% (83)	43% (63)	--	$p > 0.09$

Table 3-3: Results of $Sec24c^{+/-}$ intercrosses and backcrosses.

*All $Sec24c^{-/-}$ embryos from E9.5-E11 were absent, but yolk sacs were able to be isolated for genotyping.

	Genotype			
Parameter	*Sec24c*[+/+]	*Sec24c*[+/-]	*Sec24d*[+/GT]	*Sec24c*[+/-] *Sec24d*[+/GT]
WBC (x 10^3 cells/μL)	7.4 ± 0.9	8.2 ± 1.3	9.9 ± 1.4	9.0 ± 1.1
RBC (x 10^6 cells/μL)	8.7 ± 0.8	9.5 ± 0.3	9.7 ± 0.2	9.9 ± 0.3
HGB (g/dL)	12.2 ± 1.0	12.2 ± 0.5	13.3 ± 0.2	13.0 ± 0.5
HCT (%)	45.2 ± 2.6	45.7 ± 1.5	48.5 ± 0.7	48.4 ± 1.3
MCV (fL)	48.8 ± 0.6	48.1 ± 0.5	49.8 ± 0.4	49.1 ± 0.4
MCH (pg)	13.1 ± 0.3	12.6 ± 0.2	13.6 ± 0.1	13.1 ± 0.1
MCHC (g/dL)	26.9 ± 0.7	26.3 ± 0.5	27.2 ± 0.4	26.8 ± 0.4
RDW (%)	13.5 ± 0.3	13.5 ± 0.2	13.9 ± 0.5	14.0 ± 0.6
PLT (x 10^3 cells/μL)	976.0 ± 171.7	1221.7 ± 53.3	1186.7 ± 80.9	1076.3 ± 73.0
MPV (fL)	5.3 ± 0.6	5.0 ± 0.2	4.8 ± 0.2	5.1 ± 0.2

Table 3-4: Complete blood count analysis of mice from *Sec24c*[+/-] x *Sec24d*[+/GT] intercrosses

Tail bleeds were carried out at 8-12 weeks of age. Values were adjusted back to correct for 10-fold dilution. Data from males and females were pooled after confirming no significant difference between males and females of the same genotypes.

Genotype:	$Sec24c^{+/FL}$	$Sec24c^{+/FL}$ Cre+	$Sec24c^{FL/-}$	$Sec24c^{FL/-}$ Cre+	*p*-value

Genotype:	$Sec24c^{+/+}$ $Sec24d^{+/+}$	$Sec24c^{+/-}$ $Sec24d^{+/+}$	$Sec24c^{+/+}$ $Sec24d^{+/GT}$	$Sec24c^{+/-}$ $Sec24d^{+/GT}$	*p*-value
Expected Ratio of F1 progeny	25%	25%	25%	25%	
Observed F1 progeny (n=94)	32% (30)	26% (24)	24% (23)	18% (17)	p >0.121

Table 3-5: Results of *Sec24c^{+/-} Sec24d^{+/GT}* intercross

Mice are genotyped at 2 weeks of age. $Sec24d^{+/GT}$ mice are on C57BL/6J background ($\geq$N18).

Expected Ratio of test cross progeny	25%	25%	25%	25%	
Cre line:					
p48-Cre	20.9% (31)	25.7% (38)	27.7% (41)	25.7% (38)	*p>0.65*
SM22-Cre	32.5% (13)	15.0% (6)	22.5% (9)	30% (12)	*p>0.46*
Villin-Cre	28.7% (41)	22.4% (32)	26.5% (38)	22.4% (32)	*p>0.46*
Albumin- Cre	23.6% (29)	31.7% (39)	26.0% (32)	18.7% (23)	*p>0.10*
Meox2-Cre	37.7% (26)	20.3% (14)	42.0% (29)	0% (0)	*p<1.7 x10^{-6}*

Table 3-6: Results of tissue-specific knockout of *Sec24c*

Mice are genotyped at 2 weeks of age.

Primer Name	Sequence (5' to 3')
Genotyping and LR-PCR primers	
Sec24c-A	AAGGCGCATAACGATACCA
Sec24c-C	CTTGAGGCAAGAATGCAAAACAAGAATC
Sec24c-D	GTACTAGGTGAGCCTGAAATCAATG
Sec24c-E	ACTAAGATGGGTCCACAAAAGAGC
FLP1	GGTCCAACTGCAGCCCAAGCTTCC
FLP2	GTGGATCGATCCTACCCCTTGCG
Cre-fwd	TTACCGGTCGATGCAACGAGT
Cre-rev	TTCCATCAGTGAACGAACCTGG
24c-GF3	GCTGATACTGATACTAGGATCCACGGACAG
24c-GR4	GCACTGCTAACAGTTCGCTATTCCTTCCG
RAF5 fwd	CACACCTCCCCCTGAACCTGAAAC
Primer 7F	AAATCTGTGCGGAGCCGAAATCTG
Primer 7R	GCATGAACATGGTTAGCAGAGGCT
RT-PCR primers	
GAPDH-QFwd3	TGTGATGGGTGTGAACCACGAGAA
GAPDH-QRev3	ACCAGTGGATGCAGGGATGATGTT
24c primer F	GTGTGACCGGCGGCTCTA
24c primer R	GGGGCTGCAGGTCCTGGTTG
24d primer F	TTGGCAAAGAATTCCTCGACCTG
24d primer R	ATCTTTGAGGAGGATGGCCTGGA
Primer V	GCCCTCAACCTAATTATGAGAGCCCA
Primer W	CCCGTACACCAGTAACAAATGGCT
Primer X	TCTCAGCAGTTTGGTCCT CCATTG
Primer Y	TTTGCTGCAGCTGATAAC CAGG
Primers for transgene construction	
EA053	GAATACTCAAGCTTGCATGCCTGCAGGTCG
EA065	AGAGCATGCATGAGATCACCGGTCATCACCGTG GCGGCCGCTCATAATCAGCCATAC
EA056	ACGTACACCGGTGCTCTCATAATGAATG
EA061	CTAGTCTGCGGCCGCTTAGCTCAGTAGCTGCCG GATC
EA054	ACGTACACCGGTATTTTCATCATGAGCC
EA059	TCGCTAGGCGGCCGCTCAGTTAAGCAGTTG

Table 3-7: List of primers used in Chapter 3

All sequences are listed 5' to 3'.

Notes

This chapter is in preparation for submission under the title "Mammalian COPII

Component SEC24C Is Required For Embryonic Development In Mice" by Elizabeth J.

Adams, Xiao-Wei Chen, K. Sue O'Shea, and David Ginsburg.

Chapter IV: EXAMINING THE OVERLAP IN FUNCTION BETWEEN MOUSE SEC24C AND SEC24D

ABSTRACT

SEC24 is the COPII component thought to be primarily responsible for the recruitment of transmembrane cargoes or cargo adaptors into newly forming COPII vesicles on the ER membrane. Mammalian genomes encode four *Sec24* paralogs (*Sec24a-d)*, though little is known about their comparative functions and cargo-specificities. Based on protein sequence, SEC24A/B are more closely related to one another than they are to SEC24C/D. *Sec24b, Sec24c,* and *Sec24d* knockout mice exhibit embryonic lethality, while *Sec24a* knockouts have low cholesterol levels due to reduced secretion of PCSK9. To test the potential overlap in function between SEC24C/D, we employed dual recombinase mediated cassette exchange to generate a *Sec24c^{c-d}* allele, in which the SEC24C coding sequence has been largely replaced with SEC24D. Crossing mice with the *Sec24c^{c-d}* allele to *Sec24c* demonstrates that SEC24D, when its expression is driven by *Sec24c* regulatory elements, can rescue *Sec24c* null mice from embryonic lethality. However, *Sec24c$^{c-d/c-d}$* mice die shortly after birth suggesting that the overlap in function between SEC24C and SEC24D is incomplete.

INTRODUCTION

In eukaryotic cells, transmembrane or secretory proteins must traverse the secretory pathway before reaching their final intracellular or extracellular destinations [1,2]. The first step of this fundamental process is the concentration and packaging of newly synthesized proteins into vesicles on the surface of the ER at specific ER exit sites [3]. At these sites, cytosolic components assemble to form the COPII complex [4,5], a protein coat that generates membrane curvature and promotes the recruitment of cargo proteins into nascent COPII buds [10,72]. Central to this process is SEC24, the COPII protein component primarily responsible for interaction between transmembrane cargoes and cargo-bound receptors and the coat[73]. This critical component of the COPII inner coat forms a complex with SEC23 in the cytosol, and the SEC23/SEC24 heterodimer is drawn to ER exits sites upon activation of the GTPase SAR1[11] by its cognate ER membrane bound GEF SEC12 [110]. Once recruited to the ER membrane, SEC24 can interact with ER exit signals on protein cargoes via cargo recognition sites and facilitate vesicle formation.

There are four SE24 paralogs encoded in mammalian genomes (*Sec24a-d*). While all four contain a number of highly conserved domains at their C-termini, a hypervariable region is present in the most N-terminal third of the protein sequence. Based on protein sequence identity, the four mammalian SEC24s can be further sub-divided into two subgroups, SEC24A/B and SEC24C/D [93]. Murine SEC24A and B share 58% identity with each other and murine SEC24C and D share 60% identity with each other, yet between the two subgroups there is only 25% sequence identity (Figure 1-3), suggesting an ancient duplication and divergence of these genes. Mice with deficiencies in the

various SEC24 proteins present with a wide range of phenotypes. SEC24A-deficient mice have markedly low plasma cholesterol levels, which was recently reported to be the result of impaired secretion of PCSK9, a regulatory protein involved in LDL receptor degradation and cholesterol homeostasis [26]. SEC24B-deficient mice exhibit embryonic lethality due to a neural tube closure defect attributed to improper trafficking of the planar-cell-polarity protein VANGL2 [27]. Loss of SEC24C in the mouse causes embryonic lethality, though at a much earlier time point, with SEC24C deficient mice dying around E7.5 (*Adams et al., unpublished*). Finally, SEC24D is absolutely required for early embryonic development, as SEC24D null mice fail to develop beyond the 8-cell stage, and are completely absent by E3.5 [28].

The expansion of the number of COPII paralogs over evolutionary time possibly originated as a response to the need for variation in the secretory capacity of particular cell types or temporal contexts. Given that the mammalian *Sec24s* are broadly expressed [28,86,87], there are a few possible factors that may provide each paralog with its own particular niche in functionality: its precise expression pattern or its paralog specific interactions with a subset of cargo molecules, or perhaps other functions beyond cargo recognition. There have been a number of reports suggesting that there are particular cargo repertoires for each SEC24 paralog, including VANGL2 [27] and PCSK9 [26], as mentioned above. The serotonin receptor (SERT) was reported to be an exclusive cargo of SEC24C, with this specificity mediated by an export signal at the N-terminus of SERT [56]; mutation of the export signal on SERT can alter its preference from SEC24C to SEC24D [57]. Likewise, the GABA1 transporter directly interacts with SEC24D for ER export [58]. Yet at the same time, a number of reports have suggested that an overlap in

function exists between the mammalian paralogs of *Sec24,* particularly within the subfamilies by demonstrating that some cargo exit motifs can be recognized by multiple forms of SEC24, including the DxE signal on VSV-G, and the IxM motif on syntaxin 5, both of which confer a specificity for human SEC24A/B [23]. Additionally, the human transmembrane protein p24-p23 is reported to have a preference for SEC24C or SEC24D and is thought to be a cargo receptor for GPI-anchored CD59, which via its interaction with p24-p23, also has specificity for SEC24C/D [111].

To gain a better understanding of the degree of functional overlap between SEC24C/D *in vivo*, we employed a novel technique called dual recombinase mediated cassette exchange (dRMCE) [112], which facilitates the efficient re-engineering of existing alleles that contain heterotypic recombination sites, such as the *FRT* and *LoxP* sites found in many of the "knock-out first" alleles available from the European Mouse Mutagenesis program (EUCOMM). Modification is carried out via co-transfection of embryonic stem cells (ESCs) containing these sites with a plasmid carrying the corresponding recombinases, along with a dRMCE replacement vector containing a custom sequence flanked by the same recombination sites [113]; replacement of endogenous sequence between with recombination sites with sequence provided in the dRMCE replacement vector can occur with high efficiency (reported ranges of 13% to 69%) [112] to generate a newly modified allele.

Given the high degree of amino acid sequence identity between SEC24C and D (58%), and previous examples of similar exit motif recognition between these two paralogs, we carried out dRMCE to facilitate the targeted knock-in of the coding sequence of SEC24D into the *Sec24c* locus to test the ability of SEC24D to functionally

replace SEC24C. In this report, we demonstrate that SEC24D is able to partially rescue

the lethality observed in SEC24C-deficient mice. This *in vivo* analysis sheds light on the

overlap in function found between SEC24C and SEC24D, consistent with previous

reports, and highlights the importance of both expression pattern and protein function to

the specific role of each mammalian *Sec24* paralog.

MATERIALS AND METHODS

Cloning of *Sec24c-d* dRMCE construct

The *Sec24c-d* replacement construct (Figure 4-1) was cloned into the pUC19

vector backbone. The endogenous *Sec24c* intron 2 splice acceptor sequence was

amplified from mouse genomic DNA, and an upstream FRT sequence and SphI

restriction site for cloning was added via PCR. The SV40 polyA sequence present in the

Sec24c^{GT} allele (*Adams et al., unpublished*) was amplified from genomic DNA of a

Sec24c$^{+/GT}$ mouse, and a downstream LoxP sequence and SalI site for cloning was added

using PCR. The partial *Sec24d* coding sequence (from G^{120} to A^{3099} in cDNA sequence)

was fused to the SV40-polyA-LoxP using PCR with a "bridge primer" (half the primer

sequence complimentary to the polyA-LoxP and half complimentary to the *Sec24d*

sequence) to generate the *Sec24d*-SV40pA-LoxP fragment. This partial construct was

stitched downstream of the FRT-*Sec24c* intron2 SA fragment again using a second bridge

primer to connect the two pieces and not to introduce any additional nucleotides between

the key components of the construct. The entire cassette was inserted into pUC19 via

SalI and SphI digests, and the integrity of the sequence was confirmed via DNA

sequencing.

Plasmid purification and microinjections

pDIRE, the plasmid directing dual expression of both iCre and FLPo [112] was obtained from Addgene (Plasmid 26745). Plasmids pCAGGS-iCre and pCAGGS-Flpo (made at the transgenic core) contain the CAG promoter/enhancer [101], which drives recombinase expression of iCre or Flpo in fertilized mouse eggs [114]. All plasmids, including the *Sec24c-d* replacement construct described below, were purified using the Machery-Nagel NucleoBond® Xtra Maxi EF kit, per manufacture's instructions. All microinjections were carried out at the University of Michigan Transgenic Animal Model Core. Co-injections of the *Sec24c-d* construct with pDIRE were done on zygotes generated from the *in vitro* fertilization of C57BL/6J oocytes with sperm from $Sec24^{+/-}$ or $Sec24c^{FL/-}$ male mice (*Adams et al., unpublished*). For each microinjection, 5ng/µl of circular recombinase plasmid mixed with 5ng/µl of circular donor plasmid was administered, as described previously [115,116]. The pCAGGS promoter/enhancer [101] drives recombinase expression in fertilized mouse eggs [114]. Microinjected zygotes were then transferred to pseudopregnant foster mothers, and tail clips for genomic DNA isolation were obtained at 2 weeks of age.

Transient electroporation of ES cells

ES cell clone EPD0241-2-A11 for $Sec24c^{tm1a(EUCOMM)Wtsi}$ ($Sec24c^{gt}$ allele, *Adams et al., unpublished*) was expanded and co-electroporated with the *Sec24c-d* construct and either pDIRE or pCAGGS-iCre and pCAGGS-FLPo. After 1 week, ES cell colonies were picked to six 96 well plates, 3 for the "*Sec24c-d* + pDIRE" and 3 for "*Sec24c-d* + pCAGGS-iCre and pCAGGS-Flpo" condition to compare dRMCE efficiencies using pDIRE versus the CAG driven recombinases. Frozen stock plates and DNA plates were

generated in duplicate, as well as an additional plate for G418 screening.   To test for

G418 sensitivity, cells were fed G418 containing media for 1 week, were then fixed,

stained, and evaluated for growth. Genomic DNA was prepared from each ES cell clone

as previously described [117] and resuspended in TE to serve as template for genotyping.

Genotyping Assays

Genotypes of potential transgenic mice and ES cell clones were determined using

a series of PCR reactions at the *Sec24c* locus. All primers used in this study are listed in

Table 4-1, and expected band sizes can be found in Table 4-2.   Primers S, T, and U used

to amplify a fragment of DNA unique to the *Sec24c-d* replacement construct in Figure 4-

1 to determine the presence or absence of the replacement construct, either at the *Sec24c*

locus (targeted insertion) or elsewhere in the genome (random insertion). To verify a

targeted insertion had occurred, a primer outside the boundary of the replacement

construct (C for 5' and H,J for 3') and a primer contained within the replacement

construct (D,R for 5' and I,F for 3') were used to amplify a product of a particular size

only if the targeted insertion of the replacement construct occurred and was mediated by

the recombinases. Primers pairs for additional genotyping at the *Sec24c* locus to confirm

dRMCE with outermost FRT and LoxP sites on the *Sec24c^{gt}* allele include primers K-

N,A and E (Figure 4-1). Primers iCre10F and iCre10R and Flpo8F and Flpo8R were used

to detect integration of either the pCAGGS-iCre or pCAGGS-Flpo or pDIRE. Mice

carrying the *Sec24d^{gt}* allele were genotyped as describe previously [28].

Generation of *Sec24c$^{+/c-d}$* mice

ES cell clones were cultured as described previously [80] and expanded for
microinjection. ES cell-mouse chimeras were generated as described [81] and then bred
to B6(Cg)-Tyr^{c-2J}/J (JAX stock #000058) to achieve germ-line transmission. ES-cell
derived F1 black progeny were genotyped using primers that will detect the targeted
(*Sec24c^{c-d}*) or wild type (*Sec24c^{wt}*) allele (primers G, E, and F, Figure 4-3). The *Sec24c^{c-d}*
allele was maintained on the C57BL/6J background by continuous backcrosses to
C57BL/6J mice. Initial generations were also genotyped for iCre and Flpo insertions to
confirm the lack of random integrations in backcross progeny.

Long-Range PCR

Genomic DNA from *Sec24c$^{+/+}$* and *Sec24c$^{+/c-d}$* mice were used as templates for a
long-range PCR spanning the original arms of homology used for construction of the
Sec24c^{gt} allele (*Adams et al., unpublished*). PCR was carried out using Phusion Hot Start
II DNA Polymerase (Thermo Scientific), and products were fractionated on a 0.8%
agarose gel. For amplification of the 5' region, primers 24c-GF4 and U were used to
detect the expected 7593bp product in genomic DNA containing the *Sec24c^{c-d}* allele, and
for the 3' primers 24c-GR4 and RAF5 fwd primers were used to amplify a 5798bp
product. Neither set of primers will yield a band in the *Sec24c^{wt}* allele.

RT-PCR

Total RNA was isolated from a tail clip of *Sec24c$^{+/c-d}$* or *Sec24c$^{+/+}$* littermate mice
using the RNAeasy kit (Qiagen) per manufacturer's instructions, with the optional DNasI
digest step included. cDNA synthesis and PCR were carried out in one reaction using
SuperScript® III One-Step RT-PCR System with Platinum®*Taq* (Invitrogen) following

manufacturer's instructions. Forward primers (qF1, qF2) located in *Sec24c* exon 2 and

reverse primers (qR1, qR2) located in the *Sec24d* exons 3 and 4, respectively, were used

to detect the product of proper splicing from endogenous splice donor of exon 2 in

Sec24c into the intron 2 splice acceptor engineered into the replacement cassette.

Timed Matings

Timed matings were carried out by intercrossing $Sec24c^{+/c-d}$ mice. Embryos were

harvested at E17.5 for genotyping and histological analysis. Genotyping was performed

on genomic DNA isolated from tail clip.

Animal Care

All animal care and use complied with the Principles of Laboratory and Animal

Care established by the National Society for Medical Research. The University of

Michigan's University Committee on Use and Care of Animals (UCUCA) approved all

animal protocols in this study.

Statistical Analysis

To determine if there is a statistical deviation from the expected Mendelian ratios

of genotypes from a given cross, the *p*-value reported is the χ^2 value calculated using the

observed ratio of genotypes compared to the expected ratio.

RESULTS

Low efficiency of dRMCE by microinjection

Of the 117 total zygotes that were generated from a cross between $Sec24c^{+/-}$ males

and C57BL/6J females, and co-microinjected with pDIRE and the *Sec24c-d* replacement

construct, 65 were heterozygous for the *Sec24c⁻* allele, which contains the FRT and LoxP

sites required for dRMCE to occur. While these 65 oocytes each carried one copy of this

dRMCE FRT/LoxP "landing pad" none of them underwent targeted recombination to

generate the desired *Sec24c^{c-d}* knock-in allele. Using primers to detect *Sec24d* cDNA

present in the dRMCE *Sec24c-d* construct, we found that while no targeted insertions

were observed, 20 (17%) out of the 117 zygotes that were co-microinjected carried

random insertions of the dRMCE replacement vector.

Identification and subcloning of 6-H9

A screen of 288 ES cell clones co-electroporated with pDIRE and the *Sec24c-d*

targeting construct did not yield any properly targeted insertions, but did identify 18

random insertions of the *Sec24c-d* construct (6.25%, Table 4-3). However, in the second

set of 288 ES cell clones treated with pCAGGS-iCre and pCAGGS-Flpo instead of

pDIRE, we identified 1 potential targeted insertion of the dRMCE replacement construct

into the *Sec24c* locus via the FRT and LoxP sites present in the parental *Sec24c^{gt}* allele

(Table 4-3) and 7 random integrations of the replacement construct. Proper targeting of

this clone, 6-H9, was confirmed by PCR genotyping using primers flanking the

recombination sites on both the 5' and 3' end.

Additional genotyping at the *Sec24c* locus was necessary because the parental ES

cells were heterozygous for the *Sec24c^{gt}* allele, which can undergo recombination when

exposed to the iCre and Flpo recombinases; thus a series of PCRs were run to distinguish

the various states of the *Sec24c* locus after co-electroporation which demonstrated that

clone 6-H9 contained a mixed population of cells, some of which were properly targeted

with recombinations occurring at the outermost FRT and LoxP sites, and others that still

contained the parental $Sec24c^{gt}$ allele, consistent with the mixed resistance to G418 observed with this clone (Table 4-3). Given the possibility that the pCAGGS-iCre or pCAGGS-Flpo can themselves randomly insert into the genome, this clone was also genotyped for the presence of these "transgenes" and was found to contain both iCre and Flpo (data not shown). After one round of subcloning, six different clones (12271-12275) consisting of a pure population of properly targeted ES cells were identified, each with one wild type $Sec24c$ allele and one targeted allele in which the sequence between the outermost FRT and LoxP sites in the $Sec24c^{gt}$ allele was replaced by our dRMCE $Sec24c$-d insert (Figure 4-2A). None of these clones had random insertions of either iCre or Flpo.

Three of these clones (12273,12274, and12275) were microinjected to generate chimeras, and germline transmission of the targeted $Sec24c^{c\text{-}d}$ allele was obtained for clone 12275. The expected mendelian ratio of $Sec24c^{+/c\text{-}d}$ mice was observed in the black F1 progeny of chimeras and B6(Cg)-Tyr$^{c\text{-}2J}$/J (p <0.3), as well as in N2 progeny of subsequent backcrosses to C57BL/6J mice (p <0.3, Table 4-3). RT-PCR analysis of $Sec24c^{+/c\text{-}d}$ RNA indicates that the predicted splice product from the dRMCE $Sec24c$-d allele is produced. Primer sets flanking modified $Sec24c$ intron 2 in the $Sec24c$-d allele produce the expected product from the expected splicing from exon 2 of $Sec24c$ into the splice acceptor engineered upstream of the $Sec24d$ coding sequence (Figure 4-2B). This splice product is only observed in mice heterozygous for the $Sec24c$-d allele and is absent in the wild type mice, as expected.

The SEC24C-D fusion protein can partially rescue the loss of SEC24C

Table 4-3 also shows the results of intercrosses between $Sec24c^{+/c\text{-}d}$ mice. $Sec24c^{c\text{-}d/c\text{-}d}$ mice were not observed at 2 weeks of age (n=16, $p\sim0.02$), yet a close examination of newborn F2 litters revealed that a number of P0 progeny were dying shortly after birth (Figure 4-4). Genotyping analysis revealed that 100% of the dead pups found at P0 were $Sec24c^{c\text{-}d/c\text{-}d}$, indicating the presence of a perinatal lethality in the $Sec24c^{c\text{-}d}$ homozygotes ($p<<0.0001$). At E17.5, $Sec24c^{c\text{-}d/c\text{-}d}$ embryos are found at the expected ratios, yet are smaller than their littermate controls (Figure 4-4).

DISCUSSION

Our data demonstrate that SEC24C and SEC24D have a partial overlap in function, and that within the context of $Sec24c$ gene regulatory sequences, SEC24D can partially rescue the embryonic lethality observed in the SEC24C deficient mice. Loss of SEC24C results in embryonic lethality around E7.5, suggesting that from E7.5 to ~E17.5, SEC24C-D is able to functionally substitute for SEC24C during that developmental period to enough of an extent to prolong survival. However, the SEC24C-D chimeric protein expressed under the $Sec24c$ promoter is unable to fully replace SEC24C in these mice, as $Sec24c^{c\text{-}d/c\text{-}d}$ pups die shortly after birth, though currently the precise cause of death is unknown.

The inability of SEC24D to completely compensate for the loss of SEC24C suggests that functional differences exist between the two proteins, yet the SEC24D protein sequence downstream of Val 41 (Figure 4-1B) is unable to functionally replace the corresponding SEC24C protein sequence (with which it shares 57.5% identity),

supporting the notion that these two proteins have differing affinities for various cargoes, consistent with the paralog-specific interactions of PCSK9 [26]and VANGL2 [27]. A close analysis of the protein sequence alignment between SEC24C-D, the two splice forms of SEC24C, SEC24C-1 and SEC24C-2 (*Adams et al., unpublished*) and SEC24D (Figure 4-5) identifies a motif present in SEC24C-1 and SEC24D (I/L-DPD-A/S-IPSP) that is duplicated in SEC24C-2. The function of this duplication remains unclear, but it is possible that these additional 23 amino acids provide a platform for cargo binding, and in the absence of this sequence (as in the case of *Sec24c$^{-/-}$* mice and *Sec24c$^{c-d/c-d}$* mice), recognition and ER export of a particular protein or set of cargo proteins is diminished to a level not consistent with survival. Further analysis of the *Sec24c$^{c-d/c-d}$* embryos is necessary to determine cause of death; a detailed investigation into the precise phenotype of these mice may reveal some information about a critical cargo or pathway that has been affected by the lack of SEC24C.

The delayed lethality of the *Sec24c$^{c-d/c-d}$* mice also suggests that the precise expression pattern (spatio- or temporal and/or levels) of *Sec24c* is plays a critical role between E7.5 and E17.5; the presence of SEC24D (driven by its endogenous promoter) in *Sec24c$^{-/-}$* mice does not compensate for the SEC24C-deficinecy and mice are lost around E7.5, yet placing it under the regulatory control of *Sec24c* allows the mice to survive longer. Two possible explanations arise in light of this observation: (1) the precise expression profiles of *Sec24c* and *Sec24d* are different; SEC24D is able to functionally replace SEC24C in one context but not the other, or (2) the *Sec24c-d* mini-gene generated by dRMCE is hypomorphic, and either the SEC24C-D protein is not expressed in precisely the same pattern as endogenous SEC24C would be, or the

SEC24C-D protein is not fully functional given its chimeric nature. In either case, the

Sec24c^{c-d} cannot function as a full replacement for *Sec24c* and is unable to completely

rescue mice from SEC24C-deficency.

Based on our data from both microinjections and ESC electroporation, the

efficiency of dRMCE targeting of the *Sec24c* locus without selection is calculated to be

less than 1.5% for microinjections and less than 0.2% for ES cells, both of which are

much lower than the previously reported rate of targeted insertions for ES cells with

selection [112]. The addition of a selectable marker has its advantages in boosting

efficiency in ES cells, but once a targeted clone is identified, extra steps are then

necessary to remove the selectable marker from the genome either by another round of

electroporation or additional crosses to transgenic mice to introduce a third recombinase.

As an alternative to introducing a selectable marker into the *Sec24c-d* replacement

construct, carried out a version negative selection using G418 treatment on a duplicate

plate of cells to test for the presence of the neomycin cassette that is found in the parental

Sec24c^{gt} allele. Theoretically, those cells with our dRMCE mediated targeted

replacement would be sensitive to G418 as the neomycin cassette is recombined out of

the locus, while those that did not undergo recombination were G418 resistant.

Interestingly, cells co-electroporated with pCAGGS-iCre and pCAGGS-Flpo exhibited a

higher percentage of G418 sensitivity, suggesting that the pCAGGS promoter drives

more robust expression than the promoters contained within pDIRE. Many of the clones

had intermediate G418 sensitivity (including 6-H9), likely due to a mixed population of

cells within that well.

Aside from low efficiencies without selection, dRMCE has another critical drawback, in that it relies on the presence of two recombination sites whose location will govern where the targeted modification or insertion can occur. In this report, the location of the FRT site in intron 2 of *Sec24c* dictated what the resulting fusion protein would contain, since exons 1 and 2 of *Sec24c* will still be transcribed in the modified allele, thus resulting in the formation of a *Sec24c-d* mini gene encoding the chimeric SEC24C-D protein. Given recent advances in genome-editing technologies, it appears that the disadvantages of dRMCE mentioned above are causing it to fall out of favor and investigators are instead turning to the more precise clustered regularly interspaced short palindromic repeats, or CRISPR, system [118], which allows the user to make specific mutations at virtually any location in the genome. In contrast to dRMCE, which has not had much success in the generation of mice carrying modified alleles, the CRISPR-Cas system is developing rapidly and has already been used to in a number of studies to correct disease alleles in the mouse via pronuclear injection [119,120,121].

To further address the contributions of expression pattern and protein properties in determining the function of SEC24C and SEC24D, mouse crosses are underway to introduce $Sec24c^{+/c-d}$ into the $Sec24d^{GT}$ background to test if SEC24C-D is able to rescue the very early embryonic lethality seen in $Sec24d^{GT/GT}$ mice [28]. If *Sec24c* driven SEC24C-D can rescue or delay the lethality of these mice, this would support the notion differences in expression between *Sec24c* and *Sec24d* are not as critical as differences in the characteristics of the proteins themselves in defining paralog-specific functions. However, if SEC24C-D is unable to rescue the loss of SEC24D, this could be explained by either insufficient expression or some loss of function of the SEC24C-D compared to

SEC24D due to the first 57 amino acids of the chimeric protein originating from

SEC24C.   In either case, these data, in combination with the partial rescue of *Sec24c* null

mice, will shed more light on the *in vivo* roles of these two *Sec24* paralogs.   These studies

and future *in vivo* investigations on the contributions of the various COPII components

will contribute to our understanding of the complexity of the fundamental process of

COPII mediated ER to Golgi transport.
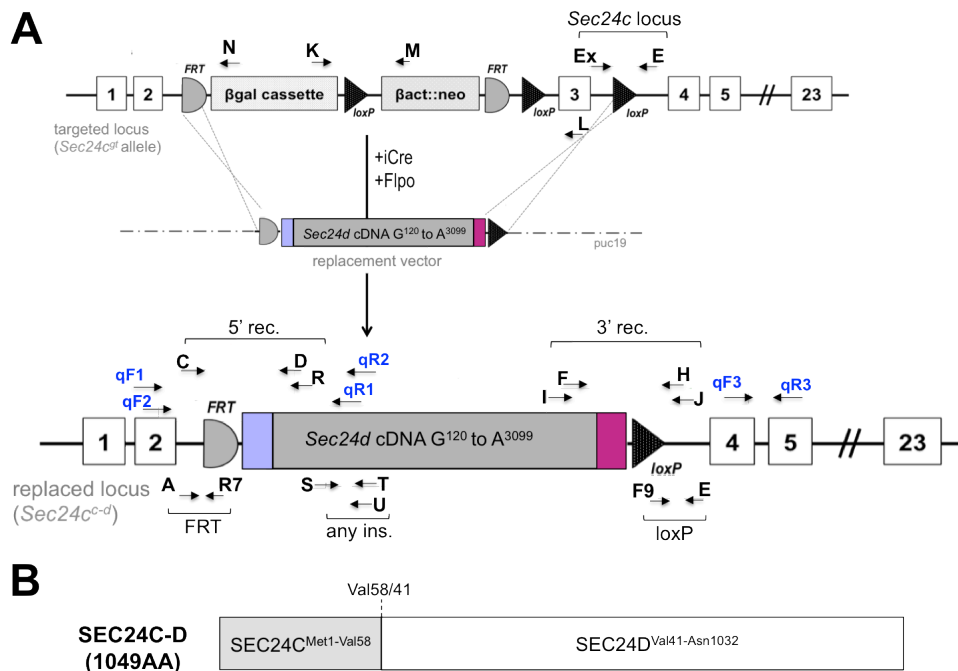
Figure 4-1: Design and generation of the chimeric *Sec24c^{c-d}* allele

(A) Schematic representation of dRMCE process to generate *Sec24c^{c-d}* allele. The replacement vector contains the *Sec24c* intron 2 splice acceptor (purple), the *Sec24d* coding sequence beginning with G^{120} (gray), and a stop codon followed by a poly A signal sequence (magenta). Arrows represent primers used for initial genotype screen to identify properly targeted clones (black) as well as those used for RT-PCR to detect proper splicing into the *Sec24c-d* cDNA from *Sec24c* exon 2 and wild type transcript (blue). (B) Depiction of the SEC24C-D fusion protein encoded by the re-engineered *Sec24c^{c-d}* allele, which contains the first 57 amino acids of SEC24C followed by the SEC24D sequence corresponding to the remaining SEC24C sequence after Val58. The residue Val58/41 (Val58 in SEC24C, Val41 in SEC24D) serves as the junction point for this chimeric protein.

101
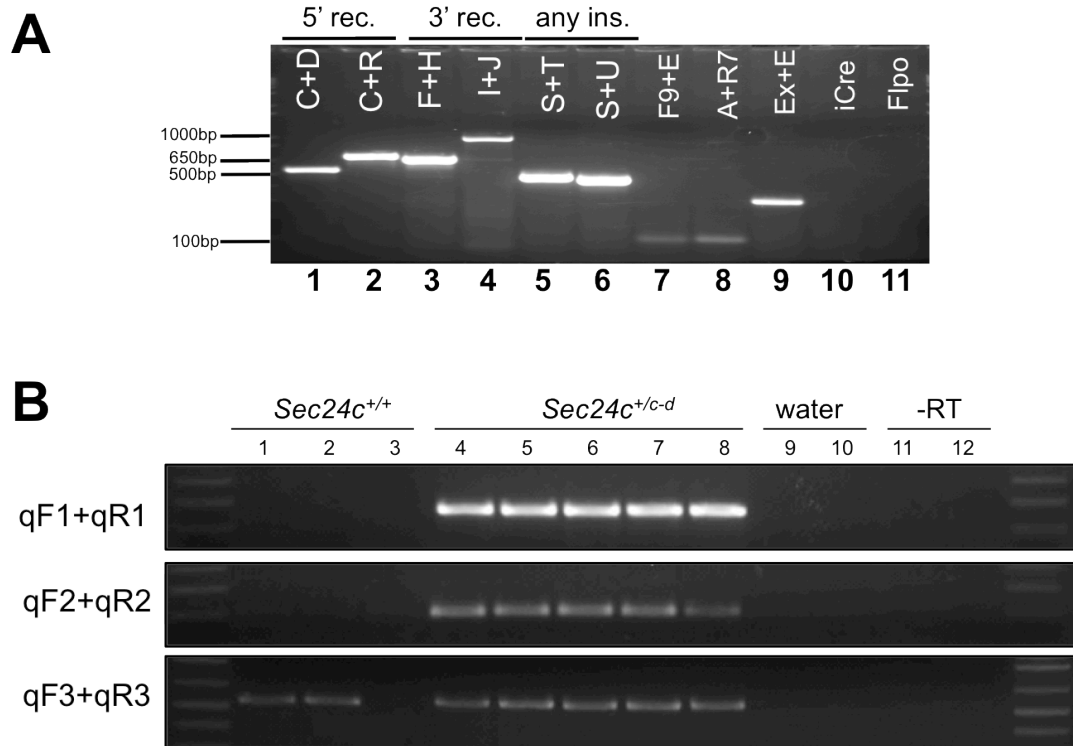
Figure 4-2: Identification of re-engineered clone and detection of proper splicing of
Sec24c^{c-d} **mRNA**

(A) PCR results for dRMCE subclone 12275. Correct targeting was observed for the 5'
recombination site (lanes 1 and 2) and the 3' recombination site (lanes 3 and 4).
Additionally, the presence of the *Sec24d* cDNA (lanes 5,6), LoxP and FRT sites was
confirmed (lanes 7,8). The *Sec24c$^+$* allele was detected in lane 9, confirming that ESC
clone 12275 is heterozygous for the *Sec24c^{c-d}* allele and do not carry any random
insertions of pCAGGS-iCre (lane 10) or pCAGGS-Flpo (lane 11). (B) *Sec24c^{c-d}* mRNA
transcript detected by RT-PCR analysis with primers flanking splice site between *Sec24c*
exon 2 and the *Sec24d* cDNA insert in *Sec24c$^{+/c-d}$* mice (lanes 4-8) but absent in wild
type littermates (lanes 1-3). Note: all primer locations for Figure 4-2 can be found in
Figure 4-1.
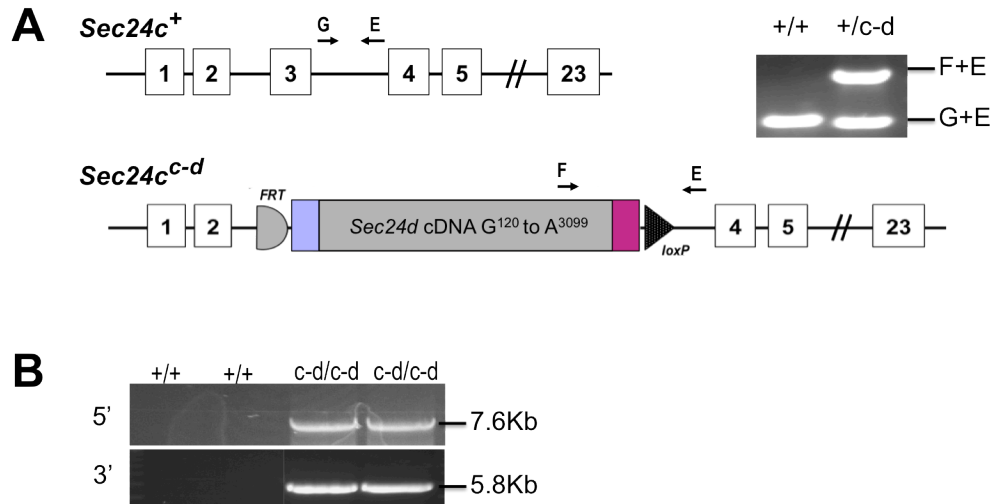
Figure 4-3: Genotyping at the *Sec24c* locus & Long Range PCR

(A) Genotyping PCR assay to distinguish between the wild type and *Sec24c^{c-d}* allele using primers E, F, and G. (B) Long range PCR confirms the targeting of the original *Sec24c^{GT}* allele was not disrupted during dRMCE process. Primers were located outside the homology arms and within the *Sec24d* cDNA, so neither set of primers will yield a band in the *Sec24c^{wt}* allele.
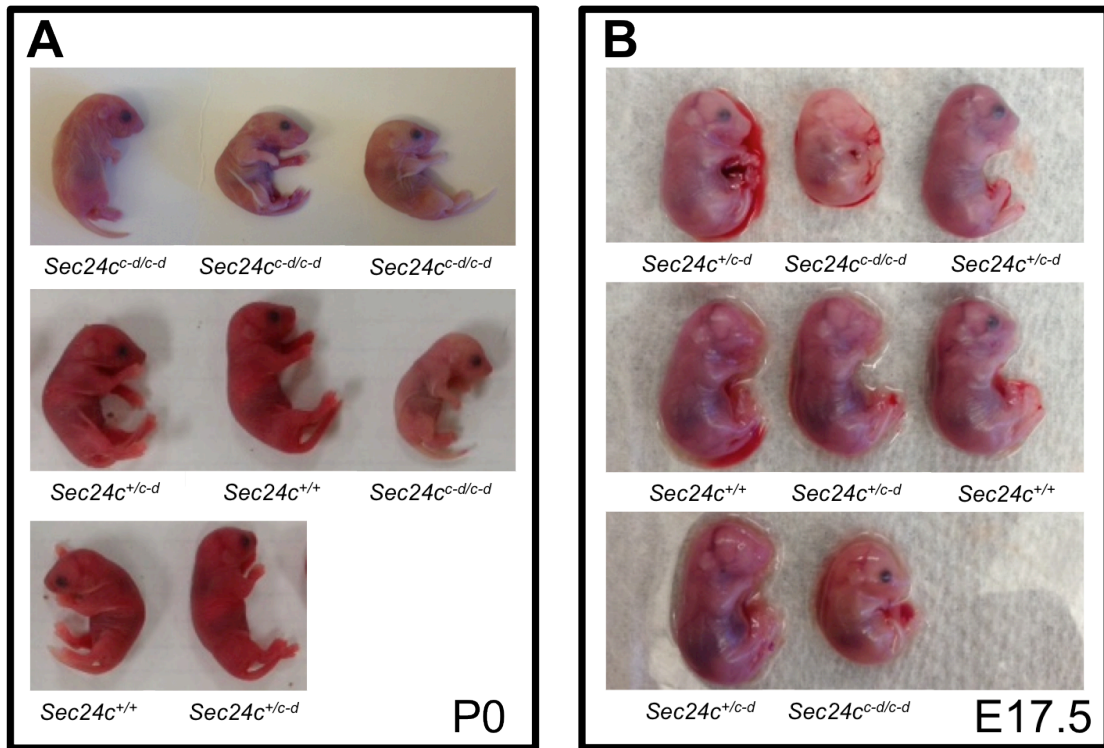
Figure 4-4: *Sec24c^{c-d/c-d}* mice exhibit perinatal lethality and growth defects at E17.5

(A) P0 pups from *Sec24c^{+/c-d}* intercross. *Sec24c^{c-d/c-d}* pups were dead at birth. (B) E17.5 embryos from *Sec24c^{+/c-d}* intercross. *Sec24c^{c-d/c-d}* embryos are markedly smaller than wild type and heterozygous littermates.

```
SEC24C-1  1    MNVNSAPPVPPYGQNQPIYPGYHQSSYGGQPGPAAPATPYGAYNGPVPGYQQAPPQGVPRAPPSSGAPPASAAQVPCGQT  80
SEC24C-2  1    MNVNSAPPVPPYGQNQPIYPGYHQSSYGGQPGPAAPATPYGAYNGPVPGYQQAPPQGVPRAPPSSGAPPASAAQVPCGQT  80
SEC24C-D  1    MNVNSAPPVPPYGQNQPIYPGYHQSSYGGQPGPAAPATPYGAYNGPVPGYQQAPPQGVMKPLGPSAAP------------  68
SEC24D    1    MSQQGYVAT----------PPYSQS----QPGMGISPPHYGHYGDP--SHASSPP-GVMKPLGPSAAP------------  51
                                                                       ↑

SEC24C-1  81   TYGQFGQGDIQNGPSSTAQMQRVPGSQQFGPPLAPVVSQPAVLQPYGPPPTSTQVTAQLAGMQISGAVAQAPPPSGLGYG  160
SEC24C-2  81   TYGQFGQGDIQNGPSSTAQMQRVPGSQQFGPPLAPVVSQPAVLQPYGPPPTSTQVTAQLAGMQISGAVAQAPPPSGLGYG  160
SEC24C-D  69   ------SGMLPPGPLP-------PGPPQFGPNGAHTPGHPPQRFP-GPPPVNSVAPSYASGQTL-------PQSSYPGPG  127
SEC24D    52   ------SGMLPPGPLP-------PGPPQFGPNGAHTPGHPPQRFP-GPPPVNSVAPSYASGQTL-------PQSSYPGPG  110

SEC24C-1  161  PPTSLASASGNFP----NSGPYGSYPQSQAPPLSQAQGHPGVQPPLRSAPPLASSFTSPASGGPQMPSMTGLLPPGQGFG  236
SEC24C-2  161  PPTSLASASGNFP----NSGPYGSYPQSQAPPLSQAQGHPGVQPPLRSAPPLASSFTSPASGGPQMPSMTGLLPPGQGFG  236
SEC24C-D  128  STSSVTQLGSQFSSMQINSYGVGATPQSQGPPGPQSAG--------------------AFQGPPQPAQPSILQPGHQVP  186
SEC24D    111  STSSVTQLGSQFSSMQINSYGVGATPQSQGPPGPQSAG--------------------AFQGPPQPAQPSILQPGHQVP  169

SEC24C-1  237  SLPVNQANHVSSPPAPALPPGTQMTGPPVP--------PPPPMHSPQQPGYQLQQNGSFGPARG-PQPNYESPYPGAPTF  307
SEC24C-2  237  SLPVNQANHVSSPPAPALPPGTQMTGPPVP--------PPPPMHSPQQPGYQLQQNGSFGPARG-PQPNYESPYPGAPTF  307
SEC24C-D  187  PPPPTALNGPGA--SPMSPPTHRQDGLPGPAPLNAQYQPPPPPGQTLGPGYPPQQATNYGPQMGGAQMSYPGGFPGGP--  262
SEC24D    170  PPPPTALNGPGA--SPMSPPTHRQDGLPGPAPLNAQYQPPPPPGQTLGPGYPPQQATNYGPQMGGAQMSYPGGFPGGP--  245

SEC24C-1  308  GSQPGPPQPLPPKRLDPDAIPSP-------------------------IQVIEDDRNNRGSEPFVTGVRGQVPPLVTTNFLV  364
SEC24C-2  308  GSQPGPPQPLPPKRLDPDAIPSPqlnelppqqktrhridpdaispIQVIEDDRNNRGSEPFVTGVRGQVPPLVTTNFLV  387
SEC24C-D  263  -AQMAGPAPQLQRKLDPDSIPSP----------------------IQVIENDRATRGGQVYTTNTRGQVPPLVTTECVI  318
SEC24D    246  -AQMAGPAPQLQRKLDPDSIPSP----------------------IQVIENDRATRGGQVYTTNTRGQVPPLVTTECVI  301

SEC24C-1  365  KDQGNASPRYIRCTSYNIPCTSDMAKQAQVPLAAVIKPLARLPPEEASPYVVDHGESGPLRCNRCKAYMCPLMTFIEGGR  444
SEC24C-2  388  KDQGNASPRYIRCTSYNIPCTSDMAKQAQVPLAAVIKPLARLPPEEASPYVVDHGESGPLRCNRCKAYMCPLMTFIEGGR  467
SEC24C-D  319  QDQGHSSPRYIRCTTYCFPCTSDMAKQAQIPLAAVIKPFADIPPNETPLYLVNHGESGPVRCNRCKAYMCPFMQFIEGGR  398
SEC24D    302  QDQGHSSPRYIRCTTYCFPCTSDMAKQAQIPLAAVIKPFADIPPNETPLYLVNHGESGPVRCNRCKAYMCPFMQFIEGGR  381
```

Figure 4-5: Multiple sequence alignment analysis of SEC24C-D

Alignment of the N-terminal third of SEC24C-D with the two splice forms of SEC24C (SEC24C-1 and -2) and SEC24D. Arrow depicts the Val58/41 junction residue present in SEC24C-D. Red indicates conserved resides and grey indicates location of insertions/deletions. The additional 23 amino acids included in SEC24C-2 are in lowercase. Repeat sequence is bracketed.

Primer Name	Sequence (5' to 3')
Genotyping and Long Range PCR primers	
A	AAGGCGCATAACGATACCA
C	TGAAGGCGCATAACGATACCACGA
D	CTGGCCTGACGCATAAGAGGGTGCCACACT
E	ACTAAGATGGGTCCACAAAAGAGC
F	GGTGGGAAGTCCACACTCTC
F9	CTTCGTATAGCATACATTATACG
G	GTACTAGGTGAGCCTGAAATCAATG
H	TCCTTCCCGTTTCCTCCTTAGCAT
I	TCGTACGTGGATTTCCTCTGCTGT
J	TGAGTTTGAGGTCCAACCTGGTCT
K	GCTACCATTACCAGTTGGTCTGGTGTCA
L	CTGACCACAAGGGACCTGTGCTGCAGAGGC
M	CCTGCGTGCAATCCATCTTGTTCAATGGC
N	CACAACGGGTTCTTCTGTTAGTCC
R	ATGGCTTCATCACGCCTAAGGGTA
R7	CCTATACTTTCTAGAGAATAGGAAC
S	GGTGCACAGATGTCTTACCCAGGAG
T	CATTCACACAGCTGCAGAATCCGCACTGATA
U	GAATCCGCACTGATAGCGCCGTCCGCCT
iCre10F	AGATGCTCCTGTCTGTGTGCAGAT
iCre10R	AGGGACACAGCATTGGAGTCAGAA
Flpo8F	TGATGAGCCAGTTCGACATCCTGT
Flpo8R	AGCATCTTCTTGCTGTGGCTGTTG
24c-GF4	CAGCTGATACTGATACTAGGATCCACGGAC
24c-GR4	GCACTGCTAACAGTTCGCTATTCCTTCCG
RAF5	CACACCTCCCCCTGAACCTGAAAC
RT-PCR primers	
qF1	GTGGGCAACCAGGACCTGCAGCCCCTGCTACTC
qF2	ACCAGTCAGCTCCACCTGTTCCACCATATGGGC
qR1	CTGGCCTGACGCATAAGAGGGTGCCACACTAT
qR2	GGAGCAGCTGAAGGCCCCAATGGCTTCATC
qF3	TCTCAGCAGTTTGGTCCTCCATTG
qR3	TTTGCTGCAGCTGATAACCAGG

Table 4-1: List of primers used in Chapter 4

All primer sequences listed 5' to 3'.

Primer Paris	$Sec24c^+$	$Sec24c^{c\text{-}d}$	$Sec24c^{gt}$	Type
C+D	--	563	--	Screening PCR
C+R	--	727	--	Screening PCR
F+H	--	689	--	Screening PCR
I+J	--	993	--	Screening PCR
S+T	--	484	--	Screening PCR
S+U	--	468	--	Screening PCR
F9+E	--	103	103	Screening PCR
A+R7	--	106	106	Screening PCR
G+E	308	--	308	Genotyping
F+E	--	568	--	Genotyping
qF1+qR1	--	266	--	RT-PCR
qF2+qR2	--	196	--	RT-PCR
qF3+qR3	238	--	--	RT-PCR
24c-GF4+U	--	7593	--	Long range PCR
RAF5+ 24c-GR4	--	5798	6647	Long range PCR

Table 4-2: Expected PCR products from various Sec24c alleles

Primer pairs listed correspond to assays carried out either screening PCRs of ESC clones, genotyping of mice potentially carrying the $Sec24c^{c\text{-}d}$ allele, and long range PCRs confirming the genomic location of the modified $Sec24c$ locus.

Recombinase Source	G418 Sensitive	G418 Resistant	Mixed	Targeted Insertion	Random Insertions
				$Sec24c^{+/c\text{-}d}$	
pDIRE *(plates 1-3)*	31 (10.7%)	247 (85.8%)	10 (3.5%)	0	18/288 (6.25%)
pCAGGS *(plates 4-6)*	42 (14.6%)	157 (54.5%)	89 (30.9%)	1*	7/288 (2.4%)
Total	73 (12.7%)	404 (70.1%)	99 (17.2%)	1 (0.0017%)	25 (4.3%)

Table 4-3: Summary of ESC co-electroporation results

Results of co-electroporation of $Sec24c^{+/GT}$ ESCs with the *Sec24c-d* replacement constructs and either pDIRE or pCAGGS-iCre and pCAGGS-Flpo. G418 resistant clones have not undergone recombination to remove the neomycin cassette present in the parental $Sec24c^{GT}$ allele. One targeted clone identified (*=6-H9) was originally screened as a mixed population of G418 sensitive and resistant ESCs.

Cross	Genotypes			
$Sec24c^{+/c\text{-}d}$ X $Sec24c^{+/+}$	$Sec24c^{+/+}$	$Sec24c^{+/c\text{-}d}$		
Expected	50%	50%	***p-value***	
Observed: chimera F1 (n=38)	58% (22)	42% (16)	*p* >0.3	
C57BL/6J N2 (n=37)	43% (16)	57% (21)	*p* >0.4	
Total	**50.6% (38)**	**49.4% (37)**	***p* >0.8**	
$Sec24c^{+/c\text{-}d}$ intercross	$Sec24c^{+/+}$	$Sec24c^{+/c\text{-}d}$	$Sec24c^{c\text{-}d/c\text{-}d}$	
Expected	25%	50%	25%	***p-value***
Observed: P14 (n=43)	28% (12)	72% (31)	0	<1.6x10^{-4}
E17.5 (n=8)	25% (2)	50% (4)	25% (2)	> 0.99
Dead at P0 (n=6)	0	0	6	*<<0.01*

Table 4-4: Results of $Sec24c^{+/c\text{-}d}$ backcrosses and intercrosses

Genotypes shown for chimera F1only include those of black chimera/ B6(Cg)-Tyr$^{c\text{-}2J}$/J F1 progeny.

CHAPTER V: CONCLUSIONS

Protein trafficking in the intracellular secretory pathway is carefully regulated to ensure the proper sorting and delivery of a diverse range of cargo proteins to many different cellular environments. The transport of proteins through the early secretory pathway is a critical cellular function that is tightly regulated and highly conserved.   This process begins in the ER with the recruitment of newly synthesized proteins into COPII coated vesicles for transport to the Golgi apparatus, where they undergo further processing and sorting to their final destinations. Central to ER exit via COPII vesicles is the selective recruitment of proteins into the nascent vesicles via SEC24. The work described in this dissertation aims to understand the specific roles of the mammalian paralogs of *Sec24 (Sec24a-d)* and to what extent their functional and expression differences play a role in determining their selective interactions with cargo proteins.

Essential role for SEC24D in early embryonic development

In Chapter II, we demonstrated that SEC24D plays a critical role in early embryonic development in the mouse. We generated two independent lines of SEC24D-deficient mice using gene trap alleles designated *Sec24d^{gt}* and *Sec24d^{gt2}*. In both gene trap lines, *Sec24d* null mice were absent when the offspring of F1 intercrosses were genotyped at two weeks of age. Genotypic analysis at the blastocyst stage showed a complete lack of *Sec24d$^{gt/gt}$* embryos in the *Sec24d^{gt}* line, indicating a requirement for SEC24D at this very early stage of development. Surprisingly, *Sec24d$^{gt2/gt2}$* blastocysts

were observed in expected numbers. Further investigation into this phenomenon revealed that $Sec24d^{gt2}$ is a hypomorphic allele, attributed to a very low level (less than 1%) of normal pre-mRNA splicing around the gene trap cassette producing a small amount of wild type $Sec24d$ transcript sufficient to bypass lethality at the blastocyst stage, but not enough to sustain later stages of development (some point between E10.5 and E13.5). The delay in lethality observed is the $Sec24d^{gt2/gt2}$ suggests that prior to E10.5, reduced levels of SEC24D-mediated ER exit are tolerated, likely due to compensation by SEC24A-C; however, at around E10.5, these low levels of SEC24D are insufficient for the faithful secretion of essential SEC24D-dependant cargo or cargoes whose transport from the ER is absolutely required for later embryonic survival.

Determining the time point of *Sec24d* null embryonic lethality

Loss of SEC24D results in embryonic lethality prior to the blastocyst stage, indicating an essential role for SEC24D in very early stages of mammalian development, and suggesting the possibility that SEC24D is absolutely required at the single cell level. Analysis at the 8-cell stage of embryonic development revealed the presence of a single embryo that was $Sec24d^{gt/gt}$. The latter could be the result of maternal contribution of *Sec24d* mRNA, which can persist during the first few stages of embryonic development [122]. Furthermore, given that <1% of normal levels of *Sec24d* transcript are sufficient to delay the lethality observed in *Sec24d* null mice, even a small amount of maternal *Sec24d* mRNA could allow for the passage of a null embryo beyond the first several rounds of cell division.

Essential role for SEC24C in embryonic development

The experiments described in Chapter III revealed a requirement for the related paralog, SEC24C, during embryonic development. Using embryonic stem cells (ESCs) carrying the $Sec24c^{tm1a(EUCOMM)Wtsi}$ ($Sec24c^{gt}$) allele, we generated two versions of SEC24C-deficient mice, and in both lines, SEC24C-deficient mice were absent when the offspring of intercrosses were genotyped at two weeks of age. A more detailed analysis of the embryonic requirement for $Sec24c$ was carried out using the $Sec24c$ null ($Sec24c^{-}$) allele. In contrast to what was observed in $Sec24d^{gt/gt}$ blastocysts, $Sec24c^{-/-}$ embryos were present in the expected Mendelian ratio at the blastocyst stage, with loss at approximately E7.5. While the exact cause of this lethality is unknown, it is likely a manifestation of a block in ER export of a cargo or set of cargoes that rely preferentially on SEC24C for recruitment into nascent vesicles, and in the absence of SEC24C, are retained in the ER unable to carry out their normal cellular function.

Normal appearance of $Sec24c^{+/-}$ and $Sec24d^{+/gt}$ and $Sec24c^{+/-}Sec24d^{+/gt}$ mice

Heterozygosity for either $Sec24c$ or $Sec24d$ results in no phenotypic abnormalities, as mice with both genotypes exhibited normal growth and fertility. Indeed, compound heterozygote ($Sec24c^{+/-} Sec24d^{+/gt}$) mice also appear phenotypically normal. Electron microscopic analysis of highly secretory liver and pancreatic tissues, as well as mouse embryonic fibroblasts derived from $Sec24d^{+/gt}$ embryos, failed to demonstrate ER dilation due to a loss of COPII function. These data, as well as the low levels of normal $Sec24d$ needed to extend the lifespan of $Sec24d$ null embryos, demonstrate that loss of a substantial amount of either SEC24C or SEC24D is surprisingly well tolerated.

Identification of a tissue-specific alternative splice form

While evaluating the mRNA expression profile of *Sec24a-d* in mouse tissues using RT-PCR, we discovered the presence of an alternative splice form of *Sec24c* in a select set of tissues, including the brain, heart, skeletal muscle, and brown fat. Interestingly, the expression of the two *Sec24c* splice forms appears relatively mutually exclusive; in tissues expressing *Sec24c-2,* very little *Sec24c-1* transcript is detected, and vice versa. The 23 additional amino acids present in the SEC24C2 appear to be a repeat of an upstream motif containing I/L-DPD-A/S-IPSP (Figure 4-5). The function of this duplication remains unclear, but it is possible that these additional 23 amino acids provide a platform for cargo binding, or have a role in regulating SEC24C's interaction with accessory proteins or other components of the COPII coat. Other work from our lab found that all four mouse SEC24s were able to bind to both SEC23 paralogs in an overexpression system; experiments are currently underway to determine if an N-terminal FLAG tagged version of SEC24C-2 has a preference for SEC23A or SEC23B. The identification of *Sec24c-2* is the first demonstration of tissue-specific alternative splicing in a mammalian COPII gene. However such alternative splicing is thought to be widespread across the genome [106,107]. Preliminary analysis of the human genome and RNA-seq data suggest that the Sec24C-2 alternative exon is conserved, though more investigation is necessary to determine the tissue-specific expression patterns of the two *Sec24c* splice forms in humans. Future experiments to determine the function of the two splice forms will be discussed in a later section.

Tissue-specific knockouts reveal that SEC24C is dispensable in a number of tissues

Our tissue-specific knockout experiments revealed that SEC24C is not required for the proper maintenance and function of a number of tissues, including highly secretory tissues such as pancreas and liver, as well as intestine and smooth muscle. Phenotypic analysis of *Sec24c*^{FL/-}*Cre+* mice (*p48-Cre* for the pancreatic acinar cells, *albumin-Cre* for hepatocytes, *villin*-Cre for intestinal epithelium, and *SM22-Cre* for smooth muscle) did not identify any abnormalities compared to *Sec24c*^{+/FL}*Cre+* littermate controls, despite a relatively high level of Cre-mediated excision in the expected tissues. The extent of excision seen, combined with the cell autonomous function of SEC24, suggests that, in all of these cell types, none of the essential proteins traversing the early secretory pathway require SEC24C and are able to rely on the three remaining SEC24 paralogs for sufficient ER export.

Examining the overlap in function between SEC24C and SEC24D

The cause of the early embryonic lethality in both the *Sec24c* and *Sec24d* null mice is unknown. An absence of either SEC24C or SEC24D likely leads to a cargo-specific defect in protein sorting and secretion of proteins with paralog-specific requirements, though it could also result in a more global disruption of the export of cargo proteins out of the ER. In either case, the lethality observed suggests that the remaining paralogs of SEC24 are unable to compensate for loss of SEC24D at any point, and are only able to compensate for the absence of SEC24C for a relatively short window during early embryonic development. However, a key question is whether the phenotypes we observe in these embryos is the result of a paralog specific function, or a

manifestation of differential temporal or tissue-specific expression patterns for SEC24C and D?

While expression of the four SEC24 paralogs has been detected in all tissues tested in the adult and in embryos from E10.5 and beyond, the precise developmental timing and expression level of each paralog at earlier stages are still unknown. Additionally, the analysis of expression pattern in the adult is at the level of whole tissues, and further investigation into the cell-type specific expression would provide additional insight into the level of functional overlap among the four SEC24 paralogs. Unfortunately, attempts to make anti-peptide antibodies against SEC24C and D in our lab were unsuccessful and commercial antibodies available against the human SEC24 paralogs are unable to recognize the corresponding mouse proteins. While this limits our ability to carry out immunohistochemistry, future studies could potentially use *in situ* hybridization to evaluate expression in various cell types within a given tissue. Mining of growing databases of RNA-seq data (from the Encyclopedia of DNA Elements (ENCODE) and other projects) provides another source of information about the levels of SEC24 paralog mRNA expression across a wide variety of tissues and cell types and at early developmental time points.

In Chapter IV, we took a genetic approach to address this central question of how much functional overlap exists between the four SEC24 paralogs. I used a recently developed technique called dual recombinase mediated cassette exchange (dRMCE) [112] to direct a targeted insertion of *Sec24d* coding sequence into the *Sec24c* locus. This modified locus should drive expression of *Sec24d* under the regulatory control elements of *Sec24c* and test its ability to rescue the loss of SEC24C-deficient embryos. This new

Sec24c^{c-d} allele was tested for its ability to functionally replace endogenous SEC24C. The perinatal lethality observed in *Sec24c^{c-d/c-d}* mice suggests that SEC24C-D, in the expression context of *Sec24c*, is able to functionally replace endogenous SEC24D to an extent that is consistent with embryonic survival, but is unable to fully function in place of endogenous SEC24C. This may also be due to some loss of function of the SEC24C-D compared to SEC24D due to the first 57 amino acids of the chimeric protein originating from SEC24C Current work to generate an N-terminal FLAG tagged version of the SEC24C-D protein is ongoing to test its interactions with SEC23A/B in the setting of an *in vitro* over-expression system.

Can SEC24C-D rescue the loss of SEC24D?

To further address the contributions of expression pattern and protein properties in determining the function of SEC24C and SEC24D, I have crossed *Sec24c^{+/c-d}* into the *Sec24d^{GT}* background to test if SEC24C-D is able to rescue the very early embryonic lethality seen in *Sec24d^{gt/gt}* mice (See Chapter IV and [28]). A cross of compound heterozygous mice (*Sec24c^{+/c-d}Sec24d^{+/gt}*), to *Sec24d^{+/gt}* mice is currently in progress. If *Sec24c* driven SEC24C-D can rescue or delay lethality in these mice, this would support the notion that differences in expression between *Sec24c* and *Sec24d* are not as critical as differences in the characteristics of the proteins themselves in defining paralog-specific functions. However, if SEC24C-D is unable to rescue the loss of SEC24D, this could be explained by either insufficient expression or again, some loss of function of the SEC24C-D compared to SEC24D. In either case, these data, in combination with the partial rescue of *Sec24c* null mice, will shed more light on the *in vivo* roles of these two SEC24 paralogs.

Testing the temporal requirement for SEC24

The early embryonic lethality of *Sec24d* and *Sec24c* null embryos precludes the study of SEC24C or SEC24D function later in development or in the adult mouse with standard gene trap alleles, however the availability of a conditional allele for *Sec24c* allows us to explore the role of SEC24C beyond the critical embryonic time point of E7.5, when *Sec24c* null embryos are lost.

Using mice carrying a ubiquitous, tamoxifen inducible Cre transgene, we can bypass the embryonic lethality to test whether SEC24C is required more broadly in the adult mouse, beyond the cell types tested by our tissue specific Cre transgenes (Chapter III). I a preliminary experiment, I found that 6 days after initial administration of tamoxifen to adult mice, these mice quickly develop a seemingly neurological phenotype, their health rapidly declines, and they die within 24 hours of initial symptoms (data not shown). The severity of the phenotype is directly correlated with the level of excision observed across a wide range of tissues. These data, while still preliminary, clearly indicate that there is more than just an embryonic requirement for SEC24C, though the precise cause of death and the cargoes responsible for this phenotype remain to be identified. Future work to carefully examine the behavior of these mice prior to expiration, as well as detailed histopathological study, is necessary to be able to conclude anything about the precise role of SEC24C in the adult mouse and what cargoes may require SEC24C for efficient ER export.

Future studies of SEC24 function *in vivo*

Traditional manipulation of the mouse genome to generate modified alleles via homologous recombination has historically been a time consuming process filled with

technical challenges that often limit their utility. However, recent advances have made

the *in vivo* study of complex questions in cell biology much more approachable than ever

before. The era of using standard transgenes or bacterial artificial chromosomes (BACs)

to introduce humanized or mutated alleles into mice may be coming to an end, as

advances such as dRMCE (Chapter IV), zinc-finger nucleases (ZFNs), transcription

activator-like effector nucleases (TALENs), and most recently the clustered regularly

interspaced short palindromic repeats/Cas 9 (CRISPR/Cas system) are becoming more

commonplace [123].

For my study of the overlap in function of SEC24C/D, I used dRMCE for a

number of reasons: (i) our *Sec24c* alleles contained the required recombination sites

necessary for dRMCE, (ii) it seemed feasible that translating the method from ESC

electroporation to pronuclear injection would be successful and ultimately take less time

than traditional BAC recombineering or homologus recombination techniques, (iii) and at

the time, ZFNs and TALENs were only beginning to be used in the mouse, and the

CRISPR/Cas system had not been developed yet. While we ultimately were able to

successfully generate the *Sec24c^{c-d}* allele via co-electroporation in ESCs, our attempt at

dRMCE using pronuclear injection was not as robust as we had hoped. Attempts at

microinjecting *in vitro* translated mRNAs for iCre and Flpo were made to address a

potential delay in recombinase expression from a plasmid; however, only a handful of

injected zygotes survived and none carried the targeted allele. At around the same time

we identified clone 6-H9 in our parallel ESC electroporation, leading us to abandon the

microinjection of mRNA in favor of sublconing 6-9H and using ESCs to generate our

Sec24c-d mouse.

As discussed previously, one major disadvantage of dRMCE is that the location of the FRT and LoxP sites dictate where targeted insertions can be made. Conversely, the CRISPR/Cas system allows for precise genome editing and targeted insertions at virtually any location in the genome [123]. In just the span of about one year, many reports of successful use of CRISPR/Cas to modify the mouse genome using microinjection have appeared [118,120,121]. Thus, future experiments to look at the role of the SEC24 paralogs will likely be undertaken using CRISPR/Cas instead of dRMCE.

Given the potential caveats that come with the chimeric SEC24C-D fusion generated by dRMCE, the use of CRISPR/Cas to direct the insertion of the entire SEC24D coding sequence into the *Sec24c* locus at the ATG has clear advantages. This allows for a true replacement of SEC24C with SEC24D, which may potentially clarify the extent of functional overlap of these two highly related proteins. However, it remains possible that this new mini-gene would still disrupt critical regulatory elements within the *Sec24c* locus (introns). To control for this, it would be crucial to also generate a *Sec24c* mini-gene that could also be knocked into the *Sec24c* locus. If mice homozygous for this allele were phenotypically normal it would demonstrate that at least the minimal regulatory elements are maintained following CRISPR/Cas mediated editing at the *Sec24c* locus.

Reciprocal experiments could be carried out to generate mice carrying *Sec24c* mini-genes under the control of the *Sec24d* promoter, but in doing so, the two alternative splice forms must be considered, as it is unclear if either or both of these splice forms are necessary for the survival of the mouse. One could potentially engineer the *Sec24c* mini-gene to include intron 6-7 where the alternative exon resides, to allow for either form of

Sec24c to be transcribed; however, this would generate a very large insertion cassette (>5kb), which may decrease the efficiency of targeted insertion. Alternatively, one could engineer the transgene to contain the appropriate splice signals sufficient to generate either splice form, and in that way reduce the size of the insertion template down to approximately 3.5kb.

To further address the role of these alternative splice forms of SEC24C, CRISPR/Cas could be used to mediate the removal of the splice signals associated with exon 6* and generate a mouse that could only produce *Sec24c-1* transcripts in all tissues. Alternatively, one could modify the region containing exons 6,6* and 7 of *Sec24c* using CRISPR/Cas to create a mouse in which only *Sec24c-2* transcript is present. At this point, the precise role of these additional 23 amino acids is unclear; answering this question will likely require experiments both *in vivo* using these engineered mice as well as *in vitro* studies evaluating the interactions between these two splice forms of SEC24C and potentially identifying cargoes with splice-form specific requirements for ER exit.

Uncovering additional SEC24 paralog-specific cargoes

When we began our tissue-specific *Sec24c* knockout study, we envisioned that if we were able to identify a tissue in which SEC24C was not required, it would provide material from which to generate primary cell lines that do not express SEC24C. However, recent advances in genome-editing technologies have provided alternative approaches for generating SEC24C- or SEC24D- deficient cell lines for the identification of paralog-specific cargoes. The CRISPR/Cas system could be used to generate cell lines deficient in a particular paralog, and from these lines (if viable), one could potentially identify proteins that require a specific SEC24 paralog for efficient ER export.

Stable isotope labeling by amino acids in cell culture (SILAC) [124] could be used to differentially label individual cell lines to allow for a direct, quantitative comparison of the export of a particular protein or set of proteins using MS/MS analysis. Depending on the fractionation technique used, one could potentially identify plasma membrane proteins, proteins destined for various intracellular compartments, or those secreted into the media, that have binding preferences for a particular SEC24 paralog.

In the case of SEC24D, where data from the mouse suggest that the complete loss of SEC24D results in death at the single cell level, an inducible knockdown approach, in combination with SILAC, may be a more appropriate choice to find paralog-specific cargoes. This approach could be used to potentially identify those proteins that exhibit a reduction in ER exit when a particular SEC24 paralog is knocked down.

Additional efforts to understand the overlap in cargo repertoires for each SEC24 paralog in our lab have been met with technical challenges. To identify additional proteins that interact with SEC24A-D, including transmembrane cargoes and cargo adaptors, as well as unidentified cytosolic regulatory factors, immunoprecipitation followed by mass spectrometry (IP-MS) was carried out. Initial attempts were plagued with low signal to noise ratios, such that it was even difficult to detect the presence of SEC24's known binding partner, SEC23, above the high background. Potential explanations include the high background of the RFP tag used on SEC24A-D, as well as the transient nature of a SEC24-cargo protein interaction. Future studies will use FLAG tags, which have less background binding than RFP, as well as crosslinking, to help capture more transient interactions occurring at the site of vesicle formation.   Results from these experiments could help to define the cargo clientele for each SEC24 paralog,

and potentially allow for the discovery of novel ER export motifs that may confer paralog-specificity. Taken together, these experiments to uncover SEC24-cargo interactions could provide enormous insight into role of individual SEC24s in cargo recognition and recruitment.

A proteomics approach to understanding SEC24 behavior

Recent analysis of SEC24A-deficient mice revealed that in the absence of SEC24A, the steady state levels of SEC24B are increased; however the levels of SEC24C/D remain the same [26]. *Sec24b* mRNA levels are unchanged in these mice, suggesting that this increase in SEC24B results from a decrease in degradation rather than an increase in expression. Given this, one could speculate that perhaps SEC24A/B and SEC24C/D are maintained in different cytoplasmic pools, and in the absence of SEC24A, a larger portion of SEC24B remains in a heterodimeric state with SEC23A/B, preventing its degradation.   If true, this represents a previously uncharacterized level of COPII vesicle regulation.

To test this hypothesis, we set out to examine the half-lives of the mammalian COPII components in the cell, which is currently unknown, and then determine if the stability of these components are altered when particular paralogs are knocked out. Using SILAC, we carried out a "pulse-chase" experiment in which we fully labeled HeLa cells by growing them in lysine-deficient media supplemented with non-radioactive "heavy" $^{13}C_6$-lysine for seven doublings. At that point, we switched them back to media containing "light" $^{12}C_6$-lysine, and collected cell lysates at various time points to measure the ratio of heavy:light peptides for a given protein by mass spectrometry. Analysis of these lysates is ongoing, but we expect to be able to determine the half-life of a given

protein in the sample by looking at the decrease in the ratio of heavy-light labeled peptides as a function of time.

Once we have established a steady-state half-life for SEC24A-D, we can repeat this experiment in cell lines that are deficient in a single paralog and determine what the half-lives of the remaining COPII proteins are in that context. Given what was seen in the SEC24A-deficient mice, one might expect that the half-life of SEC24B is increased in cells where SEC24A has been knocked out, and vice versa. Likewise, one might expect the loss of SEC24C would have an impact on the half-life of SEC24D, if indeed these two subgroups were maintained in distinct cytoplasmic pools. This proteomics based method of determining the half-lives of SEC24A-D may provide new insight into the regulation and behavior of the SEC24 paralogs, and may at the same time serve as a model for the study of protein behavior that can be extended to the entire proteome.

In summary, this dissertation has focused on defining the specific roles of the mammalian paralogs SEC24C and SEC24D and determining the features that contribute to their differential functions. These studies provide new insight into the degree of overlap in function and regulation of SEC24A-D, though many questions still remain. With the amplification of COPII paralogs in mammals, how do these proteins behave in a combinatorial fashion to respond to the dynamic needs of the cell? Do different classes of COPII vesicles exist, perhaps containing either SEC24A/B or SEC24C/D or some other combination? If so, are there particular classes of COPII vesicles involved in the trafficking of larger, bulkier cargo such as pro-collagens? Future investigations will shed light on specific contributions of SEC24 paralogs and their specialized interactions,

which will lead to a deeper understanding of the complex, fundamental process of

selective ER export and COPII vesicle biogenesis.

REFERENCES

1. Bonifacino JS, Glick BS (2004) The Mechanisms of Vesicle Budding and Fusion. 116: 153-166.

2. Palade G (1975) Intracellular aspects of the process of protein synthesis. Science 189.

3. Budnik A, Stephens DJ (2009) ER exit sites - Localization and control of COPII vesicle formation. FEBS letters 583: 3796-3803.

4. Lee MCS, Miller EA, Goldberg J, Orci L, Schekman R (2004) Bi-directional Protein Transport Between the ER and Golgi. Annual Review of Cell and Developmental Biology 20: 87-123.

5. Lee MCS, Miller EA (2007) Molecular mechanisms of COPII vesicle formation. Seminars in Cell & Developmental Biology 18: 424-434.

6. Matsuoka K, Orci L, Amherdt Mn, Bednarek SY, Hamamoto S, et al. (1998) COPII-Coated Vesicle Formation Reconstituted with Purified Coat Proteins and Chemically Defined Liposomes. Cell 93: 263-275.

7. Montegna EA, Bhave M, Liu Y, Bhattacharyya D, Glick BS (2012) Sec12 Binds to Sec16 at Transitional ER Sites. PLoS ONE 7: e31156.

8. Yorimitsu T, Sato K (2012) Insights into structural and regulatory roles of Sec16 in COPII vesicle formation at ER exit sites. Molecular Biology of the Cell 23: 2930-2942.

9. Bielli A, Haney CJ, Gabreski G, Watkins SC, Bannykh SI, et al. (2005) Regulation of Sar1 NH2 terminus by GTP binding and hydrolysis promotes membrane deformation to control COPII vesicle fission. The Journal of Cell Biology 171: 919-924.

10. Lee MCS, Orci L, Hamamoto S, Futai E, Ravazzola M, et al. (2005) Sar1p N-Terminal Helix Initiates Membrane Curvature and Completes the Fission of a COPII Vesicle. Cell 122: 605-617.

11. Aridor M, Weissman J, Bannykh S, Nuoffer C, Balch WE (1998) Cargo Selection by the COPII Budding Machinery during Export from the ER. The Journal of Cell Biology 141: 61-70.

12. Kuehn MJ, Herrmann JM, Schekman R (1998) COPII-cargo interactions direct protein sorting into ER-derived transport vesicles. Nature 391: 187-190.

13. Stagg SM, Gurkan C, Fowler DM, LaPointe P, Foss TR, et al. (2006) Structure of the Sec13/31 COPII coat cage. Nature 439: 234-238.

14. Antonny B, Madden D, Hamamoto S, Orci L, Schekman R (2001) Dynamics of the COPII coat with GTP and stable analogues. Nat Cell Biol 3: 531-537.

15. Bi X, Mancias JD, Goldberg J (2007) Insights into COPII Coat Nucleation from the Structure of Sec23-Sar1 Complexed with the Active Fragment of Sec31. Developmental Cell 13: 635-645.

16. Wieland FT, Gleason ML, Serafini TA, Rothman JE (1987) The rate of bulk flow from the endoplasmic reticulum to the cell surface. Cell 50: 289-300.

17. Pfeffer SR, Rothman JE (1987) Biosynthetic Protein Transport and Sorting by the Endoplasmic Reticulum and Golgi. Annual Review of Biochemistry 56: 829-852.

18. Klumperman J (2000) Transport between ER and Golgi. Current Opinion in Cell Biology 12: 445-449.

19. Thor F, Gautschi M, Geiger R, Helenius A (2009) Bulk Flow Revisited: Transport of a Soluble Protein in the Secretory Pathway. Traffic 10: 1819-1830.

20. Martìnez-Men·rguez JA, Geuze HJ, Slot JW, Klumperman J (1999) Vesicular Tubular Clusters between the ER and Golgi Mediate Concentration of Soluble Secretory Proteins by Exclusion from COPI-Coated Vesicles. Cell 98: 81-90.

21. Oprins A, Rabouille C, Posthuma G, Klumperman J, Geuze HJ, et al. (2001) The ER to Golgi Interface is the Major Concentration Site of Secretory Proteins in the Exocrine Pancreatic Cell. Traffic 2: 831-838.

22. Malkus P, Jiang F, Schekman R (2002) Concentrative sorting of secretory cargo proteins into COPII-coated vesicles. The Journal of Cell Biology 159: 915-921.

23. Mancias JD, Goldberg J (2008) Structural basis of cargo membrane protein discrimination by the human COPII coat machinery. EMBO J 27: 2918-2928.

24. Roberg KJ, Crotwell M, Espenshade P, Gimeno R, Kaiser CA (1999) LST1 Is a SEC24 Homologue Used for Selective Export of the Plasma Membrane ATPase from the Endoplasmic Reticulum. The Journal of Cell Biology 145: 659-672.

25. Peng R, De Antoni A, Gallwitz D (2000) Evidence for Overlapping and Distinct Functions in Protein Transport of Coat Protein Sec24p Family Members. Journal of Biological Chemistry 275: 11521-11528.

26. Xiao-Wei C, He W, Kanika B, Pengcheng Z, Zhuo-Xian M, et al. (2013) SEC24A deficiency lowers plasma cholesterol through reduced PCSK9 secretion. eLife 2.

27. Merte J, Jensen D, Wright K, Sarsfield S, Wang Y, et al. (2010) Sec24b selectively sorts Vangl2 to regulate planar cell polarity during neural tube closure. Nat Cell Biol 12: 41-46.

28. Baines AC, Adams EJ, Zhang B, Ginsburg D (2013) Disruption of the *Sec24d* Gene Results in Early Embryonic Lethality in the Mouse. PLoS ONE 8: e61114.

29. Neuhauss SC, Solnica-Krezel L, Schier AF, Zwartkruis F, Stemple DL, et al. (1996) Mutations affecting craniofacial development in zebrafish. Development 123: 357-367.

30. Driever W, Solnica-Krezel L, Schier AF, Neuhauss SC, Malicki J, et al. (1996) A genetic screen for mutations affecting embryogenesis in zebrafish. Development 123: 37-46.

31. Sarmah S, Barrallo-Gimeno A, Melville DB, Topczewski J, Solnica-Krezel L, et al. (2010) Sec24D-Dependent Transport of Extracellular Matrix Proteins Is Required for Zebrafish Skeletal Morphogenesis. PLoS ONE 5: e10367.

32. Lang MR, Lapierre LA, Frotscher M, Goldenring JR, Knapik EW (2006) Secretory COPII coat component Sec23a is essential for craniofacial chondrocyte maturation. Nat Genet 38: 1198-1203.

33. Ohisa S, Inohaya K, Takano Y, Kudo A (2010) sec24d encoding a component of COPII is essential for vertebra formation, revealed by the analysis of the medaka mutant, vbi. Developmental Biology 342: 85-95.

34. Faso C, Chen Y-N, Tamura K, Held M, Zemelis S, et al. (2009) A Missense Mutation in the Arabidopsis COPII Coat Protein Sec24A Induces the Formation of Clusters of the Endoplasmic Reticulum and Golgi Apparatus. The Plant Cell Online 21: 3655-3671.

35. Kanapin A, Batalov S, Davis MJ, Gough J, Grimmond S, et al. (2003) Mouse Proteome Analysis. Genome Research 13: 1335-1344.

36. Baines AC, Zhang B (2007) Receptor-mediated protein transport in the early secretory pathway. Trends Biochem Sci 32: 381-388.

37. Dancourt J, Barlowe C (2010) Protein Sorting Receptors in the Early Secretory Pathway. Annual Review of Biochemistry 79: 777-802.

38. Zhang B, McGee B, Yamaoka J, Guglielmone H, Downes K, et al. (2006) Combined deficiency of factor V and factor VIII is due to mutations in either LMAN1 or MCFD2. Blood 107: 1903-1907.

39. Nyfeler B, Zhang B, Ginsburg D, Kaufman R, Hauri H-P (2006) Cargo selectivity of the ERGIC-53/MCFD2 transport receptor complex. Traffic 7: 1473-1481.

40. Muñiz M, Nuoffer C, Hauri H-P, Riezman H (2000) The Emp24 Complex Recruits a Specific Cargo Molecule into Endoplasmic Reticulum-Derived Vesicles. The Journal of Cell Biology 148: 925-930.

41. Lederkremer GZ, Cheng Y, Petre BM, Vogan E, Springer S, et al. (2001) Structure of the Sec23p/24p and Sec13p/31p complexes of COPII. Proceedings of the National Academy of Sciences 98: 10704-10709.

42. Bi X, Corpina RA, Goldberg J (2002) Structure of the Sec23/24-Sar1 pre-budding complex of the COPII vesicle coat. Nature 419: 271-277.

43. Miller EA, Liu Y, Barlowe C, Schekman R (2005) ER-Golgi Transport Defects Are Associated with Mutations in the Sed5p-binding Domain of the COPII Coat Subunit, Sec24p. Mol Biol Cell 16: 3719-3726.

44. Mossessova E, Bickford LC, Goldberg J (2003) SNARE Selectivity of the COPII Coat. Cell 114: 483-495.

45. Mancias JD, Goldberg J (2007) The Transport Signal on Sec22 for Packaging into COPII-Coated Vesicles Is a Conformational Epitope. Molecular Cell 26: 403-414.

46. Miller EA, Beilharz TH, Malkus PN, Lee MC, Hamamoto S, et al. (2003) Multiple cargo binding sites on the COPII subunit Sec24p ensure capture of diverse membrane proteins into transport vesicles. Cell 114: 497-509.

47. Barlowe C (2003) Signals for COPII-dependent export from the ER: what's the ticket out? Trends in Cell Biology 13: 295-300.

48. Nishimura N, Balch WE (1997) A Di-Acidic Signal Required for Selective Export from the Endoplasmic Reticulum. Science 277: 556-558.

49. Votsmeier C, Gallwitz D (2001) An acidic sequence of a putative yeast Golgi membrane protein binds COPII and facilitates ER export. The EMBO Journal 20: 6742-6750.

50. Wang X, Matteson J, An Y, Moyer B, Yoo J-S, et al. (2004) COPII-dependent export of cystic fibrosis transmembrane conductance regulator from the ER uses a di-acidic exit code. The Journal of Cell Biology 167: 65-74.

51. Ma D, Zerangue N, Lin Y-F, Collins A, Yu M, et al. (2001) Role of ER Export Signals in Controlling Surface Potassium Channel Numbers. Science 291: 316-319.

52. Peng R, Grabowski R, De Antoni A, Gallwitz D (1999) Specific interaction of the yeast cis-Golgi syntaxin Sed5p and the coat protein complex II component Sec24p of endoplasmic reticulum-derived transport vesicles. Proceedings of the National Academy of Sciences 96: 3751-3756.

53. Nufer O, Kappeler F, Guldbrandsen S, Hauri H-P (2003) ER export of ERGIC-53 is controlled by cooperation of targeting determinants in all three of its domains. Journal of Cell Science 116: 4429-4440.

54. Nakamura N, Yamazaki S, Sato K, Nakano A, Sakaguchi M, et al. (1998) Identification of Potential Regulatory Elements for the Transport of Emp24p. Molecular Biology of the Cell 9: 3493-3503.

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

56. Sucic S, El-Kasaby A, Kudlacek O, Sarker S, Sitte HH, et al. (2011) The Serotonin Transporter Is an Exclusive Client of the Coat Protein Complex II (COPII) Component SEC24C. Journal of Biological Chemistry 286: 16482-16490.

57. Sucic S, Koban F, El-Kasaby A, Kudlacek O, Stockner T, et al. (2013) Switching the Clientele: A lysine residing in the C terminus of the serotonin transporter specifies its preference for the coat protein complex II component SEC24C. Journal of Biological Chemistry 288: 5330-5341.

58. Farhan H, Reiterer V, Korkhov VM, Schmid JA, Freissmuth M, et al. (2007) Concentrative Export from the Endoplasmic Reticulum of the Y-Aminobutyric Acid Transporter 1 Requires Binding to SEC24D. Journal of Biological Chemistry 282: 7679-7689.

59. Wendeler MW, Paccaud JP, Hauri HP (2007) Role of Sec24 isoforms in selective export of membrane proteins from the endoplasmic reticulum. EMBO Rep 8: 258-264.

60. Giraudo CG, Maccioni HJF (2003) Endoplasmic Reticulum Export of Glycosyltransferases Depends on Interaction of a Cytoplasmic Dibasic Motif with Sar1. Molecular Biology of the Cell 14: 3753-3766.

61. Venditti R, Wilson C, De Matteis MA Exiting the ER: what we know and what we don't. Trends in Cell Biology.

62. D'Arcangelo JG, Stahmer KR, Miller EA (2013) Vesicle-mediated export from the ER: COPII coat function and regulation. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1833: 2464-2472.

63. Farhan H, Weiss M, Tani K, Kaufman RJ, Hauri HÄ (2008) Adaptation of endoplasmic reticulum exit sites to acute and chronic increases in cargo load. The EMBO Journal 27: 2043-2054.

64. Aridor M, Bannykh SI, Rowe T, Balch WE (1999) Cargo Can Modulate COPII Vesicle Formation from the Endoplasmic Reticulum. Journal of Biological Chemistry 274: 4389-4399.

65. Sato K, Nakano A (2005) Dissection of COPII subunit-cargo assembly and disassembly kinetics during Sar1p-GTP hydrolysis. Nat Struct Mol Biol 12: 167-174.

66. Sato K, Nakano A (2004) Reconstitution of Coat Protein Complex II (COPII) Vesicle Formation from Cargo-reconstituted Proteoliposomes Reveals the Potential Role of GTP Hydrolysis by Sar1p in Protein Sorting. Journal of Biological Chemistry 279: 1330-1335.

67. Copic A, Latham CF, Horlbeck MA, D'Arcangelo JG, Miller EA (2012) ER Cargo Properties Specify a Requirement for COPII Coat Rigidity Mediated by Sec13p. Science 335: 1359-1362.

68. Saito K, Chen M, Bard F, Chen S, Zhou H, et al. (2009) TANGO1 facilitates cargo loading at endoplasmic reticulum exit sites. Cell 136: 891-902.

69. Kung LF, Pagant S, Futai E, Arcangelo JG, Buchanan R, et al. (2011) Sec24p and Sec16p cooperate to regulate the GTP cycle of the COPII coat. The EMBO Journal 31: 1014-1027.

70. Sharpe LJ, Luu W, Brown AJ (2011) Akt Phosphorylates Sec24: New Clues into the Regulation of ER-to-Golgi Trafficking. Traffic 12: 19-27.

71. Siddiqi S, Siddiqi SA, Mansbach CM (2010) Sec24C is required for docking the prechylomicron transport vesicle with the Golgi. Journal of Lipid Research 51: 1093-1100.

72. Bickford LC, Mossessova E, Goldberg J (2004) A structural view of the COPII vesicle coat. Current Opinion in Structural Biology 14: 147-153.

73. Miller E, Antonny B, Hamamoto S, Schekman R (2002) Cargo selection into COPII vesicles is driven by the Sec24p subunit. EMBO Journal 21: 6105-6113.

74. Kurihara T, Hamamoto S, Gimeno RE, Kaiser CA, Schekman R, et al. (2000) Sec24p and Iss1p Function Interchangeably in Transport Vesicle Formation from the Endoplasmic Reticulum in Saccharomyces cerevisiae. Mol Biol Cell 11: 983-998.

75. Tang BL, Kausalya J, Low DYH, Lock ML, Hong W (1999) A Family of Mammalian Proteins Homologous to Yeast Sec24p. Biochemical and Biophysical Research Communications 258: 679-684.

76. Boyadjiev SA, Fromme JC, Ben J, Chong SS, Nauta C, et al. (2006) Cranio-lenticulo-sutural dysplasia is caused by a SEC23A mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking. Nat Genet 38: 1192-1197.

77. Schwarz K, Iolascon A, Verissimo F, Trede NS, Horsley W, et al. (2009) Mutations affecting the secretory COPII coat component SEC23B cause congenital dyserythropoietic anemia type II. Nat Genet 41: 936-940.

78. Tao J, Zhu M, Wang H, Afelik S, Vasievich MP, et al. (2012) SEC23B is required for the maintenance of murine professional secretory tissues. Proceedings of the National Academy of Sciences.

79. Skarnes W (2000) Gene trapping methods for the identification and functional analysis of cell surface proteins in mice. Methods Enzymol 328: 592-615.

80. Kendall SK, Samuelson LC, Saunders TL, Wood RI, Camper SA (1995) Targeted disruption of the pituitary glycoprotein hormone alpha-subunit produces hypogonadal and hypothyroid mice. Genes & development 9: 2007-2019.

81. Bahou W, Bowie E, Fass D, Ginsburg D (1988) Molecular genetic analysis of porcine von Willebrand disease: tight linkage to the von Willebrand factor locus. Blood 72: 308-313.

82. Van Keuren M, Gavrilina G, Filipiak W, Zeidler M, Saunders T (2009) Generating transgenic mice from bacterial artificial chromosomes: transgenesis efficiency, integration and expression outcomes. Transgenic Research 18: 769-785.

83. Williams E AW, DeChiara TM, Gertsenstein M (2011) Combining ES Cells with Embryos. Advanced Protocols for Animal Transgenesis: 377-430.

84. Gelfand Y, Rodriguez A, Benson G (2007) TRDB, The Tandem Repeats Database. Nucleic Acids Research 35: D80-D87.

85. Stryke D, Kawamoto M, Huang CC, Johns SJ, King LA, et al. (2003) BayGenomics: a resource of insertional mutations in mouse embryonic stem cells. Nucleic Acids Research 31: 278-281.

86. Richardson L, Venkataraman S, Stevenson P, Yang Y, Burton N, et al. (2010) EMAGE mouse embryo spatial gene expression database: 2010 update. Nucleic Acids Research 38: D703-D709.

87. Krupp M, Marquardt JU, Sahin U, Galle PR, Castle J, et al. (2012) RNA-Seq Atlas, a reference database for gene expression profiling in normal tissue by next-generation sequencing. Bioinformatics 28: 1184-1185.

88. Westrick RJ, Mohlke KL, Korepta.L.M, Yang AY, Zhu G, et al. (2010) A spontaneous Irs1 passenger mutation linked to a gene targeted SerpinB2 Allele. Proceedings of the National Academy of Sciences of the United States of America 107: 16904-16909.

89. Khoriaty R, Vasievich MP, Ginsburg D (2012) The COPII pathway and hematologic disease. Blood 120: 31-38.

90. Zanetti G, Pahuja KB, Studer S, Shim S, Schekman R (2012) COPII and the regulation of protein sorting in mammals. Nat Cell Biol 14: 20-28.

91. Zhang B, Kaufman RJ, Ginsburg D (2005) LMAN1 and MCFD2 form a cargo receptor complex and interact with coagulation factor VIII in the early secretory pathway. Journal of Biological Chemistry 280: 25881-25886.

92. Nyfeler B, Michnick SW, Hauri HP (2005) Capturing protein interactions in the secretory pathway of living cells. Proceedings of the National Academy of Sciences of the United States of America 102: 6350-6355.

93. Pagano A, Letourneur F, Garcia-Estefania D, Carpentier JL, Orci L, et al. (1999) Sec24 proteins and sorting at the endoplasmic reticulum. J Biol Chem 274: 7833-7840.

94. Kawaguchi Y, Cooper B, Gannon M, Ray M, MacDonald RJ, et al. (2002) The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. Nat Genet 32: 128-134.

95. Madison BB, Dunbar L, Qiao XT, Braunstein K, Braunstein E, et al. (2002) cis Elements of the Villin Gene Control Expression in Restricted Domains of the Vertical (Crypt) and Horizontal (Duodenum, Cecum) Axes of the Intestine. Journal of Biological Chemistry 277: 33275-33283.

96. Postic C, Shiota M, Niswender KD, Jetton TL, Chen Y, et al. (1999) Dual Roles for Glucokinase in Glucose Homeostasis as Determined by Liver and Pancreatic Œ≤

Cell-specific Gene Knock-outs Using Cre Recombinase. Journal of Biological Chemistry 274: 305-315.

97. Boucher P, Gotthardt M, Li W-P, Anderson RGW, Herz J (2003) LRP: Role in Vascular Wall Integrity and Protection from Atherosclerosis. Science 300: 329-332.

98. Tallquist MD, Soriano P (2000) Epiblast-restricted Cre expression in MORE mice: A tool to distinguish embryonic vs. extra-embryonic gene function. genesis 26: 113-115.

99. Fahim AT, Wang H, Feng J, Ginsburg D (2009) Transgenic overexpression of a stable Plasminogen Activator Inhibitor-1 variant. Thrombosis Research 123: 785-792.

100. Alexopoulou AN (2008) The CMV early enhancer/chicken beta actin (CAG) promoter can be used to drive transgene expression during the differentiation of murine embryonic stem cells into vascular progenitors. BMC cell biology 9: 2.

101. Hitoshi N, Ken-ichi Y, Jun-ichi M (1991) Efficient selection for high-expression transfectants with a novel eukaryotic vector. Gene 108: 193-199.

102. Nagy A, Gertsenstein, M., Vintersten, K., Behringer, R. (2003) Manipulating the Mouse Embyro: A Laboratory Manual: CSH Press.

103. Frutkin AD, Shi H, Otsuka G, Lev√ ©en P, Karlsson S, et al. (2006) A critical developmental role for tgfbr2 in myogenic cell lineages is revealed in mice expressing SM22-Cre, not SMMHC-Cre. Journal of Molecular and Cellular Cardiology 41: 724-731.

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

105. Bianchi P, Fermo E, Vercellati C, Boschetti C, Barcellini W, et al. (2009) Congenital dyserythropoietic anemia type II (CDAII) is caused by mutations in the SEC23B gene. Human Mutation 30: 1292-1298.

106. Xu Q, Modrek B, Lee C (2002) Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. Nucleic Acids Research 30: 3754-3766.

107. Chen M, Manley JL (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat Rev Mol Cell Biol 10: 741-754.

108. Farmer SR (2008) Brown Fat and Skeletal Muscle: Unlikely Cousins? Cell 134: 726-727.

109. Yang XW, Gong S (2001) An Overview on the Generation of BAC Transgenic Mice for Neuroscience Research. Current Protocols in Neuroscience: John Wiley & Sons, Inc.

110. Barlowe C, Schekman R (1993) SEC12 encodes a guanine-nucleotide-exchange factor essential for transport vesicle budding from the ER. Nature 365: 347-349.

111. Bonnon C, Wendeler MW, Paccaud J-P, Hauri H-P (2010) Selective export of human GPI-anchored proteins from the endoplasmic reticulum. Journal of Cell Science 123: 1705-1715.

112. Osterwalder M, Galli A, Rosen B, Skarnes WC, Zeller R, et al. (2010) Dual RMCE for efficient re-engineering of mouse mutant alleles. Nat Meth 7: 893-895.

113. Osterwalder M, Lopez-Rios J, Zeller R (2010) Next generation engineering of conditional mouse alleles with loxP and FRT sites by dual RMCE.

114. Araki K, Araki M, Miyazaki J, Vassalli P (1995) Site-specific recombination of a transgene in fertilized eggs by transient expression of Cre recombinase. Proceedings of the National Academy of Sciences 92: 160-164.

115. Ohtsuka M, Ogiwara S, Miura H, Mizutani A, Warita T, et al. (2010) Pronuclear injection-based mouse targeted transgenesis for reproducible and highly efficient transgene expression. Nucleic Acids Research 38: e198-e198.

116. Ohtsuka M, Miura H, Nakaoka H, Kimura M, Sato M, et al. (2012) Targeted transgenesis through pronuclear injection of improved vectors into in vitro fertilized eggs. Transgenic Research 21: 225-226.

117. Ramirez-Solis R, Rivera-P,rez J, Wallace JD, Wims M, Zheng H, et al. (1992) Genomic DNA microextraction: A method to screen numerous samples. Analytical Biochemistry 201: 331-335.

118. Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, et al. (2013) One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering. Cell 154: 1370-1379.

119. Mashiko D, Fujihara Y, Satouh Y, Miyata H, Isotani A, et al. (2013) Generation of mutant mice by pronuclear injection of circular plasmid expressing Cas9 and single guided RNA. Sci Rep 3.

120. Zhou J, Shen B, Zhang W, Wang J, Yang J, et al. (2014) One-step generation of different immunodeficient mice with multiple gene modifications by CRISPR/Cas9 mediated genome engineering. The International Journal of Biochemistry & Cell Biology 46: 49-55.

121. Wu Y, Liang D, Wang Y, Bai M, Tang W, et al. (2013) Correction of a Genetic Disease in Mouse via Use of CRISPR-Cas9. Cell Stem Cell 13: 659-662.

122. Minami N, Suzuki T, Tsukamoto S (2007) Zygotic Gene Activation and Maternal Factors in Mammals. Journal of Reproduction and Development 53: 707-715.

123. Gaj T, Gersbach CA, Barbas CF (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends in biotechnology 31: 397-405.

124. Ong S-E, Mann M (2007) Stable Isotope Labeling by Amino Acids in Cell Culture for Quantitative Proteomics. In: Sechi S, editor. Quantitative Proteomics by Mass Spectrometry: Humana Press. pp. 37-52.