PNAS RESEARCH ARTICLE | CELL BIOLOGY

Functional overlap between the mammalian *Sar1a* and *Sar1b* paralogs in vivo

Vi T. Tang[a,b,1], Jie Xiang[b,2], Zhimin Chen[b], Joseph McCormick[b], Prabhodh S. Abbineni[b,c], Xiao-Wei Chen[d], Mark Hoenerhoff[e], Brian T. Emmer[f], Rami Khoriaty[f,g], Jiandie D. Lin[h], and David Ginsburg[b,f,h,i,3]

Contributed by David Ginsburg; received December 20, 2023; accepted March 29, 2024; reviewed by Anjon Audhya and Mondira Kundu

Proteins carrying a signal peptide and/or a transmembrane domain enter the intracellular secretory pathway at the endoplasmic reticulum (ER) and are transported to the Golgi apparatus via COPII vesicles or tubules. SAR1 initiates COPII coat assembly by recruiting other coat proteins to the ER membrane. Mammalian genomes encode two *SAR1* paralogs, *SAR1A* and *SAR1B*. While these paralogs exhibit ~90% amino acid sequence identity, it is unknown whether they perform distinct or overlapping functions in vivo. We now report that genetic inactivation of *Sar1a* in mice results in lethality during midembryogenesis. We also confirm previous reports that complete deficiency of murine *Sar1b* results in perinatal lethality. In contrast, we demonstrate that deletion of *Sar1b* restricted to hepatocytes is compatible with survival, though resulting in hypocholesterolemia that can be rescued by adenovirus-mediated overexpression of either SAR1A or SAR1B. To further examine the in vivo function of these two paralogs, we genetically engineered mice with the *Sar1a* coding sequence replacing that of *Sar1b* at the endogenous *Sar1b* locus. Mice homozygous for this allele survive to adulthood and are phenotypically normal, demonstrating complete or near-complete overlap in function between the two SAR1 protein paralogs in mice. These data also suggest upregulation of *SAR1A* gene expression as a potential approach for the treatment of SAR1B deficiency (chylomicron retention disease) in humans.

SAR1 | endoplasmic reticulum | COPII

Significance

SAR1A and SAR1B are components of the COPII coat that mediate protein transport from the endoplasmic reticulum to the Golgi. In humans, mutations in *SAR1B* cause chylomicron retention disease (CMRD), whereas mice with SAR1B deficiency exhibit embryonic lethality. No human condition or mouse phenotype has yet been associated with mutations in *SAR1A*. Here, we show that both *Sar1a* and *Sar1b* are required for murine survival. Expression of SAR1A in place of SAR1B appears to fully rescue the lethality of *Sar1b* deficiency in mice, demonstrating a high degree of functional overlap between these two paralogs. This suggests that transcriptional subfunctionalization is the primary evolutionary force maintaining both paralogs in mammals and therapeutic upregulation of SAR1A could be leveraged to treat CMRD.

The intracellular secretory pathway transports proteins from the endoplasmic reticulum (ER) to the Golgi for subsequent secretion into the extracellular space, insertion into the plasma membrane, or transport to various intracellular organelles including lysosomes and endosomes (1). Following proper folding in the ER, cargo proteins are transported to the Golgi apparatus via coat protein complex II (COPII) vesicles/tubules (2, 3). The COPII coat is a highly conserved structure consisting of five core proteins that are present in all eukaryotes: SAR1, SEC23, SEC24, SEC13, and SEC31 (4, 5). GDP-bound SAR1 is recruited to the ER exit site (ERES) by membrane-bound SEC12 to initiate coat assembly. SEC12 also functions as the guanine exchange factor for GDP-SAR1. GTP-bound SAR1 subsequently recruits the SEC23-SEC24 heterodimer complex to the ERES on the cytoplasmic face of the ER membrane via direct interaction between SAR1 and SEC23. SEC23 also acts as the guanine activating factor for SAR1, resulting in GTP hydrolysis. Following cargo recruitment and concentration in the ER lumen mediated by SEC24, SEC13 and SEC31 are recruited to form the outer coat and complete coat assembly (6).

Mammalian genomes contain multiple COPII paralogs as a result of gene duplication events, including two paralogs each for SAR1, SEC23, and SEC31 and four paralogs for SEC24 (6). Gene duplications are common evolutionary events contributing to genome expansion. One of the duplicated copies typically accumulates inactivating mutations over evolutionary time, transitioning to a pseudogene, which is eventually lost from the genome. Occasionally, one duplicate copy undergoes neofunctionalization (acquisition of new function(s) distinct from those of the ancestral gene) or subfunctionalization (division of the ancestral gene's functions among the paralogs), resulting in maintenance of both copies in the genome. Subfunctionalization can occur at either the level of protein function or transcription (7). The expansion of COPII paralogs through the course of eukaryotic evolution is thought to have accommodated increasing secretory demand. Humans and mice deficient in one paralog typically exhibit distinct phenotypes from those deficient in the other (6, 8). In mice, loss of SEC23A results in lethality due to a neural tube closure defect, whereas SEC23B deficiency leads to perinatal death from pancreatic degeneration (9, 10). However, substitution of the *Sec23a* coding sequence at the *Sec23b* locus completely rescues the perinatal lethality

Author contributions: V.T.T., J.X., X.-W.C., B.T.E., R.K., J.D.L., and D.G. designed research; V.T.T., J.X., Z.C., J.M., P.S.A., X.-W.C., M.H., and B.T.E. performed research; Z.C. and J.D.L. contributed new reagents/analytic tools; V.T.T., J.X., P.S.A., X.-W.C., M.H., B.T.E., R.K., J.D.L., and D.G. analyzed data; and V.T.T. and D.G. wrote the paper.

Reviewers: A.A., University of Wisconsin-Madison; and M.K., St. Jude children's research hospital.

The authors declare no competing interest.

[1]Present address: Department of Genetics, Harvard Medical School, Boston, MA 02115.

[2]Present address: Domino's Pizza, Inc., Ann Arbor, MI 48105.

[3]To whom correspondence may be addressed. Email: ginsburg@umich.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2322164121/-/DCSupplemental.

Published April 30, 2024.

exhibited by *Sec23b* null mice (11), demonstrating a high degree of functional overlap between these two paralogs. Mice with complete loss of SEC24A exhibit normal growth and fertility with marked hypocholesterolemia, whereas mice that lack either SEC24B, SEC24C, or SEC24D exhibit lethality at various time points during development (12–15). Both SEC24A/B and SEC24C/D exhibit at least a partial degree of functional overlap (12, 16).

The SAR1A and SAR1B paralogs differ by only 20 out of 198 amino acids in humans, and both appear to be expressed in a wide range of human and mouse tissues (Fig. 1) (17, 18). Mutations in *SAR1B* are associated with the rare autosomal recessive disorder chylomicron retention disease (CMRD, or Anderson's disease) (19), whereas a disorder due to human mutations in *SAR1A* has not been reported. Melville and colleagues previously identified three amino acid clusters differing between human SAR1A and SAR1B that alter GTPase kinetics and SEC23 affinity (20), with a recent report further demonstrating that only human SAR1A is inhibited by the alarmone ppGpp (21). While these findings suggest unique biochemical functions for these two paralogs at the molecular level, there is also evidence for some overlap in function. Genetic deletion of either *SAR1A* or *SAR1B* in Caco-2/15 cells results in decreased chylomicron secretion with a less dramatic phenotype observed in SAR1A null cells. Combined inactivation of both paralogs in Caco2/15 cells is required to recapitulate the severe phenotype seen in cells from patients with CMRD (22, 23). Similarly, while disruption of *SAR1A* is tolerated in human retinal pigment epithelial (RPE-1) cells, additional attempts to delete *SAR1B* in *SAR1A*-null cells were unsuccessful, suggesting some degree of functional compensation between these two paralogs (24).

In mice, germline deletion of *Sar1b* results in late-gestational lethality in homozygous null fetuses, with haploinsufficient animals demonstrating a defect in chylomicron secretion similar to that observed in homozygous CMRD patients (25). Liver-specific deletion of *Sar1b* results in severe hypocholesterolemia due to a lipoprotein secretion defect (26). The consequences of *Sar1a* inactivation in mice is unknown. We now report analysis of mice engineered to be genetically deficient in *Sar1a* or *Sar1b* as well as mice carrying a modified *Sar1b* locus at which the *Sar1a* coding sequence has been substituted for *Sar1b*. Our results demonstrate a high degree of functional overlap between the SAR1A and SAR1B paralogous proteins.

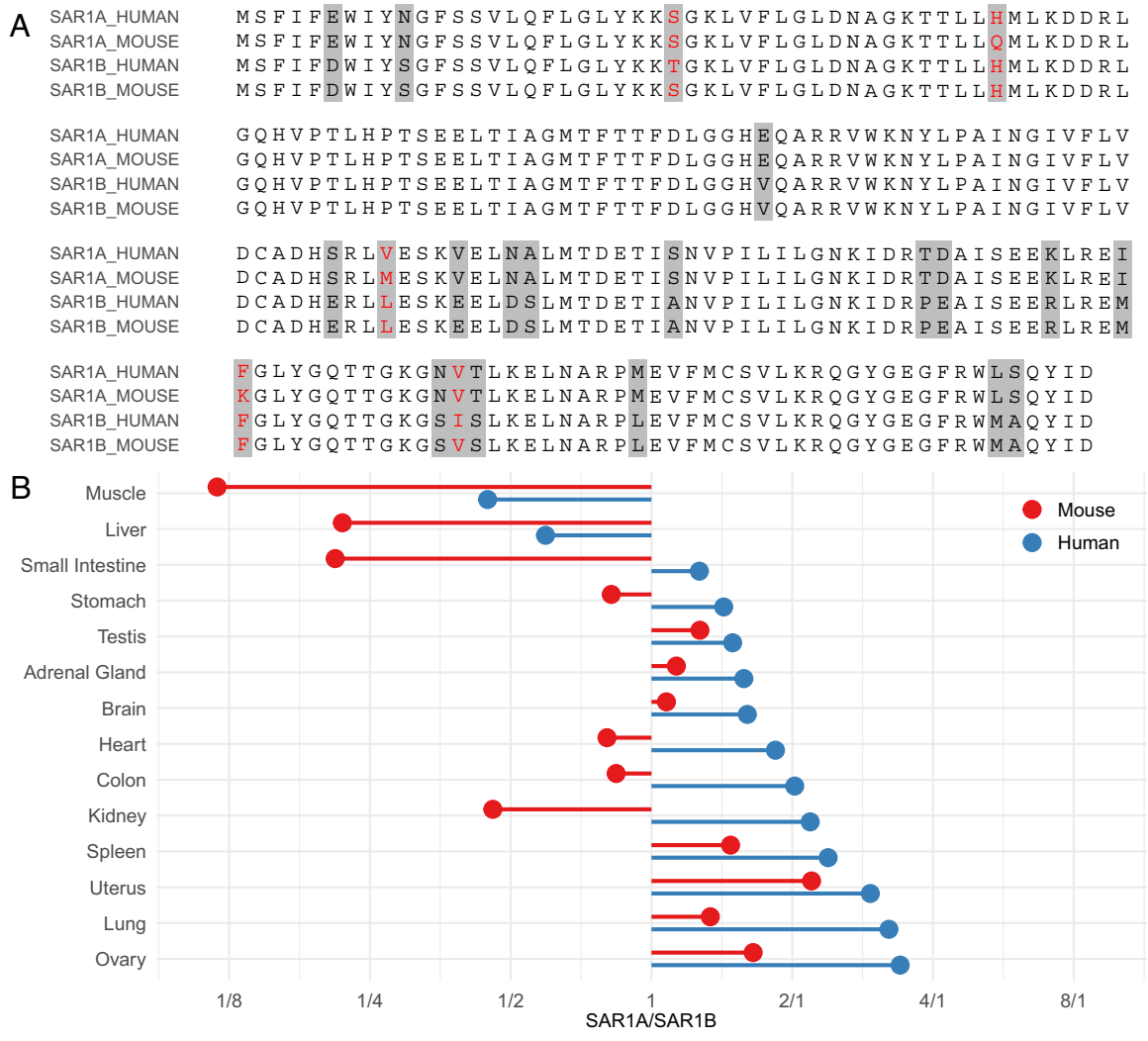


Fig. 1. SAR1A and SAR1B are highly conserved proteins ubiquitously expressed but at variable levels across mouse and human tissues. (*A*) Protein sequence alignment of human and mouse SAR1A (NP_001136120.1 and NP_001345416.1) and SAR1B (NP_001028675.1 and NP_079811.1). Amino acid residues that differ between SAR1A and SAR1B are highlighted in gray, and residues that are different between mouse and human are colored in red. (*B*) Relative *SAR1A* and *SAR1B* mRNA expression in multiple human and mouse tissues. The *x* axis represents SAR1A/SAR1B mRNA abundance ratio in each tissue. Raw RNA sequencing data were obtained from refs. 17 and 18.

Results

Germline Deletion of *Sar1a* in Mice Results in Midembryonic Lethality. To investigate SAR1A function in vivo, we generated mice carrying a *Sar1a* targeted allele in which *Sar1a* expression is disrupted by insertion of a gene trap cassette into intron 5 (27) (Fig. 2*A*). We designed a three-primer PCR assay to differentiate between the wild-type and gene trap alleles (Fig. 2*B*), with DNA sequence analysis confirming the expected configuration. Quantitative PCR (qPCR) analysis of liver cDNA generated from $Sar1a^{+/+}$ and $Sar1a^{gt/+}$ littermates demonstrated an ~50% reduction of *Sar1a* transcripts in the latter, whereas *Sar1b* mRNA abundance remained unchanged, consistent with inactivation of a single *Sar1a* allele (Fig. 2*C*). To generate mice with homozygous loss of *Sar1a*, we performed an intercross between $Sar1a^{gt/+}$ mice, with genotyping results shown in Table 1. No $Sar1a^{gt/gt}$ mice were identified among the 65 offspring genotyped at postnatal day 21 (P21). The results of genotyping for embryos harvested at embryonic days E10.5, E11, and E11.5 post coitus (pc) are shown in Table 1. Although the expected number of $Sar1a^{gt/gt}$ embryos was observed at E10.5, no viable $Sar1a^{gt/gt}$ embryos remained by E11.5. Histologic evaluation of E10.5 embryos revealed that $Sar1a^{gt/gt}$ embryos were smaller than controls but failed to identify any other obvious abnormalities (*SI Appendix*, Fig. S1).

Heterozygous $Sar1a^{gt/+}$ mice were present at the expected Mendelian ratio at weaning (P21) and exhibit normal growth and fertility. At 8 wk of age, no significant differences in body mass or plasma cholesterol levels were observed between $Sar1a^{+/+}$ and $Sar1a^{gt/+}$ mice (Fig. 2 *D* and *E*).

Germline Homozygous *Sar1b* Deletion Results in Perinatal Lethality. We generated heterozygous *Sar1b*-deficient mice ($Sar1b^{+/-}$) by crossing mice carrying a conditional *Sar1b* allele ($Sar1b^{fl/+}$) with mice carrying a ubiquitously expressed Cre (*EIIa-Cre*) transgene. Cre-mediated excision of *Sar1b* exon 5 (Fig. 3*A*) was confirmed with a three-primer PCR assay clearly differentiating the wild-type and null alleles (Fig. 3*B*). As shown in Fig. 3*C*, *Sar1b* mRNA level in livers collected from $Sar1b^{+/-}$ mice was reduced to approximately 50% of that in $Sar1b^{+/+}$ littermates, with no difference in *Sar1a* transcript abundance between the two groups. The results of a heterozygous $Sar1b^{+/-}$ intercross are shown in Table 2. Though no homozygous $Sar1b^{-/-}$ mice were observed at weaning, the expected numbers of $Sar1b^{-/-}$ embryos were present at E11.5 and E18.5, with most $Sar1b^{-/-}$ neonates appearing to die shortly after birth (Table 2). Histologic analysis of E18.5 embryos failed to identify any obvious abnormalities in $Sar1b^{-/-}$ embryos (*SI Appendix*, Fig. S2).

Heterozygous $Sar1b^{+/-}$ mice were observed at the expected numbers at weaning and exhibited normal growth and fertility (Table 2). At 3 to 4 mo of age, no differences in body mass or plasma cholesterol levels were observed between $Sar1b^{+/-}$ and $Sar1b^{+/+}$ mice (Fig. 3 *D* and *E*).

Hypocholesterolemia Resulting from Hepatic SAR1B Deficiency Is Rescued by Overexpression of either SAR1A or SAR1B. We next inactivated *Sar1b* in hepatocytes by combining an *Alb-Cre* transgene and the $Sar1b^{fl}$ allele. $Sar1b^{fl/fl}$ *Alb-Cre*$^{+}$ mice were viable and exhibited a marked reduction in plasma cholesterol levels to

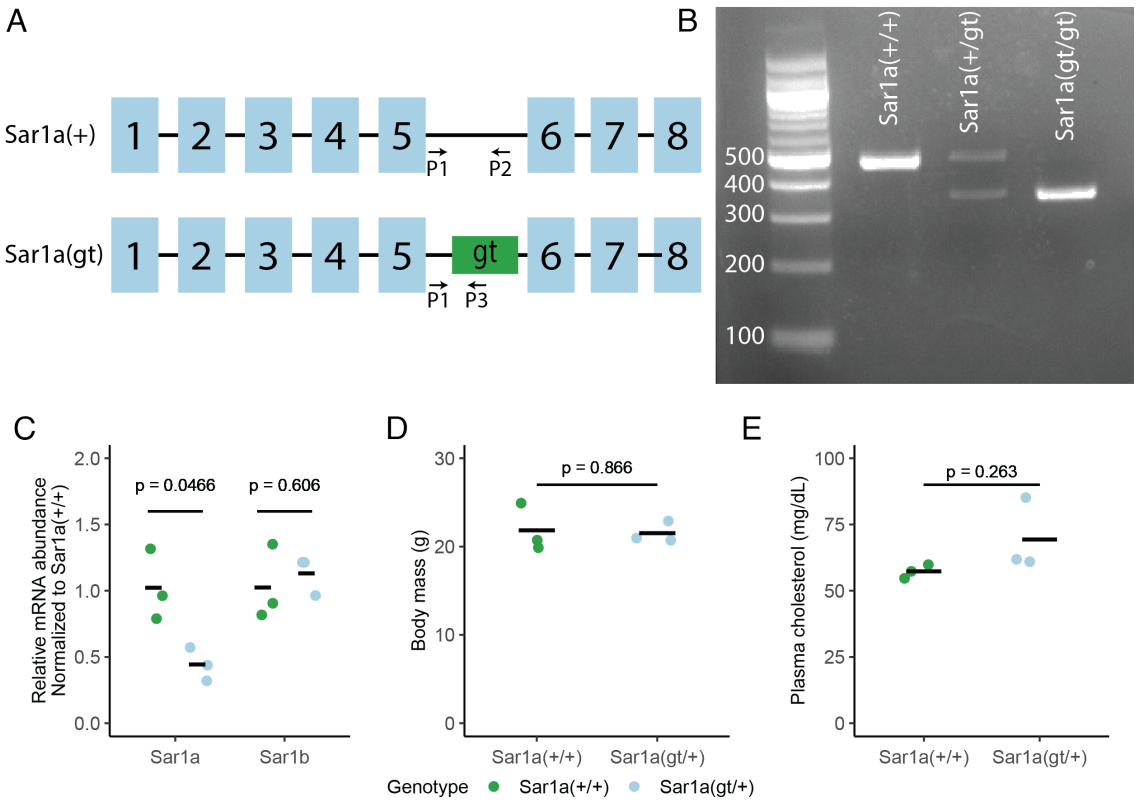


Fig. 2. Haploinsufficiency of SAR1A is tolerated in mice. (*A*) Schematic of the wild-type *Sar1a(+)* and gene trap *Sar1a(gt)* alleles The gene trap (gt) cassette is inserted into intron 5 of the *Sar1a* gene. P1, P2, and P3 denote the binding sites of genotyping primers. Rectangles represent exons, and solid line segments represent introns. Exons and introns are not drawn to scale. (*B*) Genotyping assay using the primers denoted in *A* and genomic DNA isolated from embryonic yolk sacs. The *Sar1a(+)* allele produces a PCR amplicon of 489 bp, whereas the *Sar1a(gt)* allele produces an amplicon of 352 bp. (*C*) Liver *Sar1a* and *Sar1b* mRNA abundance in $Sar1a^{gt/+}$ and littermate controls determined by qPCR (normalized to the mean level of $Sar1a^{+/+}$ samples). (*D* and *E*) Body mass and plasma cholesterol levels of wild-type and heterozygous $Sar1a^{gt/+}$ mice. For panels *C*–*E*, statistical significance was determined by a two-sided *t* test.

Table 1. Germline loss of SAR1A results in lethality during midembryogenesis

Mating genotype expected	$Sar1a^{gt/+}$ X $Sar1a^{gt/+}$			P-value (χ^2 test)
	$Sar1a^{+/+}$	$Sar1a^{gt/+}$	$Sar1a^{gt/gt}$	
	25%	50%	25%	
E10.5	10 (28%)	16 (44%)	10 (28%)	0.800
E11.0	5 (26%)	10 (53%)	4 (21%)	0.924
E11.5	7 (32%)	15 (68%)	0 (0%)*	0.025
Weaning (P21)	17 (26%)	48 (74%)	0 (0%)	7.22×10^{-6}

*Two $Sar1a^{gt/gt}$ embryos appeared dead and were partially resorbed.

approximately 20 mg/dL (compared to ~75 mg/dL in littermate controls, Fig. 4*A*). No reduction in cholesterol was observed in mice with hepatic haploinsufficiency for *Sar1b* ($Sar1b^{fl/+}$ *Alb-Cre*$^+$, Fig. 4*A*), consistent with the normal plasma cholesterol level of $Sar1b^{+/-}$ mice (Fig. 3*E*).

To examine potential overlap in function between SAR1A and SAR1B, we next tested the capacity of an adenovirus (AV) expressing either mouse SAR1A, mouse SAR1B, or GFP (as a control) to rescue the hypocholesterolemia of $Sar1b^{fl/fl}$ *Alb-Cre*$^+$ mice. AVs were injected via the tail vein into $Sar1b^{fl/fl}$ *Alb-Cre*$^+$ mice (Fig. 4*B*) and plasma cholesterol levels were determined on days 0, 3, and 7 post-injection. Injection of the SAR1B AV resulted in a rapid increase in plasma cholesterol to wild-type levels by day 3, with minimal increase observed with the control GFP AV, demonstrating the efficacy of the AV delivery and efficient expression of *Sar1b* in the targeted hepatocytes. Injection of the SAR1A AV resulted in a similar rescue of the hypocholesterolemia phenotype, indistinguishable from that observed with SAR1B AV (Fig. 4*C*). These data demonstrate that SAR1A can compensate for SAR1B deficiency in this assay, suggesting significant overlap in function between the SAR1 paralogs, at least in hepatocytes.

Expression of the *Sar1a* Coding Sequence at the *Sar1b* Endogenous Locus Rescues the Lethal *Sar1b* Null Phenotype in Mice. To further examine the overlap in function between SAR1A and SAR1B in vivo, we generated a $Sar1b^a$ allele in which the *Sar1b* coding sequence is replaced with that of *Sar1a* at the *Sar1b* endogenous locus, beginning immediately downstream of the ATG start codon in exon 2 of *Sar1b* (Fig. 5*A*). Mice homozygous for this allele should be unable to express the SAR1B protein, with SAR1A expressed under the transcriptional control of both the *Sar1a* and *Sar1b* endogenous regulatory elements. A three-primer PCR genotyping assay was developed that could clearly differentiate the wild-type and $Sar1b^a$ alleles (Fig. 5*B*).

To confirm expression of SAR1A under control of *Sar1b* endogenous regulatory elements in $Sar1b^{a/a}$ mice, we isolated RNA from the liver and small intestine of these mice and designed qPCR assays that could distinguish the endogenous *Sar1a(+)* and *Sar1b(+)* transcripts, the *Sar1b(a)* transcript, as well as primers that amplify both the *Sar1a(+)* and *Sar1b(a)* transcripts (total SAR1A-coding mRNA), or the *Sar1b(+)* and *Sar1b(a)* transcripts (total mRNA from the

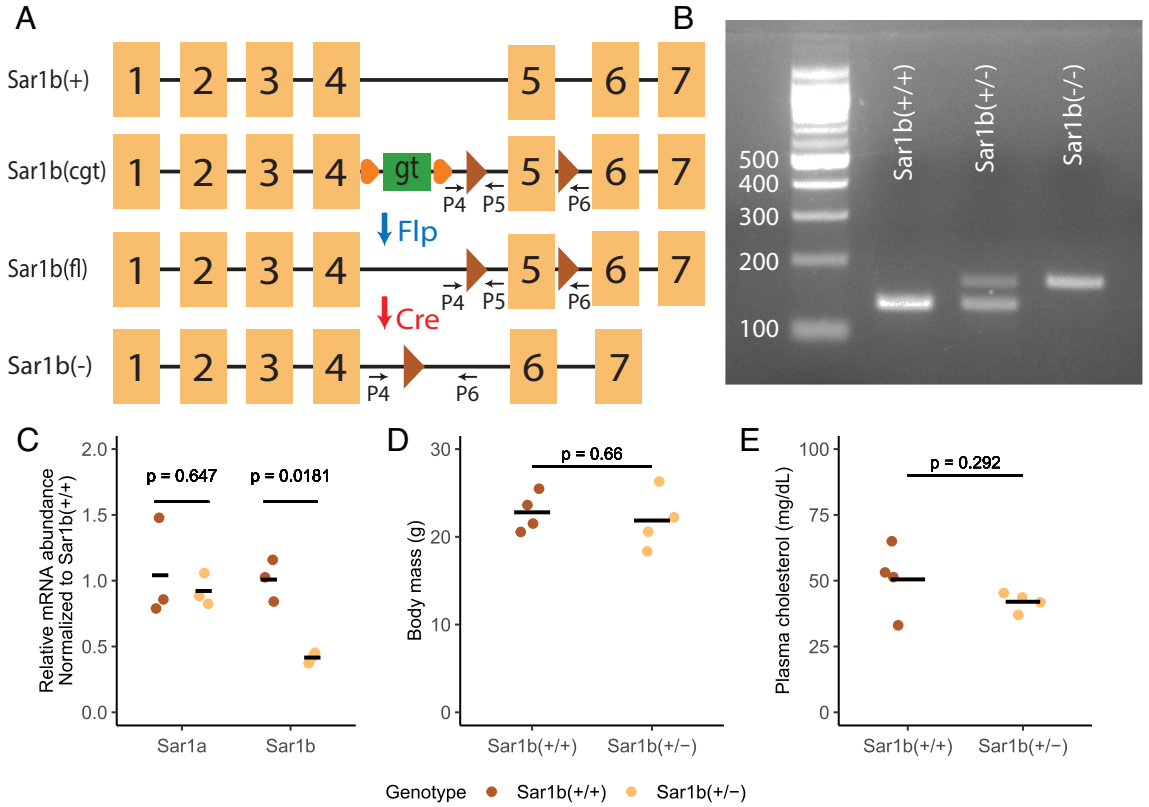


Fig. 3. Mice with heterozygous loss of *Sar1b* demonstrate normal growth and plasma cholesterol levels. (*A*) Schematic of the wild-type *Sar1b(+)*, condition gene trap *Sar1b(cgt)*, conditional *Sar1b(fl)*, and the null *Sar1b(−)* alleles. The gene trap (gt) cassette is inserted into intron 4 of the *Sar1b* gene. Orange circles represent the two FRT sites flanking the gene trap cassette. The Flp recombinase excises the gene trap cassette, resulting in the *Sar1b(fl)* allele. In the *Sar1b(fl)* allele, exon 5 of the *Sar1b* gene is flanked by two loxP sites (brown triangles), which undergo recombination in the presence of Cre recombinase, leading to excision of exon 5 in the *Sar1b(−)* allele. P4, P5, and P6 denote the binding sites of genotyping primers. Rectangles represent exons and solid line segments represent introns. Exons and introns are not drawn to scale. (*B*) Genotyping assay using primers denoted in *A* and genomic DNA isolated from tail clips. The *Sar1b(+)* allele generates a PCR product of 136 bp, whereas the *Sar1b(−)* allele produces a PCR amplicon of 169 bp. (*C*) Relative liver *Sar1a* and *Sar1b* mRNA abundance in $Sar1b^{+/-}$ mice and littermate controls determined by qPCR (normalized to the mean level of $Sar1b^{+/+}$ samples). (*D* and *E*) Body mass and plasma cholesterol levels of wild-type and heterozygous $Sar1b^{+/-}$ mice. For panels *C*–*E*, statistical significance was determined by a two-sided *t* test.

Table 2. Genetic inactivation of *Sar1b* leads to perinatal lethality

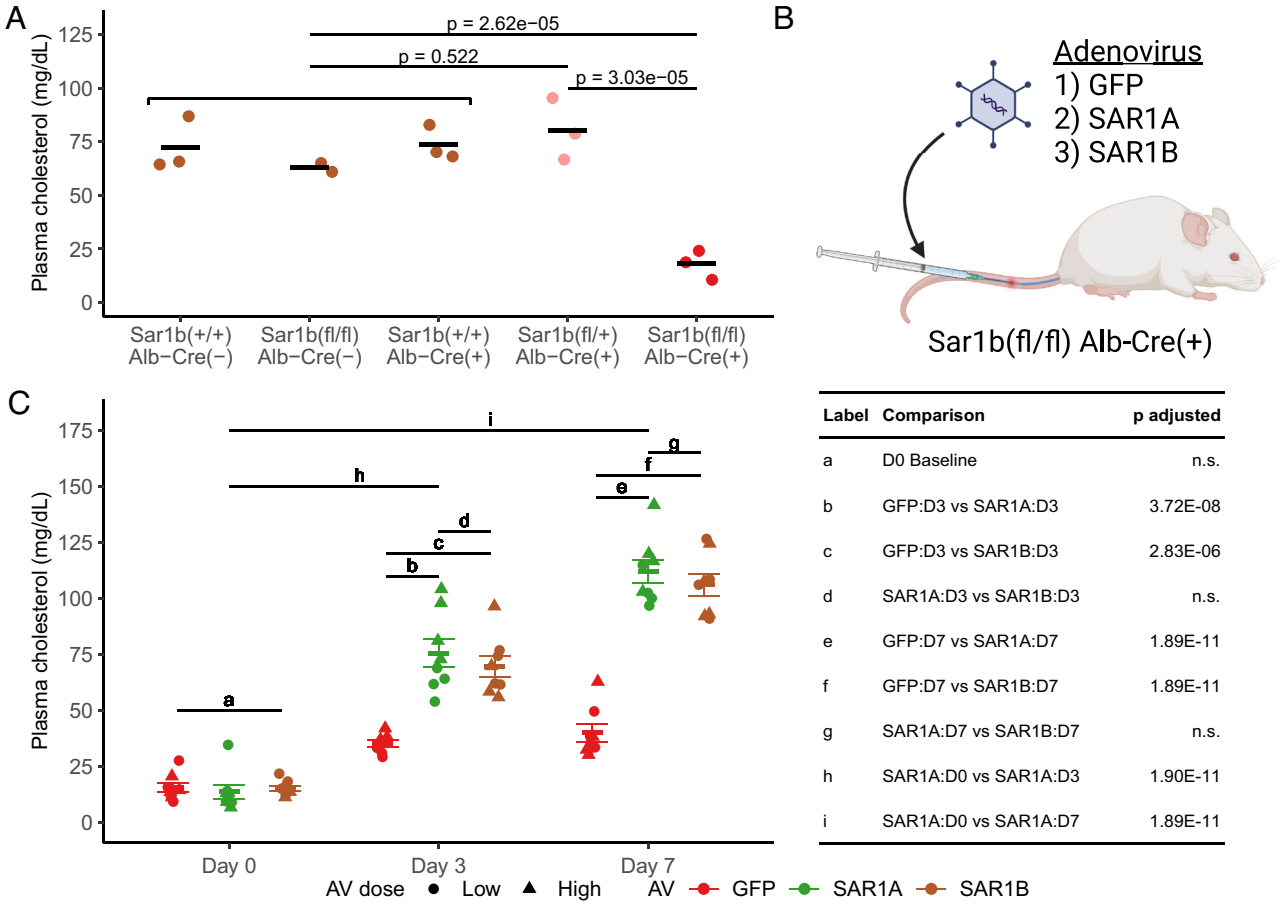
Mating genotype expected	*Sar1b*$^{+/-}$ X *Sar1b*$^{+/-}$			P-value (χ^2 test)
	Sar1b$^{+/+}$	*Sar1b*$^{+/-}$	*Sar1b*$^{-/-}$	
	25%	50%	25%	
E11.5	3 (19%)	8 (50%)	4 (31%)	0.779
E18.5	14 (26%)	24 (44%)	16 (30%)	0.665
P0	4 (31%)	7 (54%)	2 (15%)	0.707
P1	4 (57%)	3 (43%)	0 (0%)	0.095
Weaning (P21)	29 (35%)	54 (65%)	0 (0%)	9.21×10^{-7}

Sar1b locus) (Fig. 5*C*). No significant differences were observed in the abundance of endogenous *Sar1a* mRNA between *Sar1b*$^{+/+}$, *Sar1b*$^{+/a}$, *Sar1b*$^{a/a}$ mice (Fig. 5*D*). However, *Sar1b*$^{+/a}$ and *Sar1b*$^{a/a}$ mice demonstrated a dose-dependent decrease in endogenous *Sar1b* and a corresponding increase in total SAR1A-coding transcript levels as compared to *Sar1b*$^{+/+}$ mice (Fig. 5 *E* and *G*). Although *Sar1b*^{a} transcripts were only detected in *Sar1b*$^{+/a}$ and *Sar1b*$^{a/a}$ (and not *Sar1b*$^{+/+}$) mice (Fig. 5*F*), *Sar1b*^{a} mRNA expression was reduced to ~60% of the wild-type *Sar1b* gene (Fig. 5*H*).

As shown in Table 3, intercrosses between *Sar1b*$^{+/a}$ mice generated homozygous *Sar1b*$^{a/a}$ mice that are present at weaning (though at a slightly lower percentage than the expected Mendelian ratio) and demonstrate normal growth (Fig. 5*I*) and fertility. Subsequent *Sar1b*$^{+/a}$ X *Sar1b*$^{a/a}$ crosses produced *Sar1b*$^{a/a}$ offspring at the expected ratio (Table 3). Histologic examination of multiple organs did not identify any abnormalities in *Sar1b*$^{a/a}$ mice. *Sar1b*$^{a/a}$ mice also demonstrated plasma cholesterol levels indistinguishable from wild-type *Sar1b*$^{+/+}$ and heterozygous *Sar1b*$^{+/a}$ mice (Fig. 5*J*).

Discussion

In this study, we investigate the in vivo function of the two closely related SAR1 paralogs in mice, demonstrating a high degree of overlap. Though both paralogs are required for survival to weaning, complete deficiency for SAR1A results in lethality at midembryogenesis, while *Sar1b* null mice die shortly after birth. Humans with homozygous or compound heterozygous loss-of-function mutations in *SAR1A* have not been reported to date, consistent with an early embryonic lethality phenotype similar to that seen in mice. The perinatal death of *Sar1b* null mice contrasts with the human phenotype, with homozygous *SAR1B* loss-of-function mutations in humans resulting in CMRD, characterized by plasma



Label	Comparison	p adjusted
a	D0 Baseline	n.s.
b	GFP:D3 vs SAR1A:D3	3.72E-08
c	GFP:D3 vs SAR1B:D3	2.83E-06
d	SAR1A:D3 vs SAR1B:D3	n.s.
e	GFP:D7 vs SAR1A:D7	1.89E-11
f	GFP:D7 vs SAR1B:D7	1.89E-11
g	SAR1A:D7 vs SAR1B:D7	n.s.
h	SAR1A:D0 vs SAR1A:D3	1.90E-11
i	SAR1A:D0 vs SAR1A:D7	1.89E-11

Fig. 4. Hypocholesterolemia in hepatic SAR1B-deficient mice is rescued by overexpression of either SAR1A or SAR1B. (*A*) Plasma cholesterol levels in controls (*Sar1b*$^{+/+}$ *Alb-Cre*$^{-}$, *Sar1b*$^{fl/fl}$ *Alb-Cre*$^{-}$, and *Sar1b*$^{+/+}$ *Alb-Cre*$^{+}$—brown), hepatic *Sar1b* haploinsufficient (*Sar1b*$^{fl/+}$ *Alb-Cre*$^{+}$—pink), and hepatic *Sar1b* null (*Sar1b*$^{fl/fl}$ *Alb-Cre*$^{+}$—red) mice. Statistical significance was determined by a pairwise two-sided *t* test followed by Bonferroni adjustment for multiple hypothesis testing. (*B*) Schematic of the SAR1B deficiency rescue experiment. *Sar1b*$^{fl/fl}$ *Alb-Cre*$^{+}$ mice were injected with adenovirus (AV) expressing GFP, mouse SAR1A, or mouse SAR1B via the tail vein. Blood was sampled on days 0, 3, and 7 postinjection and assayed for cholesterol levels. (*C*) Plasma cholesterol levels of mice injected with AV as illustrated in *B*. Two different AV doses were used (low and high—see *Materials and Methods*; n = 4 per dose per AV). Statistical significance was determined by a two-way ANOVA test with interaction between AV type, time, and dose followed by Tukey's post hoc test. *P*-values for the labeled comparisons are listed in the table on the right; n.s., not significant. *P*-values for all comparisons are listed in *SI Appendix*, Table S1.

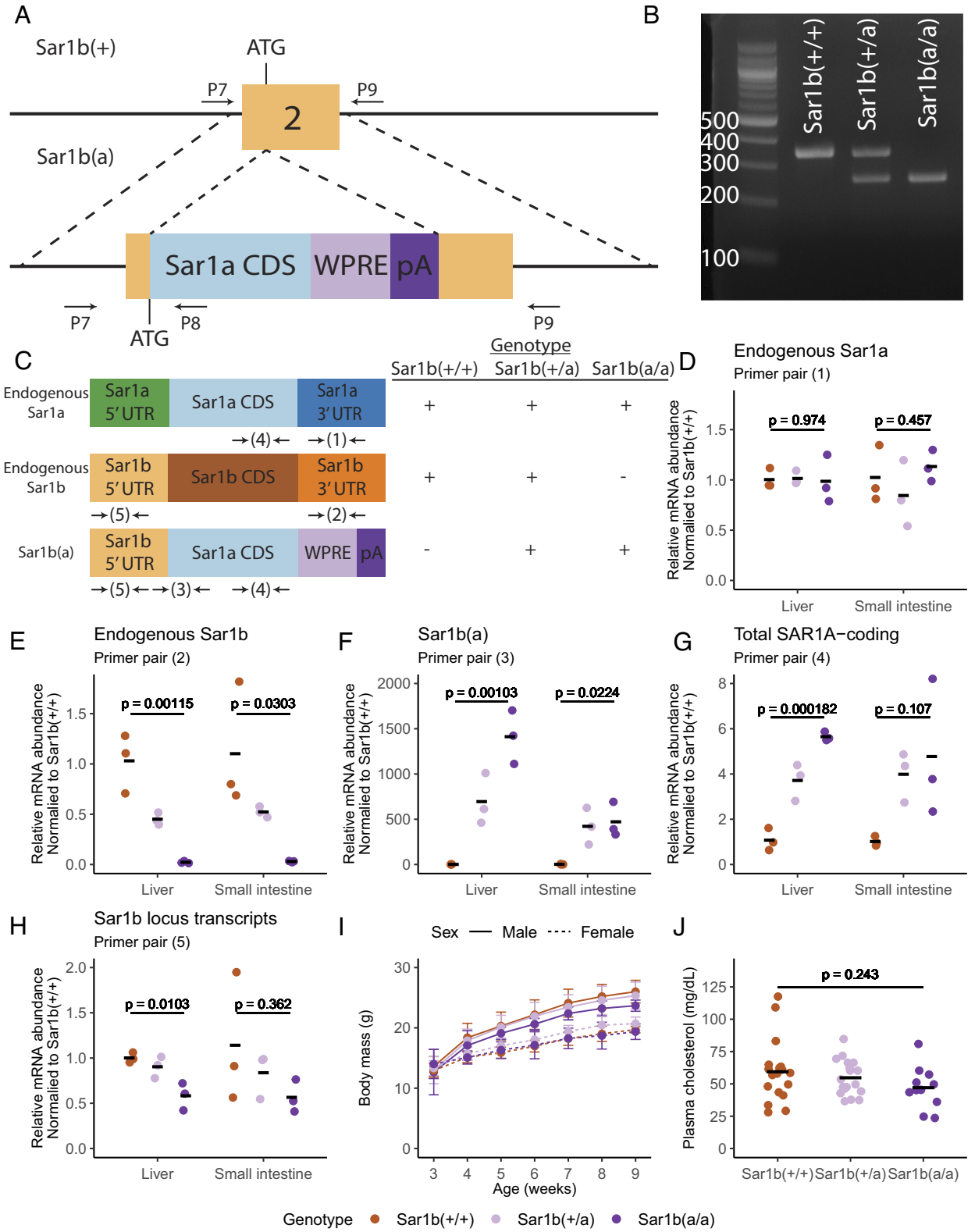


Fig. 5. SAR1A expressed under control of the *Sar1b* endogenous locus can compensate for complete loss of SAR1B. (*A*) Schematic of the *Sar1b(a)* allele. Dashed lines mark the homology arms for DNA recombination. Exons and introns are not drawn to scale. CDS, coding sequence; WPRE, woodchuck hepatitis virus posttranscriptional regulatory element; pA, polyA sequence. P7, P8, and P9 denote binding sites for genotyping primers. (*B*) Genotyping assay using primers depicted in *A* with genomic DNA isolated from tail clips. The wild-type *Sar1b(+)* allele is expected to produce a PCR product of 358 bp (primer pair P7 and P9), whereas the *Sar1b(a)* allele generates a PCR amplicon of 268 bp (primer pair P7 and P8). (*C*) Schematic of the endogenous *Sar1a*, endogenous *Sar1b*, and *Sar1b(a)* transcript. The table on the right shows whether each transcript is present in the indicated genotype. Each number (with arrows on either side) below the transcript represents a particular primer pair used for detecting that transcript by qPCR. Primer pairs (1)–(3) are specific for each of the endogenous *Sar1a*, endogenous *Sar1b*, and *Sar1b(a)* transcript, respectively. Primer pair (4) amplifies both the endogenous *Sar1a* and *Sar1b(a)* transcripts (total SAR1A-coding). Primer pair (5) amplifies both the endogenous *Sar1b* and *Sar1b(a)* transcripts (*Sar1b* locus transcript). (*D–H*) Relative abundance for transcripts as determined by qPCR using the primers illustrated in *C* (normalized to the mean level of $Sar1b^{+/+}$ samples) in livers and small intestines collected from $Sar1b^{+/+}$, $Sar1b^{+/a}$, and $Sar1b^{a/a}$ littermates. Statistical significance was determined by a one-way ANOVA test. (*I*) Body mass of male and female offspring from $Sar1b^{+/a}$ X $Sar1b^{+/a}$ intercrosses between 3 and 9 wk of age (n = 23, 35, and 6 for from $Sar1b^{+/+}$, $Sar1b^{+/a}$, and $Sar1b^{a/a}$, respectively; error bars represent the SD for each group). (*J*) Plasma cholesterol levels of $Sar1b^{+/+}$, $Sar1b^{+/a}$, and $Sar1b^{a/a}$ mice at 3 to 4 mo of age. Statistical significance was determined by a one-way ANOVA test.

Table 3. Expression of SAR1A under the control of the *Sar1b* locus rescues the lethality in *Sar1b* null mice

Mating genotype expected	*Sar1b*$^{+/a}$ X *Sar1b*$^{+/a}$			*P*-value (χ^2 test)
	Sar1b$^{+/+}$	*Sar1b*$^{+/a}$	*Sar1b*$^{a/a}$	
	25%	50%	25%	
Weaning	80 (30%)	148 (55%)	43 (15%)	0.002
Mating genotype expected	*Sar1b*$^{+/a}$ X *Sar1b*$^{a/a}$			
	Sar1b$^{+/a}$		*Sar1b*$^{a/a}$	
	50%		50%	
Weaning	23 (52%)		21 (48%)	0.763

lipid abnormality and survival to adulthood. We did not identify any structural abnormalities in *Sar1a*$^{gt/gt}$ embryos at E10.5 or *Sar1b*$^{-/-}$ fetuses at E18.5. Embryonic lethality is frequently observed in mouse knockout models, with a previous systemic analysis of ~500 such models documenting death at or before birth in approximately 29%, with half of these dying before midembryogenesis (28). The death of *Sar1a*$^{gt/gt}$ embryos at midembryogenesis could result from defective development of the placenta or cardiovascular system, as observed in other genetic models demonstrating similar embryonic loss (29). This midembryonic lethality without overt histologic abnormalities is similar to that observed in mice expressing a hypomorphic *Sec24d* allele (15), potentially consistent with a common mechanism related to impaired protein secretion. Though our analysis of both SAR1A and SAR1B haploinsufficient mice was consistent with normal growth and fertility, deeper phenotyping of *Sar1b*$^{+/-}$ mice by Auclair et al. revealed that these mice do recapitulate some features of CMRD, including lowered plasma levels of triglycerides, decreased chylomicron secretion, and inefficient intestinal fat absorption (25). Thus, *Sar1b*$^{-/-}$ pups could suffer from a more severe phenotype leading to perinatal lethality. Of note, the expected prevalence of heterozygous loss-of-function mutations in either *SAR1A* or *SAR1B* among a large cohort of human subjects in the gnomAD database (30), suggests that haploinsufficiency for these genes is well tolerated in humans.

Despite the high degree of amino acid sequence similarity between the two SAR1 paralogs, complete loss of each paralog in mice results in death at distinct developmental time points. Taken together with previously reported differences in GTP hydrolysis and affinity for SEC23B (20), these data suggest that SAR1A and SAR1B have diverged in protein function. However, we observed that overexpression of SAR1A in the liver rescued the hypocholesterolemia phenotype of mice with hepatocyte-specific SAR1B deficiency (*Sar1b*$^{fl/fl}$ *Alb-Cre*$^{+}$), demonstrating at least partial overlap in SAR1A and SAR1B function in hepatocytes, although confounding effects from SAR1A overexpression or an indirect effect of AV infection cannot be excluded. Of note, there was a small but significant increase in cholesterol levels in mice treated with the control AV (expressing GFP) in this experiment. To address these issues, and to examine the potential for paralog-specific function outside of hepatocytes, we generated mice expressing the SAR1A coding sequence under the transcriptional control of both its endogenous locus and the *Sar1b* locus. The complete rescue of the lethality caused by complete loss of SAR1B expression by the *Sar1b*^{a} allele demonstrates extensive, and potentially complete, overlap in function between the SAR1A and SAR1B proteins in vivo.

Three potential mechanisms have been proposed to explain the maintenance of paralogs such as SAR1A and SAR1B in the genome following gene duplication: neofunctionalization, subfunctionalization at the transcription level, or subfunctionalization at the protein level (7). Our findings suggest that subfunctionalization at the transcription level likely explains the maintenance of the two SAR1 paralogs in mammalian genomes. Expression of both paralogs is required, with loss of a single paralog sufficient to reduce total SAR1 below a critical level in select cell types, resulting in the unique patterns of lethality for SAR1A and SAR1B deficiency in mice and the difference in SAR1B-deficient phenotypes between mice and humans.

The phenotypic differences between humans and mice with SAR1B deficiency is reminiscent of that observed for humans and mice lacking SEC23B (10, 31). Divergence in gene expression patterns across tissues within the same species and across mammalian species could explain this phenomenon for both paralog pairs. The difference in the relative expression of SAR1A and SAR1B in mice and human small intestine, with SAR1B being the dominant paralog in mouse small intestine whereas SAR1A is expressed at a higher level than SAR1B in human small intestine (Fig. 1*B*), could explain the survival of humans carrying homozygous loss-of-function mutations in *SAR1B* while *Sar1b*$^{-/-}$ mice exhibit perinatal lethality.

As noted above, human SAR1A and SAR1B have been shown to exhibit distinct biochemical properties in vitro (20, 21). Despite these differences at the molecular level, our data demonstrate that the corresponding mouse paralogs are sufficiently similar in function such that SAR1A can compensate for the loss of SAR1B in vivo. We cannot exclude greater similarity in function for the mouse versus human paralogs or subtle phenotypic differences between *Sar1b*$^{+/+}$ and *Sar1b*$^{a/a}$ mice. Of note, we did observe a modest loss of *Sar1b*$^{a/a}$ mice in the initial *Sar1b*$^{+/a}$ X *Sar1b*$^{+/a}$ intercross (Table 3). However, a subsequent *Sar1b*$^{+/a}$ X *Sar1b*$^{a/a}$ cross yielded the expected number of *Sar1b*$^{a/a}$ offspring. These data could be the result of mosaicism in the founder mice, stochastic variation in gene expression between individual mice, minor differences in the genetic background of the founder mice and their progeny (C57BL/6N vs. C57BL/6J), or a passenger gene effect (32).

Finally, upregulation of a closely related gene paralog represents a potential strategy for the treatment of select genetic diseases. For example, upregulation of γ-globin expression has been demonstrated to provide dramatic clinical improvement in patients with sickle cell disease and β-thalassemia due to genetic abnormalities in its closely related paralog, β-globin (33, 34). Our data suggest that therapeutic upregulation of one SAR1 paralog could potentially rescue deficiency of the other. Indeed, CRISPR-mediated activation of SEC23A was shown to rescue the erythroid differentiation defect exhibit by a SEC23B-deficient erythroid cell line (35). A similar strategy could potentially be leveraged to treat patients with CMRD due to SAR1B deficiency through the activation of SAR1A.

Materials and Methods

Animal Care. All animal care and use complied with the Principles of Laboratory and Animal Care established by the National Society for Medical Research. Mice were housed in a controlled lighting (12 h light/dark cycle) and temperature (22 °C) environment and had free access to food (5L0D, LabDiet) and water. All animal protocols were approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Michigan (protocol number PRO00011038). Both male and female mice were used in this study.

Generation of *Sar1a*$^{gt/+}$ Mice. The Bay Genomics ES cell clone CSH949 (27) was obtained from the Mutant Mouse Resource & Research Centers (MMRRC). This clone is heterozygous for the *Sar1a*^{gt} allele in which a gene trap cassette is inserted into *Sar1a* gene intron 5 (Fig. 2*A*). ES cell culture, expansion, and microinjection to generate ES cell chimeric mice were performed as previously described (14, 15). Chimeric mice were bred with C57BL/6J (0006640, Jackson Laboratory)

to obtain germ-line transmission. Mice carrying the *Sar1a*^{gt} allele were maintained by continuous backcrossing to C57BL/6J mice.

Generation of *Sar1b*$^{+/-}$ Mice. Mice carrying the *Sar1b*$^{tm1a(EUCOMM)Wtsi}$ allele (36) were obtained from the European Conditional Mouse Mutagenesis (EUCOMM) program (Wellcome Trust Sanger Institute). These mice carry a conditional gene trap (*Sar1b*^{cgt}) allele (Fig. 3*A*) in which the gene trap cassette flanked by FRT sites followed by a loxP site is inserted into intron 4 of the *Sar1b* gene. An additional loxP site is inserted downstream of exon 5 of *Sar1b*. Mice carrying the *Sar1b*^{cgt} allele were maintained by continuous backcrossing to C57BL/6J mice.

To generate the conditional *Sar1b* allele (*Sar1b*^{fl}, Fig. 3*A*), *Sar1b*$^{cgt/+}$ mice were crossed with mice carrying the *Flp1* recombinase gene under control of the *Actb* promoter (005703, Jackson Laboratory). Resulting mice carrying the *Sar1b*^{fl} allele were then bred with mice carrying an *EIIa-Cre* transgene (003724, Jackson Laboratory) to generate the *Sar1b* null allele (*Sar1b*$^{-}$) (Fig. 3*A*). *Sar1b*$^{fl/+}$ and *Sar1b*$^{+/-}$ mice were maintained by continuous backcrossing to C57BL/6J mice.

Generation of *Sar1b*$^{fl/+}$ *Alb-Cre*$^{+}$ Mice. To generate *Sar1b*$^{fl/+}$ *Alb-Cre*$^{+}$ mice, the *Sar1b*$^{fl/+}$ mice generated above were crossed with mice carrying an *Alb-Cre* transgene (003574, Jackson Laboratory) (14). *Sar1b*$^{fl/+}$ *Alb-Cre*$^{+}$ mice were maintained by continuous backcrossing to C57BL/6J mice. *Sar1b*$^{fl/fl}$ *Alb-Cre*$^{+}$ mice were obtained by crossing *Sar1b*$^{fl/+}$ *Alb-Cre*$^{+}$ mice with *Sar1b*$^{fl/+}$ or *Sar1b*$^{fl/fl}$ mice.

Generation of *Sar1b*$^{+/a}$ Mice. Mice carrying the *Sar1b*^{a} allele, in which the mouse *Sar1a* coding sequence is inserted into the *Sar1b* locus immediately downstream of the ATG codon in exon 2 of the gene, were generated by Biocytogen Boston Corp. using CRISPR/Cas9 (Fig. 5*A*). Cas9 and an sgRNA targeting exon 2 of *Sar1b* along with the DNA repair template were injected into C57BL/6N zygotes. Founder mice with confirmed germ-line transmission were bred to C57BL/6N mice to obtain F1 (*Sar1b*$^{+/a}$) mice. *Sar1b*$^{+/a}$ mice were backcrossed to C57BL/6J mice for at least three generations before experimentation. *Sar1b*$^{+/a}$ mice were maintained by continuous backcrossing to C57BL/6J mice.

Mouse Genotyping. Tail clips were obtained from 2- to 3-wk-old mice for genomic DNA isolation and genotyping, as previously described (10). PCR was performed using Go-Taq Green Master Mix (Promega) and resulting products were resolved by 3% agarose gel electrophoresis. All primers used for genotyping are listed in *SI Appendix*, Table S2**.**

Timed Mating and Embryo Collection. Timed matings were performed by intercrossing heterozygous mice (*Sar1a*$^{gt/+}$ or *Sar1b*$^{+/-}$). Embryos were collected at E10.5, E11.0, E11.5, or E18.5 for genotyping and histologic analyses. Genomic DNA was isolated from the embryonic yolk sacs of E10.5–E11.5 embryos or tail clips of E18.5 embryos and genotyped as described above. Embryos were immediately fixed in Z-fix solution (Anatech Ltd.) for subsequent histologic analyses.

Blood and Tissue Collection. Blood and tissues were collected and sera separated as previously described (37). Tissues were immediately fixed in Z-fix solution or frozen in liquid nitrogen. Sera and tissues were stored at −80 °C until experimentation.

Histologic Analyses. Tissue processing, embedding, sectioning, and hematoxylin and eosin (H&E) staining were performed at the University of Michigan In-Vivo Animal Core. Slides were reviewed by the investigators and a veterinarian pathologist blinded to genotype.

Cholesterol Assays. Total cholesterol in sera was determined using a colorimetric assay (SB-1010-225, Fisher Scientific) according to the manufacturer's instructions.

qPCR Assays. Total RNA was isolated from tissues and converted to cDNA as previously described (37). Primers specific for each target (*SI Appendix*, Table S3) were designed using NCBI Primer-BLAST. qPCR were performed using Power SYBR Green PCR Master Mix (4367659, Invitrogen). *Gapdh, B2m, Rpl37,* and *Rpl38* were used as housekeeping gene controls. Normalized transcript abundance was calculated by the $2^{-\Delta\Delta Ct}$ method using the gene with the lowest SD across all samples as the endogenous control.

Adenovirus-mediated Expression of SAR1A or SAR1B. Recombinant adenoviruses expressing mouse SAR1A or SAR1B were constructed using pAdTrack-CMV and the AdEasy adenoviral vector system (Agilent Technologies). Adenoviruses were amplified in Ad293 cells and purified using CsCl gradient ultracentrifugation. Mice were transduced with Ad-GFP, Ad-SAR1A, or Ad-SAR1B via tail vein injection at 0.1 OD per mouse (low dose) or 0.3 OD per mouse (high dose), as described previously (38). Plasma cholesterol concentrations were measured prior to and 3 or 7 d after transduction as described above.

Data, Materials, and Software Availability. All study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. This work was supported by NIH grants R35HL135793 (D.G.), R35HL171421 (D.G.), R01HL148333 (R.K.), R01HL157062 (R.K.), K08HL148552 (B.T.E.), R01HL167733 (B.T.E.), and R00GM141268 (P.S.A.). V.T.T. was supported by fellowships from the American Heart Association (20PRE35210706) and the University of Michigan Horace H. Rackham School of Graduate Studies. Portions of this text are based on the doctoral dissertation of V.T.T.

Author affiliations: [a]Department of Molecular and Integrative Physiology, University of Michigan, Ann Arbor, MI 48109; [b]Life Sciences Institute, University of Michigan, Ann Arbor, MI 48109; [c]Department of Microbiology and Immunology, Loyola University Chicago Stritch School of Medicine, Maywood, IL 60153; [d]State Key Laboratory of Membrane Biology, Peking University, Beijing 100871, China; [e]Unit for Laboratory Animal Medicine, University of Michigan, Ann Arbor, MI 48109; [f]Department of Internal Medicine, University of Michigan, Ann Arbor, MI 48109; [g]Department of Cell and Developmental Biology, University of Michigan, Ann Arbor, MI 48109; [h]Department of Human Genetics, University of Michigan, Ann Arbor, MI 48109; and [i]Department of Pediatrics and Communicable Diseases, University of Michigan, Ann Arbor, MI 48109

1. G. Zanetti, K. B. Pahuja, S. Studer, S. Shim, R. Schekman, COPII and the regulation of protein sorting in mammals. *Nat. Cell Biol.* **14**, 20–28 (2011).
2. C. Barlowe, A. Helenius, Cargo capture and bulk flow in the early secretory pathway. *Annu. Rev. Cell Dev. Biol.* **32**, 197–222 (2016).
3. J. Peotter, W. Kasberg, I. Pustova, A. Audhya, COPII-mediated trafficking at the ER/ERGIC interface. *Traffic* **20**, 491–503 (2019).
4. P. Novick, C. Field, R. Schekman, Identification of 23 complementation groups required for post-translational events in the yeast secretory pathway. *Cell* **21**, 205–215 (1980).
5. C. Barlowe *et al.*, COPII: A membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* **77**, 895–907 (1994).
6. V. T. Tang, D. Ginsburg, Cargo selection in endoplasmic reticulum-to-Golgi transport and relevant diseases. *J. Clin. Invest.* **133**, e163838 (2023).
7. H. Innan, F. Kondrashov, The evolution of gene duplications: Classifying and distinguishing between models. *Nat. Rev. Genet.* **11**, 97–108 (2010).
8. R. Khoriaty, M. P. Vasievich, D. Ginsburg, The COPII pathway and hematologic disease. *Blood* **120**, 31–38 (2012).
9. M. Zhu *et al.*, Neural tube opening and abnormal extraembryonic membrane development in SEC23A deficient mice. *Sci. Rep.* **5**, 15471 (2015).
10. J. Tao *et al.*, SEC23B is required for the maintenance of murine professional secretory tissues. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E2001–E2009 (2012).
11. R. Khoriaty *et al.*, Functions of the COPII gene paralogs SEC23A and SEC23B are interchangeable in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E7748–E7757 (2018).
12. X. W. Chen *et al.*, SEC24A deficiency lowers plasma cholesterol through reduced PCSK9 secretion. *Elife* **2**, e00444 (2013).
13. J. Merte *et al.*, Sec24b selectively sorts Vangl2 to regulate planar cell polarity during neural tube closure. *Nat. Cell Biol.* **12**, 41–46 (2010).
14. E. J. Adams, X. W. Chen, K. S. O'Shea, D. Ginsburg, Mammalian COPII coat component SEC24C is required for embryonic development in mice. *J. Biol. Chem.* **289**, 20858–20870 (2014).
15. A. C. Baines, E. J. Adams, B. Zhang, D. Ginsburg, Disruption of the Sec24d gene results in early embryonic lethality in the mouse. *PLoS One* **8**, 1–10 (2013).
16. E. J. Adams *et al.*, Murine SEC24D can substitute functionally for SEC24C during embryonic development. *Sci. Rep.* **11**, 21100 (2021).
17. GTEx Consortium, The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
18. B. Li *et al.*, A comprehensive mouse transcriptomic BodyMap across 17 tissues by RNA-seq. *Sci. Rep.* **7**, 4200 (2017).
19. B. Jones *et al.*, Mutations in a Sar1 GTPase of COPII vesicles are associated with lipid absorption disorders. *Nat. Genet.* **34**, 29–31 (2003).
20. D. B. Melville, S. Studer, R. Schekman, Small sequence variations between two mammalian paralogs of the small GTPase SAR1 underlie functional differences in coat protein complex II assembly. *J. Biol. Chem.* **295**, 8401–8412 (2020), 10.1074/jbc.ra120.012964.
21. Q. Huang, D. M. E. Szebenyi, The alarmone ppGpp selectively inhibits the isoform A of the human small GTPase Sar1. *Proteins* **91**, 518–531 (2023).
22. A. Sané *et al.*, SAR1B GTPase is necessary to protect intestinal cells from disorders of lipid homeostasis, oxidative stress, and inflammation. *J. Lipid Res.* **60**, 1755–1764 (2019).
23. A. T. Sane *et al.*, Understanding chylomicron retention disease through Sar1b GTPase gene disruption: Insight from cell culture. *Arterioscler. Thromb. Vasc. Biol.* **37**, 2243–2251 (2017).
24. W. Kasberg *et al.*, The Sar1 GTPase is dispensable for COPII-dependent cargo export from the ER. *Cell Rep.* **42**, 112635 (2023).

25. N. Auclair *et al.*, Sar1b mutant mice recapitulate gastrointestinal abnormalities associated with chylomicron retention disease. *J. Lipid Res.* **62**, 100085 (2021).
26. X. Wang *et al.*, Receptor-mediated ER export of lipoproteins controls lipid homeostasis in mice and humans. *Cell Metab.* **33**, 350–366.e7 (2021).
27. D. Stryke *et al.*, BayGenomics: A resource of insertional mutations in mouse embryonic stem cells. *Nucleic Acids Res.* **31**, 278–281 (2003).
28. J. K. White *et al.*, Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. *Cell* **154**, 452–464 (2013).
29. V. E. Papaioannou, R. R. Behringer, Early embryonic lethality in genetically engineered mice: Diagnosis and phenotypic analysis. *Veterinary Pathol.* **49**, 64–70 (2012).
30. K. J. Karczewski *et al.*, The mutational constraint spectrum quantified from` variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
31. K. Schwarz *et al.*, Mutations affecting the secretory COPII coat component SEC23B cause congenital dyserythropoietic anemia type II. *Nat. Genet.* **41**, 936–940 (2009).
32. R. J. Westrick *et al.*, Spontaneous Irs1 passenger mutation linked to a gene-targeted SerpinB2 allele. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 16904–16909 (2010).
33. H. Frangoul *et al.*, CRISPR-Cas9 gene editing for sickle cell disease and beta-thalassemia. *N. Engl. J. Med.* **384**, 252–260 (2021).
34. E. B. Esrick *et al.*, Post-transcriptional genetic silencing of BCL11A to treat sickle cell disease. *N. Engl. J. Med.* **384**, 205–215 (2021).
35. R. King *et al.*, SEC23A rescues SEC23B-deficient congenital dyserythropoietic anemia type II. *Sci. Adv.* **7**, eabj5293 (2021).
36. W. C. Skarnes *et al.*, A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* **474**, 337–342 (2011).
37. V. T. Tang *et al.*, Hepatic inactivation of murine Surf4 results in marked reduction in plasma cholesterol. *Elife* **11**, e82269 (2022).
38. J. Lin *et al.*, Hyperlipidemic effects of dietary saturated fats mediated through PGC-1beta coactivation of SREBP. *Cell* **120**, 261–273 (2005).