Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

SURO-2/TMEM39 Facilitates Collagen Secretion through the Formation of Large COPII Vesicles

Kiho Lee
Brown University
Jee Young Sung
National Cancer Center
Saerom Lee
National Cancer Center
Gaeun Lim
National Cancer Center
Kyung Jin Jung
Seoul National University Bundang Hospital
Jeong Min Chung
Kangwon National University
Hyun Suk Jung
Kangwon National University
Yong Yun Kim
National Cancer Center
Jaegal Shim (jaegal@ncc.re.kr)
National Cancer Center

Article

Keywords: Collagen, COPII, Fibrosis, SURO-2, TMEM39

Posted Date: July 16th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-35924/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Fibrosis of various tissues is a typical disease caused by excessive production and secretion of extracellular matrix. We used Caenorhabditis elegans to investigate the formation of large transport vesicles to understand collagen secretion, a critical factor in fibrosis formation. The suro-2 mutant displays obvious defects in collagen secretion and cuticle structure including a rupture phenotype in early adults. Transmission electron microscopy exhibited that the cuticle thickness of the suro-2 mutant was severely reduced. SURO-2/TMEM39 has 8 transmembrane domains and localizes in the endoplasmic reticulum (ER) membrane. SURO-2 interacts directly with NPP-20/Sec13, a component of the coat protein II (COPII) complex responsible for ER-to-Golgi transport. SURO-2 and NPP-20 localized at the same large puncta, a large COPII vesicle enough to accommodate collagens. We report here that SURO-2/TMEM39 is highly conserved among animal species and is a specialized regulator of bulky collagen secretion rather than general transport in C. elegans.

Introduction

Extracellular matrix (ECM) plays an essential role in animal physiology, including cell adhesion and migration and organ formation, but also in pathological conditions, such as connective tissue disorders, autoimmune diseases, muscular dystrophy, and infectious diseases [1, 2]. ECM is composed of approximately 300 proteins, including fibrous proteins and proteoglycans [2, 3], collagens are most abundant proteins of ECM. Collagens constitute up to 90% of ECM proteins and 30% of total proteins in the human body [4, 5]. In humans, 28 types of collagens have been identified, and more than 1300 mutations in collagen-associated genes are linked with genetic diseases [4, 6]. As with the ECM deficiency problem, excessive secretion and accumulation of ECM causes serious fibrotic diseases such as liver, kidney, heart and lung fibrosis [7, 8, 9]. Fibrosis that forms in solid cancer tissue is necessary or accelerated for cancer growth and metastasis [5]. In this regard, it is important to understand the regulation and secretion of collagen and other ECM proteins.

Biosynthesis and secretion of collagen proteins is a complex process with multiple extracellular and intracellular stages. Collagen biosynthesis includes transcription and post-translational regulation, translation regulation, post-translational modification, and extracellular crosslinking. The most common and primitive collagen has a long, uninterrupted collagen repeat of typical 1,000 amino acids and consists of many Gly-X-Y tripeptide repeats where X and Y are often proline and hydroxyproline [10].

Current research on collagen secretion focuses on the early secretory pathway from the endoplasmic reticulum (ER) to Golgi apparatus. In general, secreted ER cargo is exported through coat protein II complex (COPII) vesicles. Because collagen consists of rigid triple helices up to 400 nm in length, typical COPII vesicles with a diameter of 60–90 nm are not sufficient for collagen secretion [11]. In the yeast *Saccharomyces cerevisiae,* individual genes encode inner coatomer of COPII (Sec23-Sec24 heterodimer) and outer coatomer (Sec13-Sec31 heterotetramer). However, the vertebrate genome contains several

copies of COPII genes except for *Sec13* [12]. Collagen transport is known to rely on several gene products including Sec23A [13, 14], Sec24D [15, 16], and Sec13 [17]. Collagen secretion is also affected by the depletion of Sar1b, a small GTP-binding protein that initiates COPII coat assembly in the ER membrane [12, 18, 19]. The Sar1 cycle regulator Sedlin [20], and the Sec16-interacting protein TFG1 also contribute to collagen secretion [21].

Several studies have demonstrated potential mechanisms for the formation of large transport vesicles, including collagen packaging and COPII vesicle size increases. The single transmembrane domain protein TANGO1 plays an important role in procollagen VII transport and formation of large vesicles [22]. TANGO1 and cTAGE delay COPII vesicle formation until large enough to accommodate collagen packaging through direct binding of procollagen VII at the ER lumen and interaction with the Sec23-Sec24 heterodimer in the cytoplasm [22, 23, 24, 25]. Sec31 ubiquitination by the E3 ligase Cullin 3 and its specific adapter protein KLHL12 contributes to an increase in COPII coat size [26]. Recruitment and interaction of these proteins with COPII-related vesicles may be sufficient for the export of bulky collagen from ER [11]. A recent study reported that TMEM131, ranging from nematodes to humans, commonly plays an important role in regulating collagen secretion [27]. However, the understanding of the underlying mechanism of collagen secretion through evolutionarily large vesicles is still limited. For instance, simple animals like *Caenorhabditis elegans* do not have the TANGO1 homologue.

There are two main types of collagen in *C. elegans:* cuticle and basement collagen. Basement collagen contains two types of IV and one XVIII collagen, but cuticle collagen contains more than 170 cuticle collagen genes that support the rapid molting process associated with the synthesis and secretion of new cuticles at each stage of development [28, 29]. *C. elegans* cuticle mutants exhibit a broad range of phenotypes such as the lethal (Let), abnormal embryogenesis (Emb), dumpy (Dpy), long (Lon), blister (Bli), squat (Sqt) and roller (Rol) phenotypes [28, 29]. Genetic analysis of these cuticle mutants has provided valuable information regarding collagen function and ECM organization. We performed a suppressor screen using a transgenic strain of *C. elegans* with the Rol phenotype to identify genes involved in cuticle formation, processing, and secretion. In this study, we characterized *suro−2 (suppressor-of-rolling−2)* mutants in *C. elegans* and determined a cooperative role of SURO−2/TMEM39 and COPII in the formation of large ER vesicles essential for collagen secretion and cuticle formation.

Results

The *suro−2* mutant has serious defects in collagen secretion and cuticle structure

The *suro−2* mutant was isolated from a suppressor screen of *C. elegans* with the Rol phenotype. The *suro−2* mutation completely inhibits the Rol phenotype of the *jgIs4* expressing mutant ROL−6collagen derived from the *rol−6 (su1006)* mutant[30, 31, 32]. The *suro−2* mutant is a small and sick Dpy, and *suro−2 jgIs4* exhibits a severe Dpy phenotype compared to a single *suro−2* mutant (Fig. 1a).

Another interesting *suro–2* mutant phenotype is young adult lethality. Wild-type adults begin to die after laying eggs, usually 6–7 days of adult life, but retain their body shape. However, some *suro–2* mutants were ruptured (Fig. 1b) and approximately 18% of *suro–2* mutants died between 1–3 days (Fig. 1c). Since this burst phenotype may be related to cuticle abnormalities or disorders, we used electron microscopy to observe the cuticle structure in the *suro–2* mutant. The cuticle thickness of the adult *suro–2* mutant was reduced to nearly one third of the wild type (Fig. 1d). The average cuticle thickness was 604 nm in the wild type and decreased to 233 nm in the *suro–2* mutant. These results suggest that SURO–2 is required for proper collagen secretion and normal cuticle formation in *C. elegans.*

We further examined the effect of SURO–2 on the cuticle structure using two collagen markers, ROL–6::GFP [33] and COL–19::GFP [34]. These fusion proteins exhibit a striped pattern in the wild-type cuticle. Collagen-GFP expression was significantly reduced in the *suro–2* mutant and the wild-type stripe pattern disappeared from the cuticle (Fig. 1e). Therefore, SURO–2 is required for the secretion of collagen, and a rupture phenotype of the *suro–2* mutant may occur due to insufficient collagen supply.

SURO–2 is a homologue of mammalian transmembrane protein 39

We isolated two *suro–2* alleles, *jg35* and *jg92.* Sequencing of the D1007.5 using *suro–2 (jg35)* genomic DNA revealed a point mutation that resulted in the substitution of amino acid 187 with a stop codon (W187*). The wild-type *suro–2* gene putatively encodes 477 amino acids (Supplementary fig. 1), while the *suro–2 (jg35)* mutant encodes a shorter protein composed of only 186 amino acids whose *suro–2* mRNA may be degraded by nonsense RNA-mediated decay.

The *suro–2 (jg92)* mutantdisplays a mild Dpy phenotype compared with *suro–2 (jg35).* Genomic DNA sequencing of *suro–2 (jg92)* revealed an A>T mutation located in the noncoding region between *rps–10* and *suro–2 (*Fig. 2a). Because D1007.4, *suro–2,* and *rps–10* are putatively transcribed as a single operon, we compared the mRNA expression level of these genes using RT-PCR. We found that *suro–2* mRNA expression decreased significantly compared to the other two genes (Supplementary fig. 1e). Because knockdown of each gene using RNAi resulted in different phenotypes (Supplementary fig. 1b) and *suro–2 (jg92)* showed a mild phenotype compared with the putative null *suro–2 (jg35)* mutant, the *suro–2 (jg92)* mutation and associated reduction of *suro–2* mRNA expression are reliable. We speculate that *jg92* allele has a problem in trans-splicing after transcription. In fact, 70% of *C. elegans* genes are trans-spliced after transcription using SL1 or SL2 leader sequences [35]. Although recent studies revealed that conserved sequences are required for trans-splicing [36, 37] in *C. elegans,* we could not link the *suro–2 (jg92)* mutation to any of these sequences.

SURO–2 is a homologue of human transmembrane protein 39 (TMEM39) and contains eight transmembrane (TM) domains predicted by SMART (smart.embl-heidelberg.de) and Phobius (phobius.sbc.su.se). The *C. elegans* genome includes one *TMEM39* gene, but higher organisms have two

TMEM39 genes, *TMEM39A* and *TMEM39B.* The sequence of SURO–2 shares 27.5% and 25.9% identities with human TMEM39A and TMEM39B, respectively. The *C. elegans suro–2* gene expresses two isoforms resulting from alternative splicing, SURO–2A and SURO–2B (www.wormbase.org), which include 8 and 6 TM domains, respectively (Supplementary fig. 1f).

SURO–2 is a putative ER membrane protein

We generated GFP reporter constructs using the *suro–2* promoter to examine the *suro–2* expression pattern. Because the transgenic strain with a proximal upstream of *suro–2* coding region did not express GFP, we generated another transgenic strain with a distal promoter upstream of the *rps–10* coding region, and observed GFP expression in many tissues, such as the pharynx, hypodermis, body wall muscles, vulva, spermatheca, several neurons, and the intestine (Fig. 2b). We generated several GFP::SURO–2 fusion constructs using the Y37A1B.5 promoter (*Y37A1B.5p)* that enables to express gene products in the hypodermis [38]. Because we could not detect GFP expression from transgenic strains with terminal SURO–2::GFP fusions, we generated three additional GFP::SURO–2 constructs in which GFP was inserted between the cytoplasmic domains of SURO–2. The pJG720 plasmid displayed the best GFP expression, which showed fluorescence around the nuclei and relatively specific fluorescence in the cytoplasm (Supplementary fig. 2A)

We also generated several organelle markers by expressing GFP fusion proteins in the hypodermis to determine the exact subcellular localization of SURO–2. TRAM–1::GFP was used for the ER marker, and AMAN–2::GFP for the Golgi apparatus marker. TRAM–1 is the *C. elegans* orthologue of the translocating chain-associated membrane protein 1 [39], and AMAN–2 is an alpha-mannosidase II which functions at the Golgi apparatus [40]. The pJG735 plasmid is a tdTomato version of pJG720 and expresses tdTomato::SURO–2 in the hypodermis. Most tdTomato::SURO–2 proteins overlapped exactly with TRAM–1::GFP, but not with AMAN–2::GFP puncta (Fig. 2c). Because other ER markers such as GFP::KDEL and GFP::PISY–1 also overlapped with tdTomato::SURO–2 (Supplementary fig. 2B), we concluded that SURO–2 mainly localized to the ER membrane.

Because *suro–2* encodes two isoforms via alternative splicing, we examined the subcellular localization and functional differences of these two proteins. To examine the expression of SURO–2A and SURO–2B in the hypodermis, we generated a transgenic strain expressing tdTomato::SURO–2A and GFP::SURO–2B simultaneously. We found that these two proteins overlapped in most area (Fig. 2d), suggesting SURO–2A and SURO–2B function at the same ER membrane. Next, we performed rescue experiments using the *suro–2 jgIs4* strain to compare protein function involved in cuticle formation. We crossed *suro–2 jgIs4* with transgenic strains expressing GFP::SURO–2A or GFP::SURO–2B. The *suro–2 jgIs4* strain expressing the GFP::SURO–2A transgenic product recovered the Rol phenotype, whereas the *suro–2 jgIs4* strain expressing GFP::SURO–2B did not (Fig. 2e). Therefore, hypodermal SURO–2A expression is required and sufficient for proper cuticle formation.

SURO–2 is closely related to COPII in relation to collagen secretion

A large-scale yeast two-hybrid screen for protein networks and a genome-wide prediction of *C. elegans* genetic interactions proposed associations between D1007.5 (SURO–2) and nucleopore protein–20 (NPP–20) [41, 42], the *C. elegans* orthologue of human protein Sec13. Sec13 is a component of the COPII complex originally discovered in *S. cerevisiae* and participates in the early collagen secretory pathway [43, 44]. Another role of Sec13 relates to its function as a nucleopore protein during nuclear envelope reassembly [45] and is associated with the GATOR2 complex that regulates the mTORC1 complex [46]. Amidst varied functions, we focused on its potential role in the collagen secretory pathway to understand the relationship between SURO–2 and NPP–20, because SURO–2 is localized at the ER and participates in cuticle formation. We examined the knockdown phenotype of *npp–20* via RNAi to compare its phenotype with that of the *suro–2* mutant. Like *suro–2* mutants, *npp–20* RNAi suppressed the Rol phenotype of *jgIs4* (Fig. 3a), implying a close relationship between *suro–2* and *npp–20.* Therefore, we performed RNAi-mediated knockdowns of other COPII components to examine their functions in cuticle formation. Knockdown of most COPII genes resulted in Rol suppression. Knockdown of *sar–1* and *tfg–1* also resulted in Rol suppression (Fig. 3b). SAR–1 is the orthologue of human Sar1 GTPase which is the key regulator of COPII vesicle formation [19]. TFG–1 associates with SEC–16 and also functions in early collagen secretion [47]. Collectively, these results suggest that SURO–2 and COPII may function in the same secretory pathway.

Following Rol suppression by COPII knockdown, we determined collagen secretion was affected by COPII RNAi. We compared COL–19::GFP expression in the cytoplasm to identify potential deficiencies in collagen secretion when SURO–2 or NPP–20 was depleted by RNAi. COL–19::GFP was evenly distributed in the wild-type background, but COL–19::GFP accumulated in the cytoplasm by SURO–2 and NPP–20 knockdowns (Fig. 3c). We observed larger COL–19::GFP puncta in SURO–2-depleted worms than in NPP–20-depleted worms, implying that SURO–2 is more specialized in collagen secretion than NPP–20. Other components of COPII were depleted by RNAi and also resulted in COL–19::GFP accumulation in the cytoplasm (Fig. 3d). COPII plays a major role in early secretion of collagens from ER-to-Golgi apparatus [17, 23], and our experimental data supported this role as described in previous studies. Therefore, SURO–2 and COPII are closely related with respect to collagen transport and cuticle formation.

We validated the expression of COPII proteins if they colocalize in *C. elegans.* We generated transgenic strains expressing NPP–20::GFP and SEC–24.1::tdTomato simultaneously. Many NPP–20::GFP and SEC–24.1::tdTomato proteins were punctate and overlapped in the cytoplasm, whereas NPP–20::GFP also expresses in the nucleus (Fig. 3e).

SURO–2 interacts directly and is located together with NPP–20

SURO–2 and COPII are collectively involved in collagen secretion, and hypodermal expression is important for cuticle formation. We performed a glutathione S-transferase (GST) pulldown analysis to examine the direct interaction of SURO–2 and NPP–20. Because we were unable to express a SURO–2 protein including its TM domains in *E. coli,* we used its two long loop domains exposed to cytoplasm (Fig 4A). We found that NPP–20 and SURO–2 interacted directly through the 6th loop (L6) of SURO–2 using *in vitro* binding analysis (Fig. 4b).

We constructed 11 plasmids expressing series of YFP::SURO–2 deletion proteins to examine *in vivo* expression and function of SURO–2 domains (Supplementary fig. 3A). Most YFP::SURO–2 proteins exhibited similar expression patterns to intact protein excepting SURO–2C3, which has only C-terminus of SURO–2 and expressed like simple GFP expression in the cytoplasm and nucleus, while some proteins yielded strong punctate fluorescence. In particular, YFP::SURO–2 proteins having the L6 domain with the C-terminus formed large bright puncta (Supplementary fig. 3B). These results correlate with *in vitro* interactions of SURO–2 and NPP–20 and suggest the importance of the L6 domain of SURO–2 for *in vivo* complex formation.

We used the *jgIs47* strain which expresses tdTomato::SURO–2 and NPP–20::GFP simultaneously to examine *in vivo* colocalization of SURO–2 and NPP–20. In most cases, tdTomato::SURO–2 localized at the ER membrane as indicated by comparison of ER markers (Fig. 2b). Occasionally, the tdTomato::SURO–2 expressing strain displayed large fluorescent puncta, and their ER membrane-associated expression was relatively weak in the same cell. NPP–20::GFP appeared as variable-sized puncta. When tdTomato::SURO–2 proteins formed large fluorescent puncta, large NPP–20::GFP puncta overlapped with tdTomato::SURO–2. When tdTomato::SURO–2 spread throughout the ER membrane, NPP–20::GFP proteins were dispersed as smaller puncta (Fig. 4c). Similarly, SURO–2::GFP and SEC–24.1::tdTomato colocalized as large puncta, and small SEC–24.1::tdTomato puncta exist solitarily (Fig. 4d). To further clarify the relationship between SURO–2 and NPP–20, fluorescence was quantified by observing at high resolution using a confocal microscope. When NPP–20::GFP independently formed small puncta, the intensity of NPP–2::GFP and tdTomato::SURO–2 was low and fluorescence peaks did not overlap (Fig. 4e left panel). When NPP–20::GFP formed large puncta with tdTomato::SURO–2, the intensity of the two fluorescence was high and peaks overlapped (Fig. 4e right panel).

Super resolution images using Airyscan determined the identity of the puncta formed by NPP–20::GFP and tdTomato::SURO–2. The processed images of both small and large puncta appeared as circular forms. As a result, each punctum can be assumed to be a fluorescence image of the vesicle. The small puncta composed of NPP–20::GFP was around 100 nm in diameter (Fig. 4f left panel). Whereas, large puncta of NPP–20::GFP and tdTomato::SURO–2 were often larger than 400 nm in diameter (Fig. 4f right panel). This result implies that SURO–2 is required and facilitates large COPII vesicle formation.

Large COPII vesicles were actively produced during molting

We observed fluorescence over time from mid L4 larvae to investigate when large vesicles formed. In the mid L4 stage, tdTomato::SURO–2 expressed at ER membrane and NPP–20::GFP formed small puncta. From the late L4 stage, large puncta composed of tdTomato::SURO–2 and NPP–20::GFP began to form, and increased until young adults. The peak time of large vesicle formation represented by NPP–20::GFP puncta size and intensity is around 24 hours from the mid L4 stage. Passing the gravid adult stage, tdTomato::SURO–2 expression disappeared in the hypodermis and NPP–20::GFP expressed as small puncta (Fig. 5a). This observation implies that large vesicle formation is highest when new cuticles form, and most SURO–2 proteins in the hypodermis are consumed for making large vesicle in contrast to NPP–20 which existed as small vesicles in late adults.

Following strong interaction of SURO–2 and NPP–20, we examine the relationship of these proteins by depleting SURO–2 or NPP–20 using RNAi. SURO–2 RNAi resulted in reduction of NPP–20::GFP expression, both of intensity and large puncta formation in compared to control RNAi. When NPP–20 was depleted by RNAi, tdTomato::SURO–2 disappeared completely (Fig. 5b). However, small NPP–20::GFP puncta existed after SURO–2 depletion. These results suggest that SURO–2 is specialized for large vesicles rather than general COPII vesicles. The mutual stabilization and cooperation of SURO–2 and NPP–20 proteins seem to form large vesicles to facilitate collagen transport from the ER.

We also investigated if COPII regulators are required for large vesicle formation as their knockdown suppressed the Rol phenotype (Fig. 3b). We performed *sar–1* RNAi using the *jgIs56* strain expressing tdTomato::SURO–2 and NPP–20::GFP. To observe young adults, we transferred L4 larvae to *sar–1* RNAi plates because *sar–1* RNAi from early larvae resulted in larval lethality. After 24 hours from L4, *sar–1* knockdown resulted in significant reduction of large vesicles compared with control RNAi. The mean vesicle size and intensity of NPP–20::GFP decreased by SAR–1 depletion (Fig. 5c). From this result, we know that SURO–2 associated large vesicle formation is basically dependent on COPII and its regulator.

Cytoplasmic long loop domains are essential for SURO–2 expression and function

The domain study and *in vitro* binding assay of SURO–2 implied the importance of cytoplasmic long loop domains, L4 and L6. We generated transgenic strains expressing NPP–20::GFP together with L4 deletion form of tdTomato::SURO–2A (tdTomato::SURO–2ΔL4) or L6 deletion (tdTomato::SURO–2ΔL6) to prove critical roles of L4 or L6 domains *in vivo.* In the wild-type background, tdTomato::SURO–2ΔL4 forms large puncta with NPP–20::GFP superior to tdTomato::SURO–2A and tdTomato::SURO–2B. In contrast to tdTomato::SURO–2ΔL4, tdTomato::SURO–2ΔL6 could not form large vesicles (Supplementary fig. 4A). This result implies that L6 is required for stable SURO–2 expression and large vesicle formation when the endogenous wild-type SURO–2 exists. We wondered that L4 domain functions as a negative role for large vesicle formation on the contrary to L6. To examine this, we mated these transgenic strains with the *suro–2 (jg35)* mutant. The result is that tdTomato::SURO–2ΔL4 and tdTomato::SURO–2ΔL6 rarely expressed in the *suro–2* mutant background. Only tdTomato::SURO–2A

expressed well and rescued *suro–2,* but others including tdTomato::SURO–2B didn't rescue. Like these results, NPP–20::GFP expressed together with tdTomato::SURO–2ΔL4, tdTomato::SURO–2ΔL6 and tdTomato::SURO–2B did not form large vesicles well in the *suro–2* mutant background (Fig. 6a). This result means that the previous expression data were biased by endogenous SURO–2, and both L4 and L6 are essential for SURO–2 expression and large vesicle formation.

We further investigated to find out important regions of L4 and L6 domains by sequence alignment of SURO–2 and human TMEM39A and B. The posterior region of L4 domain and the anterior region of L6 domain are highly conserved compared with other regions. Seven out of ten amino acids from 232 to 241 and eight out of eleven amino acids from 318 to 328 are positive among three proteins (Fig. 6b). We generated 10 aa deletion (tdTomato::SURO–2Δ10) or 11 aa deletion (tdTomato::SURO–2Δ11) expression constructs to examine the importance of these conserved amino acids in L4 and L6 domains. In the wild-type background, tdTomato::SURO–2Δ10 expression was stronger than tdTomato::SURO–2A similar to tdTomato::SURO–2ΔL4 expression. tdTomato::SURO–2Δ11 and tdTomato::SURO–2Δ10Δ11 did not express well similar to tdTomato::SURO–2ΔL6 (Supplementary fig. 4B). In the *suro–2* mutant background, any tdTomato::SURO–2 deletion proteins did not express well. As expected, the size and intensity of NPP–20::GFP puncta when expressed together with tdTomato::SURO–2Δ10 or tdTomato::SURO–2Δ11 decreased significantly in the *suro–2* mutant (Fig. 6c).

SURO–2 is required for SURO–1/Carboxypeptidase A secretion

SURO–1 is a carboxypeptidase A (CPA) and is required for normal cuticle formation. Since the *suro–1* mutant exhibited a mild Dpy compared with the *suro–2* mutant that is a severe small Dpy, SURO–1 is one candidate cargo protein of the SURO–2 vesicle. In particular, SURO–1::DsRed is also secreted to the cuticle as a large vesicle [32]. To examine that SURO–2 is required for SURO–1 secretion, the *jgIs32* integration line expressing SURO–1::DsRed was crossed with the *suro–2* mutant. SURO–1::DsRed exhibited a typical cuticle pattern in wild type, but it was barely detected in the *suro–2* mutant excepting marginal regions of the cuticle. Whole fluorescence intensity of SURO–1::DsRed also decreased significantly in the *suro–2* mutant (Fig. 7a). Next, we generated a transgenic strain expressing SURO–1::GFP and tdTomato::SURO–2 to examine that SURO–1 is localized at the SURO–2 vesicle. Most SURO–1::GFP proteins were exactly localized at the same puncta of tdTomato::SURO–2 (Fig. 7b). These results indicate that one of the collagen-modifying enzymes, SURO–1/CPA, is secreted by large vesicles composed of SURO–2.

In conclusion, the results of this work are summarized in a simple model (Fig. 7c). Soluble proteins are transported from the ER by conventional COPII vesicles with an average diameter of 80 nm. On the other hand, many proteins involved in ECM should be transported from the ER using large vesicles. Both types of vesicles, large and small, contain COPII and commonly require regulators such as Sar1. The conserved

protein SURO–2/TMEM39 is necessary for the formation of this large vesicle and may function as the basic mechanism of bulky secretion, from simple animals to mammals.

Discussion

We have identified novel genes involved in collagen secretion and ECM remodeling using *C. elegans* to provide basic knowledges for solving fibrotic diseases. SURO–2/TMEM39, a *suro–2* gene product, is a highly conserved protein found in animals. The secretion level of collagen proteins was drastically reduced in the *suro–2* mutant, indicating SURO–2, rather than TANGO1/cTAGE5, may play a fundamental role in bulky collagen secretion like recently reported TMEM131 [27]. Since TANGO1/cTAGE5 have only been found in higher animals, the TANGO1-cTAGE system may have evolved to enable transport of diverse collagens. In fact, collagen VII, a main cargo of TANGO1-associated COPII vesicles, does not exist in *C. elegans.* However, we cannot dismiss the possible existence of protein other than SURO–2 that is directly involved in bulky collagen secretion because the *suro–2 (jg35)* mutant phenotype was less severe than that of a single essential collagen mutant. In addition, the cuticle still formed without SURO–2, indicating cuticle formation is not wholly depended on SURO–2 (Fig. 1d). Nevertheless, our data suggest that SURO–2 is a common key component for collagen secretion among diverse animal species.

Two TMEM39 genes exist in the human genome, as opposed to a single gene in *C. elegans*. TMEM39A seems to be closely associated with the human immune system, as several genome-wide association studies have reported relationships between TMEM39A polymorphisms and several autoimmune diseases, such as multiple sclerosis (MS) [48] and systemic lupus erythematosus [49]. For example, the first genome-wide association study of MS patients identified *KIF21B* and *TMEM39A* as MS susceptibility loci [48]. MS, which affects the central nervous system, is likely to be linked to the collagen secretion function of TMEM39A in the fact that it is clinically indistinguishable from collagen disorders [50, 51]. KIF21B is a kinesin motor protein that functions in Golgi-to-ER retrograde transport mediated by COPI vesicles [52]. Likewise, Arf1 which is a key regulator of COPI vesicle formation, seems to be required for proper cuticle formation, because knockdown of *arf–1,* the *C. elegans Arf1* orthologue, resulted in Rol suppression (our unpublished data). Defects in either TMEM39A or KIF21B could lead to perturbation of the ER-to-Golgi or Golgi-to-ER secretion of certain proteins involved in the immune system and MS. No functional study of TMEM39B has been published to date, although one previous report indicates TMEM39B is highly expressed in diffuse large B-cell lymphomas [53]. Thus, our results regarding the functionality of SURO–2/TMEM39 could provide preliminary information to guide studies of TMEM39A/B and their roles in human disease.

COPII and its associated gene products are required for collagen transport from the ER in vertebrates [12]. In *C. elegans,* SEC–23/Sec23 was reported to function in collagen secretion during embryogenesis [54]. Our RNAi experiments showed that most COPII genes suppressed the Rol phenotype and induced collagen accumulation in the cytoplasm as observed in *suro–2* knockdown. However, SURO–2 depletion exhibited

more obvious phenotypes than the COPII gene knockdown in collagen secretion (Fig. 3). Because COPII is essential for the export of general proteins including soluble and membrane proteins, COPII depletion resulted in more pleiotropic and detrimental phenotypes than SURO–2 was depleted. The *suro–2 (jg35)* mutant, which is expected to be a null mutation, is alive, while most of the COPII gene RNAi leads to embryonic death. Thus, SURO–2 may specifically facilitate bulky secretion of collagens and collagen processing enzymes, rather than of general proteins. However, SURO–2 may be involved in secretion of other proteins besides collagens, as *suro–2* was expressed throughout various cell types that do not produce collagens (Fig. 2b).

SURO–2 protein has eight putative TM domains and two long cytoplasmic loop domains. We determined that the L6 cytoplasmic domain directly binds NPP–20/Sec13. Due to its eight TM domains, SURO–2/TMEM39's movement might be more restricted than that of NPP–20/Sec13, a soluble cytoplasmic protein. Therefore, SURO–2 may recruit Sec13-Sec31 outer coatomers through direct interaction with Sec13 when a large vesicle is forming. SURO–2 and NPP–20 localization patterns also suggest that SURO–2 recruits NPP–20 to large collagen transport vesicles. Large puncta of NPP–20::GFP localization overlapped exactly with those of tdTomato::SURO–2, although smaller puncta of NPP–20::GFP did not (Fig. 4). Secretion of SURO–1, a CPA involved in collagen processing, was also dependent on SURO–2 function (Fig. 7a). SURO–1::DsRed proteins were secreted as large puncta in the cytoplasm and initially considered to be secretory granules because of their large size [32]. Taken together, SURO–1 function is interpreted as a subset of SURO–2 because SURO–1 and SURO–2 localized at the same large puncta (Fig. 7b), and the *suro–1* mutant has a much weaker phenotype than the *suro–2* mutant.

Cytoplasmic regions of SURO–2 are composed of 5 domains including N-terminus, L2, L4, L6 and C-terminus (Fig. 4a). We tried to clarify the function around L4 and L6 domains, because they are long enough to have roles compared with the short L2 domain. Both L4 and L6 domains are required for SURO–2 expression, and particularly, L6 domain binds NPP–20/Sec13 directly. Because large vesicles are composed of SURO–2 and COPII, NPP–20 interacting L6 domain has a critical role in collagen secretion. SURO–2 and human TMEM39 proteins have highly conserved regions in L4 and L6 domains. These conserved 10 amino acids in L4 and 11 amino acids in L6 are required for L4 and L6 function to express SURO–2 itself and form large vesicles. The greatest difference between L4 and L6 was that the L4 deletion form increased in the wild type but was not expressed well in the *suro–2* mutant. When endogenous SURO–2 and L4-deficient SURO–2 coexist, there appears to be a complex interaction between SURO–2 proteins.

Because the secondary protein structure and binding partner of SURO–2 are unlike TANGO1, a single-TM protein that interacts with Sec23-Sec24 inner coatomers [24, 41], mechanisms of collagen-packaging vesicle formation between the two proteins may differ. TANGO1 has a long tail and delays COPII vesicle formation by interacting with collagen VII directly [41]. However, both termini of SURO–2/TMEM39 are exposed to the cytoplasm, and only four short loop regions (L1, L3, L5 and L7) composed of less than 20 amino acids are located on the luminal side of the ER (Fig. 4a). We propose that SURO–2 directly incorporates into the COPII membrane and reduces its flexibility via SURO–2's eight TM domains and

four luminal loop domains. This complex of ER membrane and SURO–2 proteins may look like sewn buttons on fabric. The inflexible domains of the COPII membrane formed by SURO–2 may inhibit the formation of standard-sized COPII vesicles and instead induce spontaneous formation of a larger COPII-SURO–2 vesicle. This hypothesis requires additional investigation using biochemical, structural, and biophysical approaches, although we propose a model for early collagen secretion based on our data presented here (Fig. 7c). In mammals, HSP47, which functions as a collagen chaperone, is present in the ER lumen [55], and the presence of HSP47 orthologue has not been reported in nematodes. Recently reported TMEM131 recruits premature collagen monomers through the bacterial PapD chaperone-like domain and functions in collagen assembly and secretion. The C-terminal region of TMEM131 interacts with TRAPPC8, a component of the TRAPP tethering complex, and functions in collagen secretion [27]. In this regard, SURO–2/TMEM39 and TMEM–131/TMEM131 seem to cooperate in collagen secretion.

In summary, we elucidated novel functions of SURO–2/TMEM39 in *C. elegans,* thus providing valuable insights into bulky protein secretion, matrix biology, and TMEM39A-associated diseases including cancer fibrosis in higher animals.

Methods

C. elegans cultures and strains

The wild-type N2 strain and other mutant strains were fed on *E. coli (*OP50) growing in the standard nematode growth media (NGM) as described previously [56]. The *C. elegans* strains used in this study are *jgIs4, jgIs5, jgIs32, jgIs47, jgIs56,* and TP12 (*kaIs12)*.. The *jgIs4* strain expresses mutant ROL–6 collagen originated from *rol–6 (su1006),* and *jgIs5* expresses ROL–6::GFP. The *jgIs32* strain expresses SURO–1::DsRed, and *jgIs47* expresses both tdTomato::SURO–2 and NPP–20::GFP by the Y37A1B.5 promoter. The *jgIs56* strain expresses tdTomato::SURO–2 and NPP–20::GFP by the *col–12* promoter. These integration lines were generated by UV irradiation and outcrossed with N2 males four times.

Mapping and cloning of *suro–2* mutants

We used a SNP mapping method to identify *suro–2 (jg35)* mutation as previously described [56]. More than 1,000 F2 progeny that exhibited Dpy and non-Rol phenotypes were selected from the cross of *suro–2 (jg35) jgIs4* and Hawiian CB4856 male. Several F3 worms from each F2 plate were lysed, and their genomic DNA was used for PCR templates. PCR products were digested by restriction enzymes specific to each SNP and analyzed by 2% agarose gel electrophoresis. The RNAi clones for 12 candidate genes in the mapped region were originated from the Ahringer library [57]. Several L4-stage *jgIs4* worms were transferred to each RNAi plate, and F1 progeny were observed. The WRM067cA7 fosmid and 8 kb PCR products, including three genes in the same operon, were used for the rescue experiment. We generated transgenic worms and mated each transgenic line with *suro–2* mutants because direct microinjection ruptured *suro–2* mutants. Finally, PCR-amplified D1007.5 from the *suro–2* mutant was sequenced to

detect molecular resin (Cosmogenetech, Seoul, South Korea). Whole-genome sequencing of *suro−2 (jg92)* was performed by Macrogen (Seoul, South Korea).

Fluorescent protein fusion constructs and microinjection

We used GFP-encoded pPD95.77 and pPD117.01 Fire vectors and generated genes encoding additional fluorescent proteins (FP), such as YFP, DsRed, and tdTomato, via PCR for incorporation into Clontech vectors, such as pEYFP. For hypodermal expression, we used the Y37A1B.5 promoter or *col−12* promoter, and pJG452 (*Y37A1B.5p::*::GFP) was constructed from *Y37A1B.5p* and pPD95.77 using *Hind*III and *Pst*I. Primers used for amplification of the Y37A1B.5 promoter region were Y37A1B.5−03 (5 TAT *AAG CTT* AGC GAG TGA CAT TTG CCG 3) and Y37A1B.5−02 (5 TAT *CTG CAG* TTT GGT TTT TGG GAT TTT TGA TCT GC 3). We observed strong GFP expression in the hypodermis of the transgenic strain carrying pJG452. The final concentration of microinjection mixture was 150 μg/ml, including 75 μg/ml pRF4 (Rol) or 5 μg/ml pCFJ90 (*myo−2p::*::mCherry), which were used as injection markers. SURO−2-or COPII-FP fusion plasmids were injected at 20−50 μg/ml concentration, and the pBluescript SK+ plasmid was added to adjust the final DNA concentration when necessary.

Plasmids used in this study were pJG405 (*suro−2p::*::NLS::GFP), pJG476 (*Y37A1B.5p::*::GFP::SURO−2L4), pJG477 (*Y37A1B.5p::*::SURO−2N::GFP::SURO−2C), pJG493 (*Y37A1B.5p::*::NPP−20::GFP), pJG496 (*Y37A1B.5p::*::NPP−20::DsRed), pJG519 (*Y37A1B.5p::*::AMAN−2::GFP), pJG527 (*Y37A1B.5p::*::GFP::TRAM−1), pJG536 (Y37A1B.5*p::*::GFP::PISY−1), pJG601 (*sec−24.1p::*::NLS::GFP), pJG602 (*sec−24.2p::*::NLS::GFP), pJG663 (*Y37A1B.5p::*::GFP::KDEL), pJG683 (*Y37A1B.5p::*::SEC−24.1::tdTomato), pJG719 (*Y37A1B.5p::*::SP::GFP), pJG720 (*Y37A1B.5p::*::GFP::SURO−2A), pJG735 (*Y37A1B.5p::*::tdTomato::SURO−2A), pJG789 (*Y37A1B.5p::*::GFP::SURO−2B), pJG819 (*Y37A1B.5p::*::YFP::SURO−2), pJG1079 (*Y37A1B.5p::*::SURO−2N::GFP::SURO−2C), pJG1090 (*Y37A1B.5p::*::tdTomato::SURO−2L4), pJG1111 (*Y37A1B.5p::*::YFP::SURO−2L4), pJG1114 (*Y37A1B.5p::*::YFP::SURO−2L6). pJG1115 (*Y37A1B.5p::*::YFP::SURO−2(SP)), pJG1117 (*suro−2p::*::GFP), and pJG1124 (*Y37A1B.5p::*::YFP::SURO−2AC).

RNA interference

Most RNAi clones that were originated from the Ahringer library [57] were kindly provided by Dr. J. Lee (Seoul National University). Additionally, we constructed RNAi clones for *suro−2* using its cDNA and L4440 vector. The 1.4 kb *suro−2* cDNA was amplified by PCR using the wild-type N2 cDNA as templates. The PCR fragments were cloned into L4440 vector using *Sal*I and *BamH*I sites. Primers used for *suro−2* RNAi clone were D1007.5−03 (5AT*G TCG AC*A TGC CGC CTC GAA GAC G 3) and D1007.5−23 (5 T*GG ATC C*TT ATC GTT CTT GTT GAA GTT GAT GC 3). We transferred different staged worms for each gene RNAi according to the extent of RNAi effect. L4 stage worms were transferred for control, *sec−31* and *suro−2* RNAi, and gravid adult worms for *npp−20* and *sec−16* RNAi. F1 progeny were observed in these cases. However, the effect of *sec−12, sec−23, sec−24.1, sar−1* and *tfg−1* RNAi was very severe, and L1-

L2 larvae were transferred and observed when they grew up. The empty vector L4440 was used for control.

GST pull-down and western blot analyses

The pGEX4T1 GST vector system (GE helthcare, Wauwatosa, WI, USA) or GST-SURO–2 and the pRSET vector system (Thermo-Fisher, Waltham, MA, USA) were used for 6×His-NPP–20 protein expression in *E. coli* (BL21). The soluble fraction of the bacterial lysate was prepared by sonication and centrifugation. The soluble fraction containing GST or GST-SURO–2 was mixed with glutathione-agarose beads (G–4510; Sigma, St. Louis, MO, USA) and washed five times with sonication buffer (100 mM NaCl, 25 mM Tris-HCl (pH8.0), 10% glycerol). The bacterial lysate containing 6×His-NPP–20 was mixed with G-agarose beads bound to GST or GST-SURO–2 and incubated overnight at 4°C. Each G-agarose bead suspension was then washed five times with sonication buffer and boiled for 2 min after SDS-PAGE sample buffer was added. Each sample was loaded and electrophoresed on the 12% SDS-PAGE gel. Rabbit anti-NPP–20 antibodies, generated by immunizing synthetic NPP–20 peptides composed of 14 amino acids (IKVPRDNKEREGMS) in C-terminal region of NPP–20 (Peptron, Daejeon, South Korea), were used to detect 6×His-NPP–20. Anti-NPP–20 serum was purified by affinity column chromatography (SulfoLink Immobilization kit; Thermo-Fisher) according to the supplier's instruction.

Transmission electron microscopy

The wild-type N2 and *suro–2 (jg35)* mutant were fixed with 2.5% glutaraldehyde in 0.1 M sodium phosphate buffer (PB, pH 7.4) overnight at room temperature. The samples were then post-fixed in 1% OsO_4 and 1.5% potassium ferrocyanide for 1 h at room temperature and stained by 0.5% uranyl acetate overnight at 4°C and dehydrated in the graded ethanol series. After polymerization in Epon for 2 days at 60°C, the glass coverslips were detached from epon block. Samples were sectioned by EM UC7 (Leica, Wetzlar, Germany) ultramicrotome to 70 nm thickness and contrasted with 0.5% uranyl acetate and lead citrate. The specimens were examined in Tecnai G2 Spirit Twin transmission electron microscopy (FEI, Hillsboro, OR, US; Instrumentation was used in Korea Basic Science Institute) operated at 120 kV. Images were recorded on 4Kx4K Ultrascan 895 CCD (Gatan, Pleasanton, CA, USA).

Fluorescence microscopy and statistical analysis

Fluorescence microscopy was performed using an Axio Imager A2 compound and LSM880 confocal microscopes (Zeiss, Oberkochen, Germany) and images were processed using Axiovision Rel.4.5 software (Zeiss), Photoshop (Adobe systems, San Jose, CA, USA) and Image J (https://imagej.nih.gov/ij). Worms were prepared with M9 buffer containing 2.5 mM levamisole. Statistical analyses including *t*-test in this study were carried out with SigmaPlot software (Systat software, San Jose, CA, USA). *P* values of <0.05, <0.01, and <0.001, indicated statistical significance and are denoted with *, **, or ***, respectively.

References

1.Bateman JF, Boot-Handford RP, Lamande SR. Genetic diseases of connective tissues: cellular and extracellular effects of ECM mutations. *Nature reviews Genetics 10,* 173–183 (2009).

2.Mouw JK, Ou G, Weaver VM. Extracellular matrix assembly: a multiscale deconstruction. *Nature reviews Molecular cell biology 15,* 771–785 (2014).

3.Gillet LC*, et al.* Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Molecular & cellular proteomics: MCP 11,* O111 016717 (2012).

4.Ricard-Blum S. The collagen family. *Cold Spring Harbor perspectives in biology 3,* a004978 (2011).

5.Gilkes DM, Semenza GL, Wirtz D. Hypoxia and the extracellular matrix: drivers of tumour metastasis. *Nature reviews Cancer 14,* 430–439 (2014).

6.Myllyharju J, Kivirikko KI. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends in genetics: TIG 20,* 33–43 (2004).

7.Bataller R, Brenner DA. Liver fibrosis. *J Clin Invest 115,* 209–218 (2005).

8.Humphreys BD. Mechanisms of Renal Fibrosis. *Annu Rev Physiol 80,* 309–326 (2018).

9.Meyer KC. Pulmonary fibrosis, part I: epidemiology, pathogenesis, and diagnosis. *Expert Rev Respir Med 11,* 343–359 (2017).

10.Hynes RO. The evolution of metazoan extracellular matrix. *The Journal of cell biology 196,* 671–679 (2012).

11.Malhotra V, Erlmann P. The pathway of collagen secretion. *Annual review of cell and developmental biology 31,* 109–124 (2015).

12.Unlu G, Levic DS, Melville DB, Knapik EW. Trafficking mechanisms of extracellular matrix macromolecules: insights from vertebrate development and human diseases. *The international journal of biochemistry & cell biology 47,* 57–67 (2014).

13.Boyadjiev SA*, et al.* Cranio-lenticulo-sutural dysplasia is caused by a SEC23A mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking. *Nature genetics 38,* 1192–1197 (2006).

14.Lang MR, Lapierre LA, Frotscher M, Goldenring JR, Knapik EW. Secretory COPII coat component Sec23a is essential for craniofacial chondrocyte maturation. *Nature genetics 38,* 1198–1203 (2006).

15.Sarmah S, Barrallo-Gimeno A, Melville DB, Topczewski J, Solnica-Krezel L, Knapik EW. Sec24D-dependent transport of extracellular matrix proteins is required for zebrafish skeletal morphogenesis. *PloS one 5,* e10367 (2010).

16.Ohisa S, Inohaya K, Takano Y, Kudo A. 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 (2010).

17.Townley AK*, et al.* Efficient coupling of Sec23-Sec24 to Sec13-Sec31 drives COPII-dependent collagen secretion and is essential for normal craniofacial development. *Journal of cell science 121,* 3025–3034 (2008).

18.Dannoura AH*, et al.* Anderson's disease: exclusion of apolipoprotein and intracellular lipid transport genes. *Arteriosclerosis, thrombosis, and vascular biology 19,* 2494–2508 (1999).

19.Aridor M*, et al.* The Sar1 GTPase coordinates biosynthetic cargo selection with endoplasmic reticulum export site assembly. *The Journal of cell biology 152,* 213–229 (2001).

20.Venditti R*, et al.* Sedlin controls the ER export of procollagen by regulating the Sar1 cycle. *Science 337,* 1668–1672 (2012).

21.McCaughey J*, et al.* TFG Promotes Organization of Transitional ER and Efficient Collagen Secretion. *Cell reports 15,* 1648–1659 (2016).

22.Saito K*, et al.* TANGO1 facilitates cargo loading at endoplasmic reticulum exit sites. *Cell 136,* 891–902 (2009).

23.Malhotra V, Erlmann P. Protein export at the ER: loading big collagens into COPII carriers. *The EMBO journal 30,* 3475–3480 (2011).

24.Saito K*, et al.* cTAGE5 mediates collagen secretion through interaction with TANGO1 at endoplasmic reticulum exit sites. *Molecular biology of the cell 22,* 2301–2308 (2011).

25.Tanabe T, Maeda M, Saito K, Katada T. Dual function of cTAGE5 in collagen export from the endoplasmic reticulum. *Molecular biology of the cell 27,* 2008–2013 (2016).

26.Jin L*, et al.* Ubiquitin-dependent regulation of COPII coat size and function. *Nature 482,* 495–500 (2012).

27.Zhang Z*, et al.* Broadly conserved roles of TMEM131 family proteins in intracellular collagen assembly and secretory cargo trafficking. *Sci Adv 6,* eaay7667 (2020).

28.Taffoni C, Pujol N. Mechanisms of innate immunity in *C. elegans* epidermis. *Tissue barriers 3,* e1078432 (2015).

29.Page AP, Johnstone IL. The cuticle. *WormBook: the online review of C elegans biology,* 1–15 (2007).

30.Kramer JM, French RP, Park EC, Johnson JJ. The *Caenorhabditis elegans rol–6* gene, which interacts with the *sqt–1* collagen gene to determine organismal morphology, encodes a collagen. *Molecular and*

cellular biology 10, 2081–2089 (1990).

31.Mello CC, Kramer JM, Stinchcomb D, Ambros V. Efficient gene transfer in *C.elegans:* extrachromosomal maintenance and integration of transforming sequences. *The EMBO journal 10,* 3959–3970 (1991).

32.Kim TH, Kim YJ, Cho JW, Shim J. A novel zinc-carboxypeptidase SURO–1 regulates cuticle formation and body morphogenesis in *Caenorhabditis elegans. FEBS letters 585,* 121–127 (2011).

33.Kim TH, Kim DH, Nam HW, Park SY, Shim J, Cho JW. Tyrosylprotein sulfotransferase regulates collagen secretion in *Caenorhabditis elegans. Molecules and cells 29,* 413–418 (2010).

34.Thein MC, McCormack G, Winter AD, Johnstone IL, Shoemaker CB, Page AP. *Caenorhabditis elegans* exoskeleton collagen COL–19: an adult-specific marker for collagen modification and assembly, and the analysis of organismal morphology. *Developmental dynamics: an official publication of the American Association of Anatomists 226,* 523–539 (2003).

35.Blumenthal T. Trans-splicing and operons. *WormBook: the online review of C elegans biology,* 1–9 (2005).

36.Allen MA, Hillier LW, Waterston RH, Blumenthal T. A global analysis of *C. elegans* trans-splicing. *Genome research 21,* 255–264 (2011).

37.Ma X*, et al.* Analysis of *C. elegans* muscle transcriptome using trans-splicing-based RNA tagging (SRT). *Nucleic acids research,* (2016).

38.McKay SJ*, et al.* Gene expression profiling of cells, tissues, and developmental stages of the nematode *C. elegans. Cold Spring Harbor symposia on quantitative biology 68,* 159–169 (2003).

39.Rolls MM, Hall DH, Victor M, Stelzer EH, Rapoport TA. Targeting of rough endoplasmic reticulum membrane proteins and ribosomes in invertebrate neurons. *Molecular biology of the cell 13,* 1778–1791 (2002).

40.Paschinger K*, et al.* A deletion in the golgi alpha-mannosidase II gene of *Caenorhabditis elegans* results in unexpected non-wild-type N-glycan structures. *The Journal of biological chemistry 281,* 28265–28277 (2006).

41.Li S*, et al.* A map of the interactome network of the metazoan *C. elegans. Science 303,* 540–543 (2004).

42.Zhong W, Sternberg PW. Genome-wide prediction of *C. elegans* genetic interactions. *Science 311,* 1481–1484 (2006).

43.Kaiser CA, Schekman R. Distinct sets of SEC genes govern transport vesicle formation and fusion early in the secretory pathway. *Cell 61,* 723–733 (1990).

44.Schekman R. Lasker Basic Medical Research Award. SEC mutants and the secretory apparatus. *Nature medicine 8,* 1055–1058 (2002).

45.Vasu SK, Forbes DJ. Nuclear pores and nuclear assembly. *Current opinion in cell biology 13,* 363–375 (2001).

46.Bar-Peled L*, et al.* A Tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. *Science 340,* 1100–1106 (2013).

47.Witte K*, et al.* TFG–1 function in protein secretion and oncogenesis. *Nature cell biology 13,* 550–558 (2011).

48.Comprehensive follow-up of the first genome-wide association study of multiple sclerosis identifies KIF21B and TMEM39A as susceptibility loci. *Hum Mol Genet 19,* 953–962 (2010).

49.Lessard CJ*, et al.* Identification of IRF8, TMEM39A, and IKZF3-ZPBP2 as susceptibility loci for systemic lupus erythematosus in a large-scale multiracial replication study. *American journal of human genetics 90,* 648–660 (2012).

50.Fulford KW, Catterall RD, Delhanty JJ, Doniach D, Kremer M. A collagen disorder of the nervous system presenting as multiple sclerosis. *Brain 95,* 373–386 (1972).

51.Kurne A*, et al.* A clinically isolated syndrome: a challenging entity: multiple sclerosis or collagen tissue disorders: clues for differentiation. *J Neurol 255,* 1625–1635 (2008).

52.Marszalek JR, Weiner JA, Farlow SJ, Chun J, Goldstein LS. Novel dendritic kinesin sorting identified by different process targeting of two related kinesins: KIF21A and KIF21B. *The Journal of cell biology 145,* 469–479 (1999).

53.Kim SJ*, et al.* Gene expression profiles for the prediction of progression-free survival in diffuse large B cell lymphoma: results of a DASL assay. *Annals of hematology 93,* 437–447 (2014).

54.Roberts B, Clucas C, Johnstone IL. Loss of SEC–23 in *Caenorhabditis elegans* causes defects in oogenesis, morphogenesis, and extracellular matrix secretion. *Molecular biology of the cell 14,* 4414–4426 (2003).

55.Ferreira LR, Norris K, Smith T, Hebert C, Sauk JJ. Hsp47 and other ER-resident molecular chaperones form heterocomplexes with each other and with collagen type IV chains. *Connect Tissue Res 33,* 265–273 (1996).

56.Brenner S. The genetics of *Caenorhabditis elegans. Genetics 77,* 71–94 (1974).

57.Ashrafi K, *et al.* Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes. *Nature 421,* 268–272 (2003).

Declarations

Acknowledgements

We appreciate Dr. Junho Lee and his laboratory members for providing RNAi clones, and Dr. Donghoon Lee for providing useful comments. Nematode strains such as TP12 were provided by the *Caenorhabditis* Genetics Center, which is funded by the NIH National Center for Research Resources. This work was supported in part by grants from the National Cancer Center of Korea (NCC-1610290, NCC-1810932) and the National Research Foundation (NRF-1531940). EM study was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT and Future Planning (NRF-2015R1C1A1A01053611).

Author information

These authors contributed equally: Kiho Lee and Jee Young Sung

Affiliations

Contributions

K.L., J.Y.S., Y.K., H.S.J. and J.S. designed research; K.L., J.Y.S., S.L., G.L., K.J.J, J.M.C. and J.S. performed experiments; Y.K., H.S.J. and J.S. analyzed data and wrote the paper.

Corresponding authors

Yong-Nyun Kim. Research Institute, National Cancer Center, 323 Ilsan-ro, Goyang-si, Gyeonggi-do, 10408, Korea. Tel: 82-31-920-2415. E-mail: ynk@ncc.re.kr

Jaegal Shim. Research Institute, National Cancer Center, 323 Ilsan-ro, Goyang-si, Gyeonggi-do, 10408, Korea. Tel: 82-31-920-2262. E-mail: jaegal@ncc.re.kr

Ethics declarations

Competing interests

The authors declare no competing interests

Figures

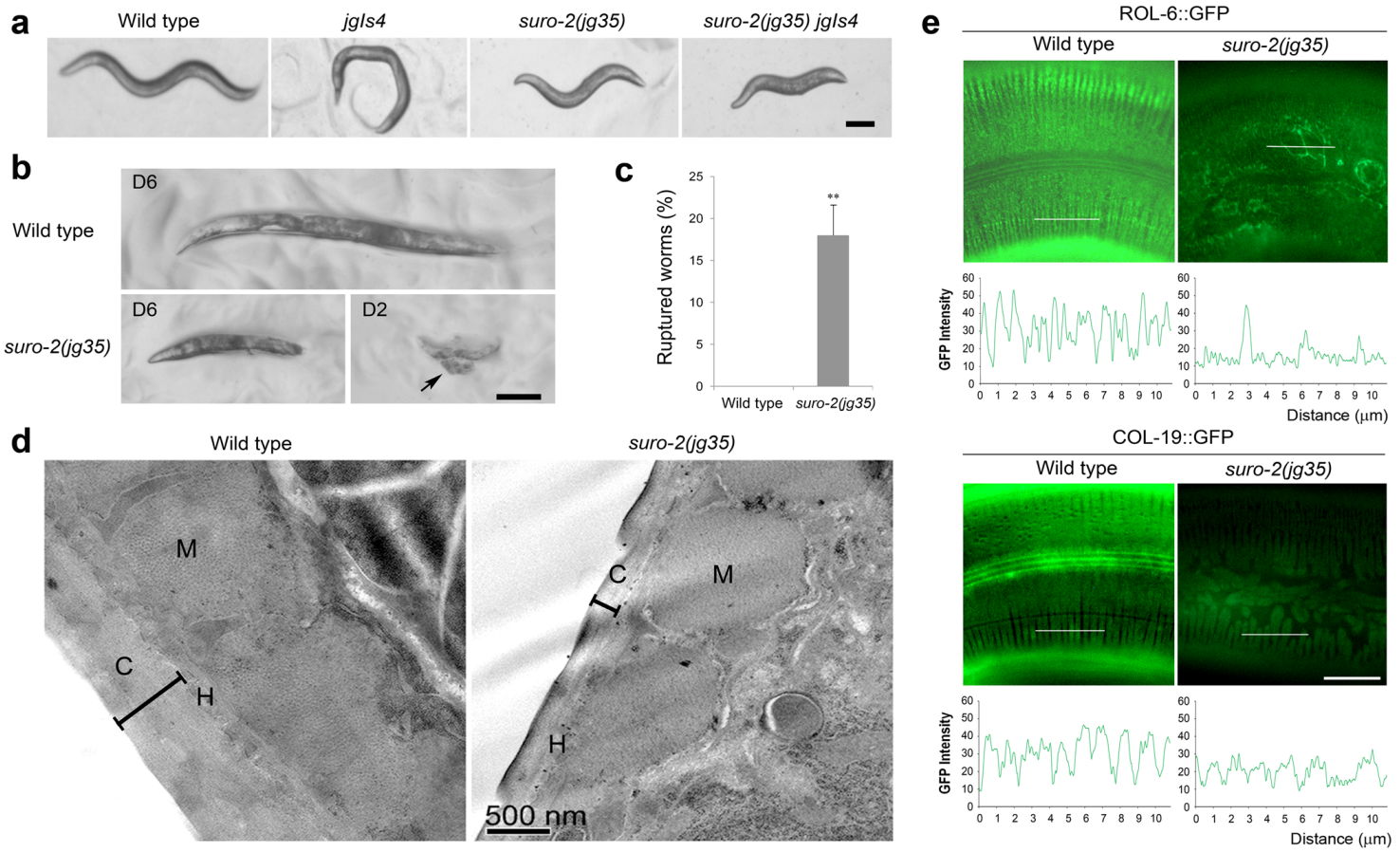

Figure 1

Cuticle malformation of suro-2 mutants. a Dpy and Rol suppression phenotype of suro-2 mutants. The suro-2 jgIs4 strain is more severe than a single suro-2 mutant. Scale bar= 100 μm. b Dead worms of wild type and suro-2 mutants (D6= adult day 6). The suro-2 mutant worms often ruptured at the young adult stage (D2= adult day 2). The ruptured suro-2 mutant released its contents including the intestines and gonads (arrow). Scale bar= 150 μm. c Approximately 18% of suro-2 mutants ruptured and died within 3 days of becoming adults. Error bar= a standard error. d Electron microscopy of adult cuticle. Twenty adult animals of wild type or suro-2 mutants were analyzed. The average cuticle thickness in wild type and suro-2 mutants was 604 ± 23.2 nm and 233 ± 18.7 nm, respectively. Black H lines indicate cuticle thicknesses. C= cuticle, M= body wall muscle, and H= hypodermis. ± sign is standard deviation. Scale bar= 500 nm. e Expression of two collagen markers, ROL-6::GFP and COL-9::GFP in adult cuticle. Scale bar= 10 μm. The lower panel quantifies GFP fluorescence along thin white lines in each figure to represent ROL-6::GFP and COL-19::GFP as wild type and suro-2 mutants. In the suro-2 mutant, ROL-6::GFP decreased from 31.6 ± 9.93 to 15.9 ± 5.95, and COL-19 :: GFP decreased from 30.8 ± 8.98 to 19.9 ± 5.23 compared to the wild type. ± sign is standard error.

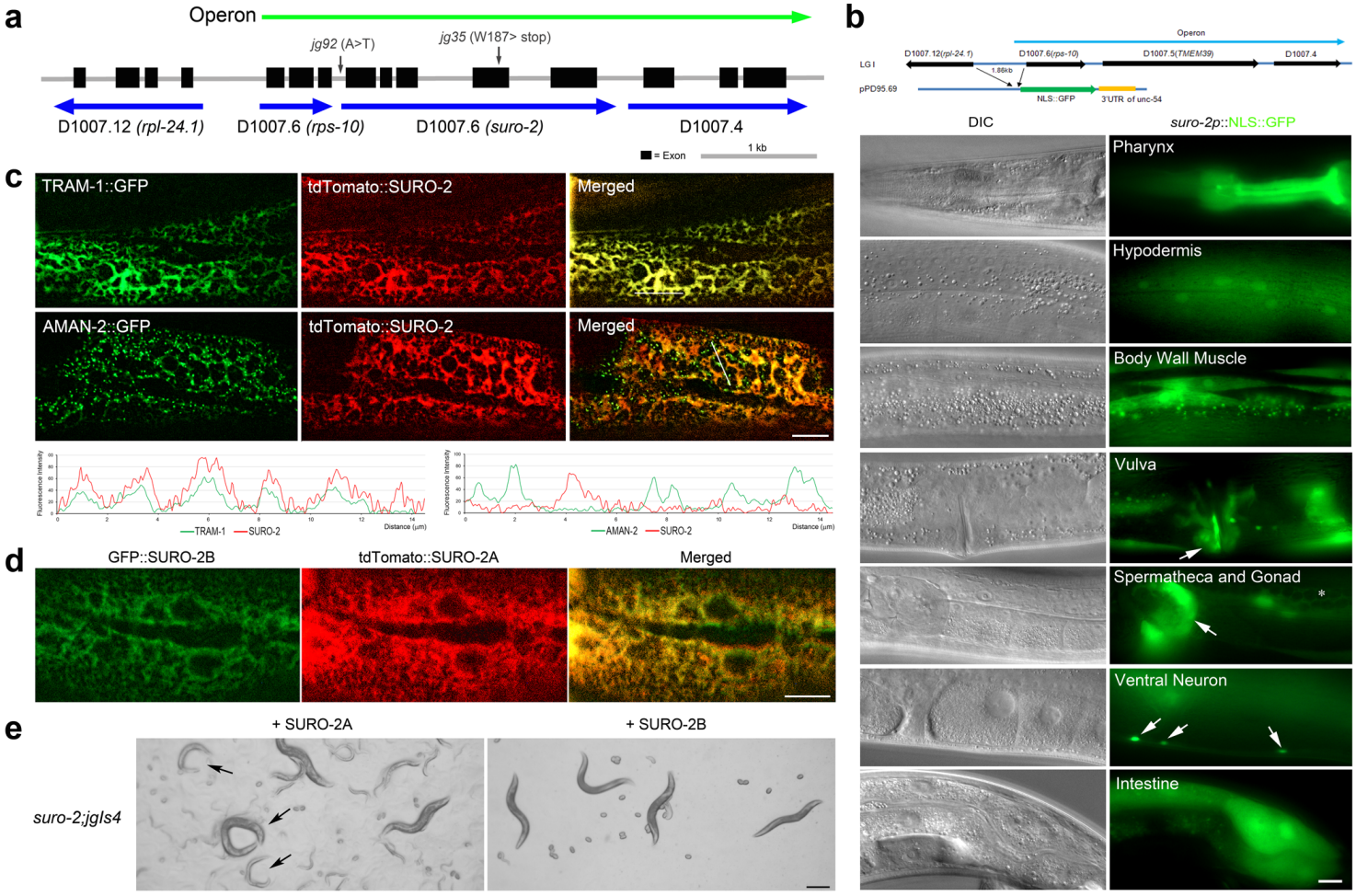

Figure 2

Genomic structure and expression pattern of suro-2. a Gene and genomic structures around suro-2 (D1007.5). The two alleles of the suro-2 mutant are indicated by black arrows. Blue arrows indicate the direction of transcription. b GFP expressions by the distal promoter of suro-2. The 1.86-kb promoter of rps-10 drove expression of GFP in many tissues and organs, such as the pharynx, hypodermis, body wall muscle, vulva (arrow), spermatheca (arrow), oocytes (*), some neurons (arrows), and the intestine. Scale bar= 10 μm. c Subcellular localization of SURO-2. tdTomato::SURO-2 and TRAM-1::GFP (ER marker) or AMAN-2::GFP (Golgi marker) proteins were expressed in the hypodermis by the Y37A1B.5 promoter. Most tdTomato::SURO-2 and TRAM-1::GFP proteins are colocalized. Scale bar= 10 μm. The bottom panel quantifies fluorescence along the thin white line in each figure. d Colocalization of tdTomato::SURO-2A and GFP::SURO-2B in the hypodermis. Most tdTomato::SURO-2A and GFP::SURO-2B are in the same location. Scale bar= 10 μm. e Rescue experiment of suro-2 (jg92) jgIs4 using GFP::SURO-2A or GFP::SURO-2B. Arrows indicate rolling worms rescued by GFP::SURO-2A. Scale bar= 150 μm.

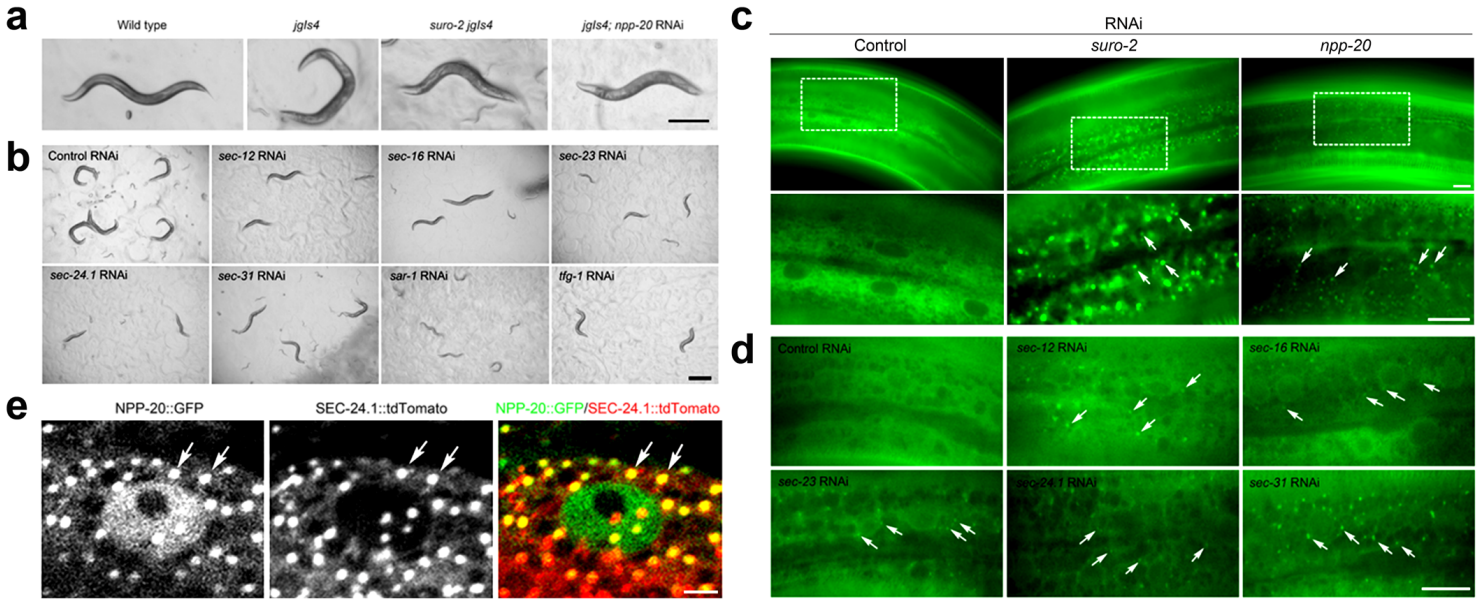

Figure 3

Inhibition of collagen secretion by knockdown of COPII gene. a RoI suppression phenotype after npp-20 knockdown. Scale bar= 200 μm. b RoI suppression phenotype by knockdown of COPII components and related genes using RNAi. Scale bar= 200 μm. c Accumulation of COL-19::GFP in the cytoplasm after knockdown of suro-2 or npp-20. White boxes of the upper panel are enlarged in the bottom panel. Arrows indicate COL-19::GFP puncta. Scale bars= 10 μm. d Accumulation of COL-19::GFP in the cytoplasm after knockdown of COPII genes. Arrows indicate COL-19::GFP puncta. Scale bar= 10 μm. e Confocal microscopy for NPP-20::GFP and SEC-24.1::tdTomato. These COPII components were expressed as the same punctuated pattern in the cytoplasm. In the cytoplasm, most of the NPP-20::GFP and SEC-24.1::tdTomato proteins were colocalized (arrows). Scale bar= 2 μm.

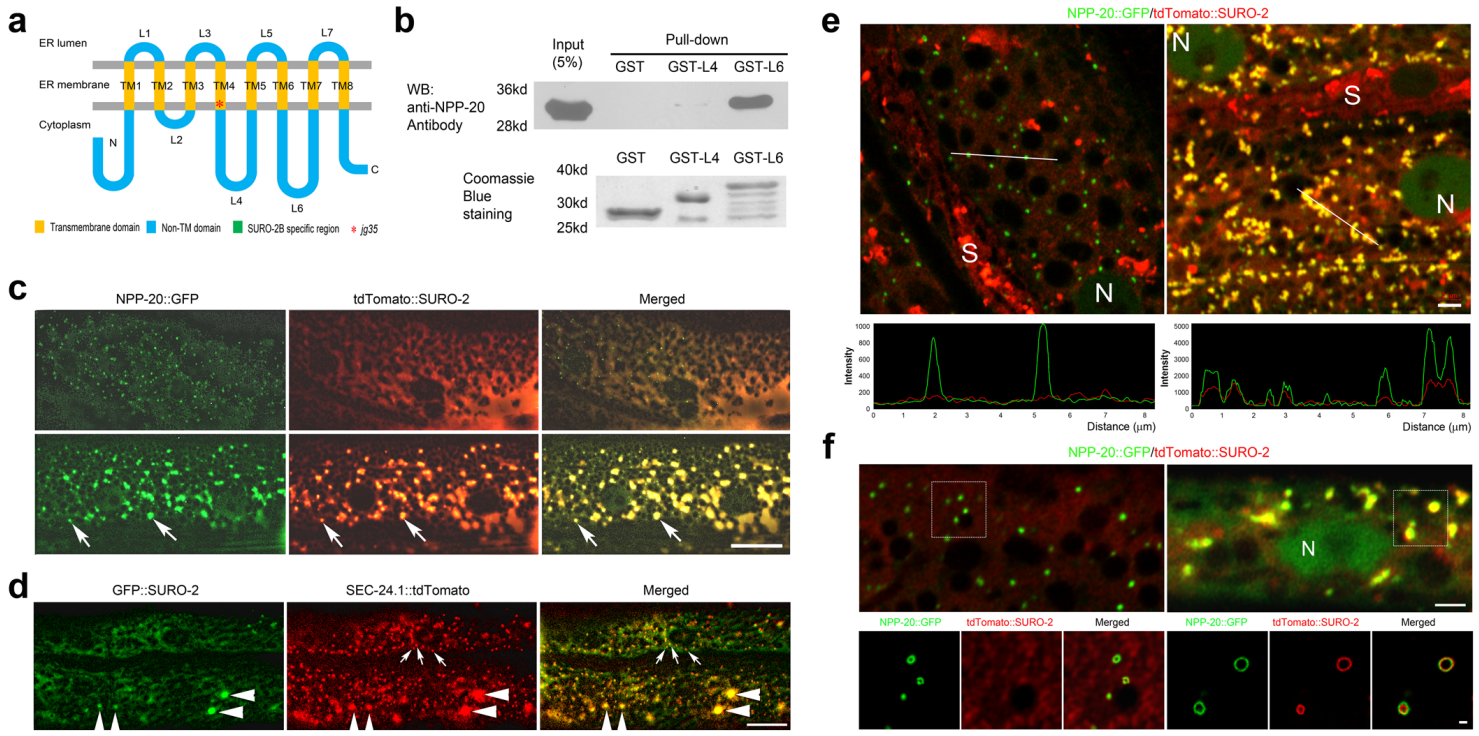

Figure 4

Large ER vesicles made by direct interaction of SURO-2 and NPP-20. a The secondary structure of SURO-2. SURO-2A has two putative long loop domains. The fourth loop (L4) is composed of 66 amino acids, and the sixth loop (L6) is composed of 85 amino acids. The asterisk indicates the position of the suro-2 (jg35) mutation (W187*). TM= transmembrane domain. L= loop. b Results of GST pulldown analysis of SURO-2 and NPP-20. The sixth loop of SURO-2 (GST-L6) binds well to 6xHis-NPP-20 recombinant protein compared to the fourth loop of SURO-2 (GST-L4) and GST. 6xHis-NPP-20 was detected by western blotting with rabbit anti-NPP-20 antibody (top panel). The purified GST fusion protein is presented by Coomassie Blue staining (bottom panel). c Localization of tdTomato::SURO-2 and NPP-20::GFP. In general, fluorescent microscopy showed two types of expression patterns of SURO-2; an ER membrane type (upper panel) and a punctate type of large spots (bottom panel). NPP-20::GFP was colocalized with tdTomato::SURO-2 large puncta (arrows). Scale bar=10 μm. d Colocalization of GFP::SURO-2 and SEC-24.1::tdTomato. According to GFP::SURO-2 localization patterns, SEC-24.1::tdTomato was expressed in either small (arrows) or large puncta (arrowheads). Large puncta of GFP::SURO-2 and SEC-24.1::tdTomato are overlapping, but small puncta of SEC-24.1::tdTomato are solitary. Scale bar=10 μm. e Colocalization of SURO-2 and NPP-20. The integration line jgIs47 expresses both tdTomato::SURO-2 and NPP-20::GFP in the hypodermis. Two expression patterns of tdTomato::SURO-2 were observed with the endoplasmic membrane type (left panel) and the punctate type of large spot (right panel). According to the tdTomato::SURO-2 localization pattern, NPP-20::GFP was expressed in small or large puncta. Large puncta of tdTomato::SURO-2 and NPP-20::GFP are overlapping, but small puncta of NPP-20::GFP are solitary. The bottom panel shows the fluorescence quantification of each area indicated by the white line in the top panel. S= seam cell. N= nucleus of the hypodermal cell. Scale bar= 2 μm. f Higher resolution confocal microscopy of NPP-20::GFP and tdTomato::SURO-2. The left panel shows small NPP-20::GFP puncta, and the right panel shows large puncta using Airyscan. The bottom panel shows an enlarged image of each box area of the top panel after light subtraction. N= nucleus. Scale bars= 2 μm (top panel) and 200 nm (bottom panel).

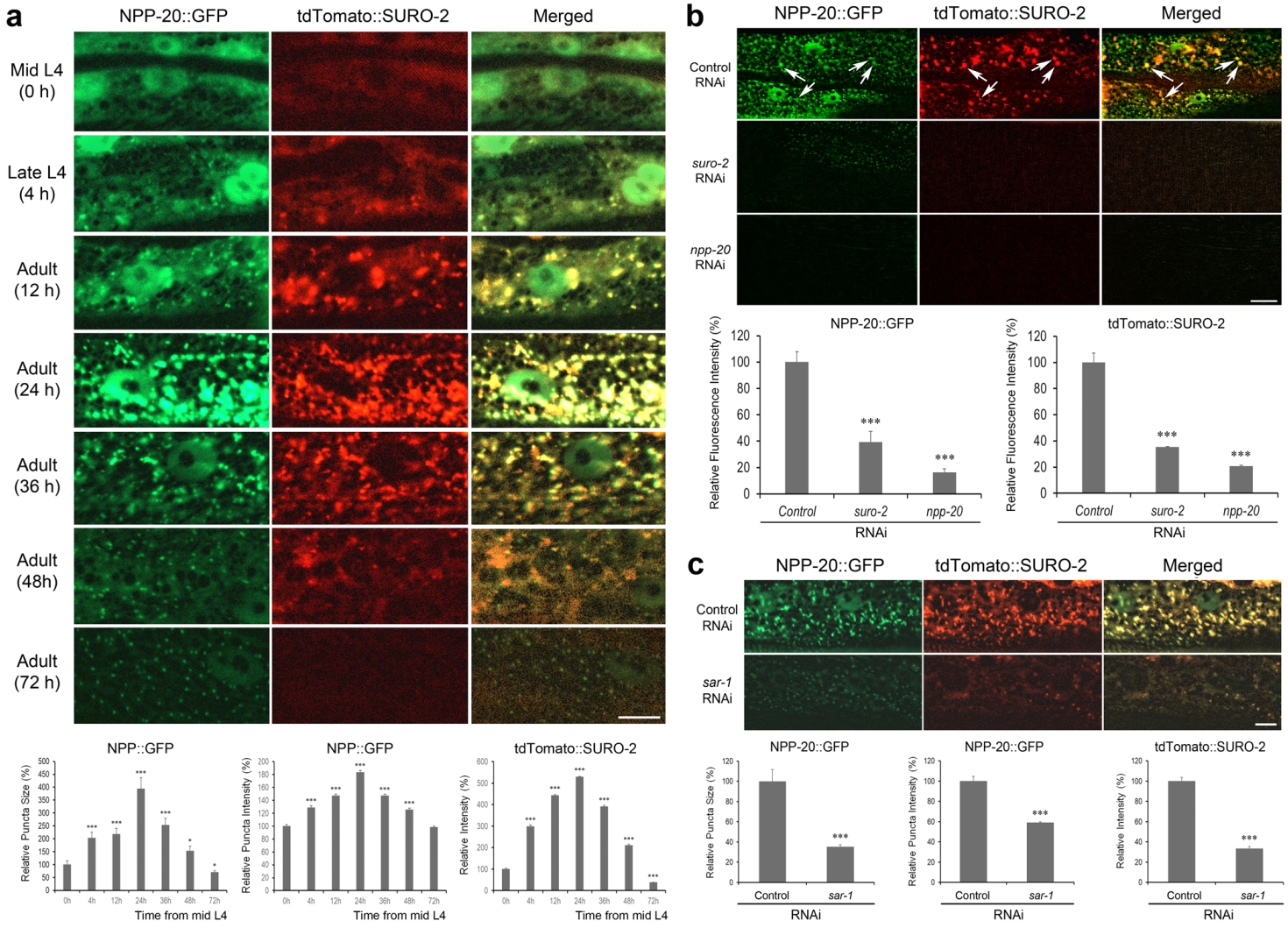

Figure 5

Large COPII vesicle formation during molting. a Large vesicle formation during molting from L4 to adult. The jgIs56 strain expresses both NPP-20::GFP and tdTomato::SURO-2 in the hypodermis using the col-12 promoter. Large vesicles composed of SURO-2 and NPP-20 formed from late L4 larvae to young adults. In old adults, tdTomato::SURO-2 disappeared in contrast to NPP-20::GFP which still expressed as small puncta. Scale bar= 10 μm. The lower panel shows quantification of NPP-20::GFP puncta size, NPP-20::GFP puncta intensity, and tdTomato::SURO-2 intensity. b Knockdown of SURO-2 or NPP-20 reduced large vesicles. Arrows indicates large vesicles consisting of tdTomato::SURO-2 and NPP-20::GFP. SURO-2 depletion abolished large NPP-20::GFP puncta, but not small NPP-20::GFP puncta. Scale bar= 10 μm. The lower panel shows quantification of NPP-20::GFP puncta intensity, and tdTomato::SURO-2 intensity. c SAR-1 is required for large vesicle formation. The number and intensity of large vesicles was reduced by sar-1 knockdown. Scale bar= 10 μm. The lower panel shows quantification of NPP-20::GFP puncta size, NPP-20::GFP puncta intensity, and tdTomato::SURO-2 intensity.

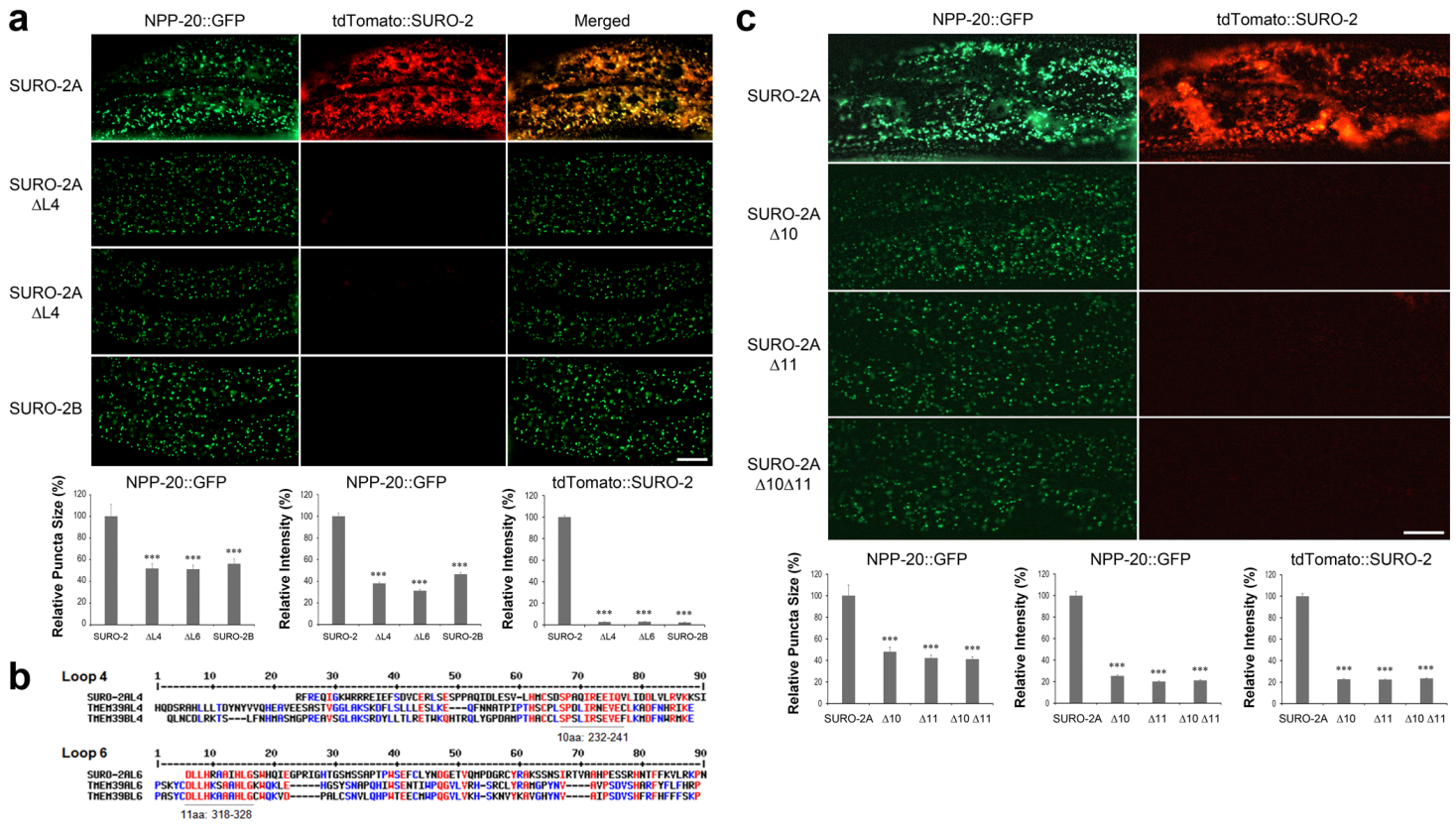

Figure 6

Essential role of SURO-2 long loop domains in large vesicle formation. a The long loop domain exposed to the cytoplasm of SURO-2 is necessary for the formation of large vesicles. Coexpression of NPP-20::GFP and tdTomato::SURO-2 including full-length, L4 deletion, L6 deletion and SURO-2B in the suro-2 (jg35) mutant. Only the full-length tdTomato::SURO-2A protein was well expressed in the suro-2 mutant and large vesicles well formed. Scale bar= 10 μm. The lower panel shows quantification of NPP-20::GFP puncta size, NPP-20::GFP puncta intensity, and tdTomato::SURO-2 intensity. b Sequence alignment of amino acids in the L4 and L6 domains of SURO-2, human TMEM39A and TMEM39B. The underlined region was highly conserved among the three proteins and was used for deletion constructs for the next experiment. The blue letters represent the same amino acids in both proteins, and the red letters represent the same amino acids in all three proteins. c The essential role of conserved regions in L4 and L6 domains in SURO-2 expression and large vesicle formation. tdTomato::SURO-2Δ10 and Δ11 proteins were poorly expressed well in suro-2 mutants, and NPP-20::GFP was unable to form large vesicles with tdTomato::SURO-2 Δ10 or Δ11. Scale bar= 10 μm. The lower panel shows quantification of NPP-20::GFP puncta size, NPP-20::GFP puncta intensity, and tdTomato::SURO-2 intensity.

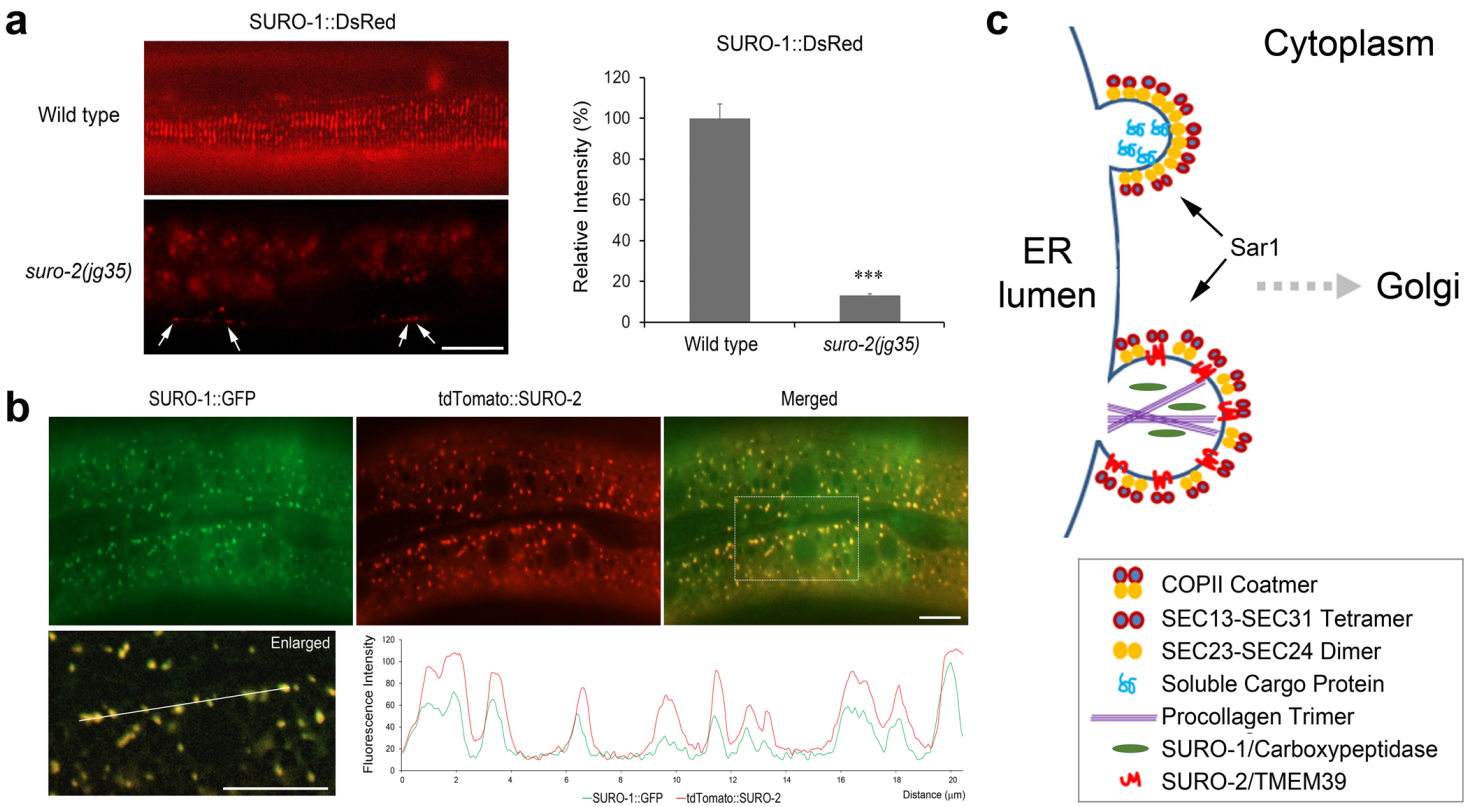

Figure 7

SURO-2-dependent SURO-1 secretion. a SURO-1/CPA expression in cuticle. The SURO-1::DsRed protein was expressed and secreted to the cuticle in the wild type. In contrast, only the marginal cuticle showed weak fluorescence in the suro-2 mutant (arrow). Most red signals of the suro-2 mutant are autofluorescence of intestinal granules. The graph on the right shows the average intensity of SURO-1::DsRed. Scale bar= 10 μm. b Colocalization of SURO-1 and SURO-2. SURO-1::GFP and tdTomato::SURO-2 are located at the same large puncta in the hypodermis. The bottom panel is an enlarged white box on the top panel, and the fluorescence is quantified along the white line. Scale bars= 10 μm. c SURO-2 functional model in the formation of large ER vesicles. Conventional ER vesicles for the transport of soluble cargo proteins were coated with COPII. However, large ER vesicles for the transport of bulky proteins such as collagens and SURO-1/CPA require SURO-2/TMEM39 proteins in addition to COPII. When new cuticle components were synthesized and secreted, SURO-2 forms large vesicles with NPP-20 / SEC13 from the ER membrane. SAR-1/Sar1 is required for both conventional COPII vesicle and large vesicle formation.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfigurelegends.docx