1    **SLC25A48 is a human mitochondrial choline transporter**

2

3    Suraj Patil [1,6,9], Oleg Borisov [2], Nora Scherer [2,6], Christophe Wirth [3], Pascal Schlosser [2],

4    Matthias Wuttke [2], Kai-Uwe Eckardt [4,5], Carola Hunte [3,7,8], Björn Neubauer [1], Anna Köttgen [2,7],

5    Michael Köttgen [1,7]

6

7

8    **Affiliations:**

9

10   [1] Department of Medicine IV - Nephrology and Primary Care, Faculty of Medicine and Medical
11   Center, University of Freiburg, Freiburg, Germany

12

13   [2] Institute of Genetic Epidemiology, Faculty of Medicine and Medical Center–University of Freiburg,
14   Freiburg, Germany

15

16   [3] Institute of Biochemistry and Molecular Biology, ZBMZ, Faculty of Medicine, University of Freiburg,
17   Freiburg, Germany.

18

19   [4] Department of Nephrology and Medical Intensive Care, Charité–Universitätsmedizin Berlin,
20   Germany

21

22   [5] Department of Nephrology and Hypertension, University Hospital Erlangen, Friedrich-Alexander-
23   Universität Erlangen-Nürnberg, Erlangen, Germany

24

25   [6] Spemann Graduate School of Biology and Medicine (SGBM), University of Freiburg, Freiburg,
26   Germany

27

28   [7] CIBSS - Centre for Integrative Biological Signalling Studies, University of Freiburg, Freiburg, Germany

29

30   [8] BIOSS-Centre for Biological Signalling Studies, University of Freiburg, Freiburg, Germany.

31

32   [9] Faculty of Biology, University of Freiburg, Germany

33

34   **Correspondence:**

35

36   Anna Köttgen, MD MPH
37   Institute of Genetic Epidemiology
38   Medical Center – University of Freiburg
39   Hugstetter Str. 49, 79106 Freiburg, Germany
40   anna.koettgen@uniklinik-freiburg.de

41

42   Michael Köttgen, MD
43   Department of Medicine IV - Nephrology and Primary Care
44   Medical Center – University of Freiburg
45   Institute for Disease Modeling and Targeted Medicine
46   Breisacherstr. 113, 79106 Freiburg, Germany
47   michael.koettgen@uniklinik-freiburg.de

**Abstract**

Choline has important physiological functions as a precursor for essential cell components and signaling molecules including phospholipids and the neurotransmitter acetylcholine. Choline is a water-soluble charged molecule and therefore requires transport proteins to cross biological membranes. Membrane transport of choline is incompletely understood. Here we show that SLC25A48 is a human mitochondrial choline transporter. Loss-of-function mutations in *SLC25A48* are associated with elevated urine and plasma choline levels resulting from impaired choline transport into mitochondria.

56 **Main**

57 Membrane transport proteins play a crucial role in the movement of ions and metabolites

58 across biological membranes. Their proper functioning is essential for many physiological

59 processes. Despite significant progress in our understanding of these proteins, a substantial

60 number of them remain uncharacterized, often referred to as orphan membrane transport

61 proteins[1]. Deorphanization efforts are essential to unveil the hidden cellular functions and

62 roles of these proteins in health and disease. Recent large-scale genome-wide association

63 studies of metabolite levels (mGWAS) provide links between common genetic variants in

64 membrane transporter-encoding genes and metabolite levels[2,3]. This generates testable

65 hypotheses to identify the physiological substrates of established and orphan human

66 transport proteins *in vivo*.

67

68 Choline is an essential nutrient with important roles in a variety of physiological processes and

69 metabolic pathways. It is an essential component of phospholipids in cell membranes such as

70 phosphatidylcholines, a precursor of the neurotransmitter acetylcholine and of the

71 osmoregulatory betaine, and an important player in lipid metabolism[4]. Recent studies have

72 uncovered a growing body of evidence linking choline to various diseases such as neurological

73 disorders, metabolic syndromes, and liver disease[4]. Several choline transport proteins have

74 been identified and characterized in model systems, including SLC5A7, SLC44A1, SLC44A2,

75 SLC44A4, FLVCR1, and FLVCR2[5-10]. However, it is unclear which transporters affect systemic

76 levels of free choline in humans and if they are linked to human genetic variation.

77

78 We previously found through mGWAS that common *SLC25A48* variants associate with altered

79 choline levels in urine (**Fig. 1a**)[2], and now established that the genetic basis of this association

80    is shared with plasma choline as well as *SLC25A48* transcript levels in brain tissue

81    (**Supplementary Fig. 1**), supporting the hypothesis that the orphan solute carrier SLC25A48

82    may be a choline transporter. Overexpressed SLC25A48 localizes to mitochondria in human

83    epithelial cells (**Fig. 1b**). We therefore isolated mitochondria of cells over-expressing SLC25A48

84    and measured mitochondrial uptake of radio-labelled choline (**Fig. 1c**). Mitochondria from

85    cells over-expressing SLC25A48 showed a significant increase of choline uptake compared to

86    mitochondria from control cells (**Fig. 1d**). Mitochondrial uptake of choline was time- and

87    concentration-dependent (**Fig. 1e, f**). These data establish that SLC25A48 is a newly identified,

88    high-affinity mitochondrial choline transporter. SLC25A48 transports choline at nanomolar

89    concentrations (**Fig. 1f**), and thus operates at physiological concentrations of free choline in

90    humans (7-20 µM) [11].

91

92    To test whether there is a causal link between altered choline levels in humans and impaired

93    mitochondrial choline transport, we investigated the aggregated impact of rare, putatively

94    damaging variants in *SLC25A48* and found significant associations with urine and plasma

95    choline levels (burden test P-values 1.4e-17 and 4.2e-06, respectively; **Supplementary Table**

96    **1**). Carriers of driver variants (Methods) showed significantly higher levels of choline

97    compared to non-carriers, which was more pronounced in urine than in plasma (**Fig. 2a**).

98    Identified driver variants mapped into different regions of *SLC25A48* (**Fig. 2b**).

99

100   Choline has been implicated in various human conditions ranging from neurological diseases

101   to metabolic traits, but there is limited evidence linking genetic defects in choline transport

102   to human disease. We therefore tested whether choline-related, putative loss-of-function

103   (pLoF) mutations in *SLC25A48* were associated with any of the human traits and diseases

104     ascertained in the UK Biobank that are related to tissues where *SLC25A48* is highly expressed

105     (Methods). Although none of the binary (**Supplementary Table 2**) or quantitative

106     (**Supplementary Table 3**) traits was significantly associated with *SLC25A48* pLoF carrier status

107     after correction for multiple testing, we note several suggestive associations with congenital,

108     musculoskeletal, neurological, and eye diseases or traits.

109

110     To investigate whether implicated variants were indeed causally related to altered choline

111     levels, we generated four mutations in the *SLC25A48* cDNA using site-directed mutagenesis

112     and over-expressed these variants in human cell lines to study their effect on protein

113     expression, localization, and function. Western blot analyses showed reduced protein

114     expression of some but not all investigated *SLC25A48* mutations compared to wild-type

115     protein (**Fig. 2c**). Interestingly, some mutations showed mis-localization from mitochondria

116     (p.R179P and p.R243*), with a significant reduction of the mitochondrial co-localization index

117     compared to wild-type SLC25A48 (**Fig. 2d, e**; **Supplementary Fig. 2**), whereas other mutations

118     localized to mitochondria like wild-type transporters (p.D27G and p.R64L; **Supplementary Fig.**

119     **2**). These data show that only some of the investigated *SLC25A48* mutations result in reduced

120     protein expression and/or mis-localization.

121

122     We subsequently measured choline uptake in mitochondria to investigate which mutations

123     affect SLC25A48 transport function. Despite differences in abundance and localization, all

124     tested SLC25A48 mutations significantly impaired mitochondrial choline uptake (**Fig. 2f**).

125     These data establish a causal link between impaired mitochondrial choline transport via

126     SLC25A48 and altered plasma and urine choline levels in humans. From a mechanistic

127     standpoint, our data suggest that the investigated mutations cause loss-of-function through

128   different pathogenic mechanisms, including reduced expression, mis-localization, and

129   impaired substrate turnover with normal expression level and localization.

130         To gain a better mechanistic understanding why specific mutations cause impaired

131   transport, we analyzed structural models of SLC25A48 in two conformations (**Fig. 2g,**

132   **Supplementary Fig. 3**). The SLC25A48 models were in very good agreement with experimental

133   structures of other SLC25 family members, such as the ADP/ATP carrier and mitochondrial

134   uncoupling protein (UCP) (**Supplementary Fig. 3**). Notably, residue D27 is part of a matrix-salt-

135   bridge network of conserved residues, which is a key component of the matrix gate of these

136   transporters and thus essential for transport function. This in line with the strongly impaired

137   choline transport of p.D27G despite its normal localization[12,13]. The other mutations were also

138   localized on the matrix side of the transporter and may affect its folding and stability.

139

140   In summary, we show that the physiological function of SLC25A48 in humans is choline import

141   into mitochondria. Loss-of-function mutations in *SLC25A48* impair choline transport into

142   mitochondria, thereby causing higher levels of free choline in urine and plasma. The

143   deorphanization of SLC25A48 defines its molecular function in humans and enables future

144   well-powered studies addressing its role in health and disease.

145

**Figure legends**

**Fig. 1: SLC25A48 is a human mitochondrial choline transporter. (a)** Regional association plot showing association of common genetic variants in the *SLC25A48* locus with choline levels in urine. **(b)** Flag-tagged SLC25A48 localizes to mitochondria as shown by co-localization with the mitochondrial protein COX4. Colors in merged image: SLC25A48 (turquoise), COX4 (violet), and DAPI (blue). scale bar: 5 µM. **(c, d)** Mitochondrial uptake of radio-labelled $^3$H-choline in cells expressing SLC25A48 compared to mock-transfected control cells (n=10). Average counts per minute (CPMA). *** p<0.001. **(e)** Time course of relative choline uptake in mitochondria. **(f)** Concentration-dependence of choline uptake in mitochondria from cells expressing SLC25A48 compared to mock-transfected controls.

**Fig. 2: Rare damaging variants in *SLC25A48* impair choline transport. (a)** Comparison of inverse normal transformed choline levels in urine and plasma among carriers (N=47) and non-carriers (N=5,572 for urine; N=4,666 for plasma) of putative rare damaging driver variants in *SLC25A48* (P-value unpaired t-test (two-tailed): 3.8e-21 for urine and 1.3e-07 for plasma). **(b)** Localization of rare, damaging driver variants with respect to their protein position in SLC25A48 (Q6ZT89 corresponding to transcript ENST00000681962.1, domains based on InterPro, x-axis). Symbol shape corresponds to variant consequence and the size represents the positive effect size of each individual variant on urine choline levels (**Supplementary Table 1**). Individual variant association P-values with urine choline levels are shown on the y-axis. Variants selected for subsequent functional analyses are labeled. **(c)** Western blot analysis of wild type versus mutant flag-tagged SLC25A48. Actin was used as loading control. **(d)** Cellular localization of the SLC25A48 missense mutation p.R179P compared to wild type SLC25A48.

7

170    Indirect immunofluorescence of SLC25A48-flag and COX4 as mitochondrial marker. Colors in

171    merged image: SLC25A48 (turquoise), COX4 (violet), and DAPI (blue). scale bar: 5 µM. **(e)**

172    Quantification of co-localization of SLC25A48 and COX4 (Pearson correlation coefficient)

173    shows a significant reduction of mitochondrial localization of SLC25A48-R179P compared to

174    wild type (**** $p<0.0001$; see **Supplementary Fig. 2** for other mutants). **(f)** Relative

175    mitochondrial uptake of radio-labelled $^{3}$H-choline in cells expressing wild-type and mutant

176    SLC25A48. Transport of mutant SLC25A48 was normalized to wild-type transport. * $p<0.05$,

177    **** $p<0.0001$. **(g)** Position of damaging mutations in model of SLC25A48 (AlphaFold) in

178    intermembrane space (IMS)-facing conformation (helices numbered). Left: Side chains of

179    mutant residues are highlighted in pink. The protein part that would be truncated in p.R243*

180    is colored in pink. Center, right: Residues of the conserved matrix-salt-bridge network (dotted

181    line) including D27 (mutation in p.D27G) are highlighted in red and blue sticks for acidic and

182    basic residues, respectively.

183

184 **Methods**

185

186 **Molecular biology**

187 pDONR221_SLC25A48 was a gift from the RESOLUTE Consortium (Addgene plasmid # 131995;

188 http://n2t.net/addgene:131995; RRID:Addgene_131995). To incorporate a 5'-Mlu1- and a 3'-

189 Not1-restriction site into the SLC25A48 transcript, PCR was carried out using Pfu Ultra enzyme

190 and the following primers CGC GGG ACG CGT GCC ACC ATG GGC AGC TTC CAG CTG GA and

191 CGC GGG GCG GCC GCC TGG GGA TGT CAC TGC GTG. The resulting fragment was ligated into

192 the pcDNA6.flag vector. To generate plasmids with specific damaging variants, site-directed

193 mutagenesis was performed using the following primers: hSLC25A48.flag _D27G: CTT GTC TTC

194 hSLC25A48.flag_F64L: GCT GGC AGT GGG TAA AGA CAT GCC CTT AAA GAA GCC; GGC TTC TTT

195 AAG GGC ATG TCT TTA CCA CTG CCA GC; hSLC25A48.flag_R179P: GCT TGC TCC CGG ATA CAG

196 TCC TGC AGT CCC; GGG ACT GGC AGG ACT GTA TCC GGG AGC AAG C; To obtain the truncated

197 version SLC25A48.flag_R243*, a shorter transcript was derived from the

198 hSLC25A48.flag.pcDNA6 plasmid through PCR with the primers CGC GGG ACG CGT GCC ACC

199 ATG GGC AGC TTC CAG CTG GA and CGC GGG GCG GCC GCC AGA CTT CAC CAC GTC CAT which

200 then was subcloned into pcDNA6.flag vector as described before.

201

202 **Cell Culture**

203 HEK293T and HeLa cell lines were obtained from the American Type Culture Collection. Both

204 cell lines were cultured as adherent monolayers in Dulbecco's Modified Eagle Medium

205 (DMEM, Gibco), supplemented with 10% heat-inactivated fetal bovine serum (FBS, Biochrom)

206 and Penicillin-Streptomycin (Sigma Aldrich P8781) in a humidified incubator with 5% $CO_2$ at a

207 temperature of 36.5°C. Cells were routinely passaged every 3–4 days using 0.05% trypsin-

208    EDTA (Invitrogen) for detachment. For transfection, HEK293T cells were transfected with

209    either 10 µg of plasmid per 55 cm$^2$ dish or 100 µg of plasmid per 500 cm$^2$ plates using the

210    calcium phosphate transfection method. HeLa cells were transfected using FuGENE HD

211    transfection reagent according to the manufacturer's instructions (Promega E2311).

212

213    **Protein isolation, gel electrophoresis, and Western blot**

214    Cells were harvested 48 hours post transfection, and protein isolation and processing were

215    conducted following established procedures as described previously[14]. Briefly, cell samples

216    were lysed in a cold IP buffer (1% Triton X-100, 20 mM Tris-HCl pH 7.5, 50 mM NaCl, 50 mM

217    NaF, 15 mM Na$_4$P$_2$O$_7$, and 0.1 mM EDTA pH 8) supplemented with 2 mM Na$_3$VO$_4$ and complete

218    protease inhibitor cocktail tablets (Roche 11697498001). The lysates were then subjected to

219    centrifugation at 4°C for 15 minutes at 10,000 g. Supernatants were prepared for further

220    analysis by denaturation for 30 minutes at 42°C with 2x Laemmli buffer.

221        The denatured protein samples were separated using SDS-PAGE, employing precast

222    Mini-PROTEAN TGX 4-15% gels (BioRad 4561086). The separated proteins were then

223    transferred onto PVDF membranes via a wet blot system for 1 hour at 100 V (BioRad).

224    Membranes were blocked with 5% BSA and subsequently incubated with primary antibodies

225    anti-flag (1:3000; Sigma Aldrich F3165and anti-ß-actin (1:5000; Sigma Aldrich A1978)

226    overnight at 4°C. For detection secondary antibodies (anti-rabbit (GE Healthcare NA934V) and

227    anti-mouse (Dako P0447), were used at a dilution of 1:10,000). Chemiluminescence signals

228    were acquired using the Intas ChemoCam system (Intas Science Imaging) within the dynamic

229    range of the charge-coupled device sensor, ensuring that none of the analyzed bands were

230    saturated. The Western blot in Fig. 2c is representative of three experiments with similar

231    results.

232

233 **Immunofluorescence**

234 300,000 Hela cells were seeded in μdish 35mm dishes (ibidi, 81156) and transfected the

235 following day with 2.5 μg of plasmid in 7.8 μl of FuGENE HD transfection reagent (Promega,

236 E2311) and 117.2 μl of water. Two days after transfection, cells were fixed with 3.2 %

237 paraformaldehyde (PFA) for 7 minutes 30 seconds and then permeabilized with a 0.05%

238 Triton-X solution in PBS for 15 minutes. A blocking step was carried out using a mixture of 5%

239 horse serum and 1% BSA in PBS (IF blocking buffer) for 1 hour at RT.

240 　　　　The cells were then incubated overnight with primary antibodies in IF blocking buffer,

241 including anti-flag (Sigma Aldrich F3165, 1:1000) and anti-COX4 (Cell Signaling 4850, 1:200).

242 The following day, the samples were incubated with secondary antibodies for 90 minutes

243 (Hoechst 33342 (Thermo Fisher Scientific, H1399, 1:10000), Alexa Fluor 647 goat anti-rabbit

244 (Invitrogen, A21245, 1:1000), and anti-mouse Alexa Fluor 488 (Molecular Probes, A-11029)).

245 washed, and mounted for microscopy (Dako Glycergel, Agilent). IF-Imaging was performed on

246 a on a LSM980 MP AiryScan 2 equipped with Plan-Apochromat 63x 1.4 oil (Zeiss) within the

247 dynamic range of the detector (pixel size 59 nm x 59 nm). A quantitative colocalization analysis

248 of Flag-tagged SLC25A48 (Alexa Fluor 488) and COX4 (Alexa Fluor 647) was performed with

249 the Colocalization plugin in the Zeiss ZEN blue 3.4 software. 55 to 86 ROIs per group

250 comprising whole cells without nucleus originating from three independent transfections

251 were manually determined. Costes optimal threshold was applied and pixel-based

252 colocalization between two channels was measured to calculate Pearson's correlation

253 coefficient per ROI.

254

255

**Mitochondrial isolation**

HEK293T cells were seeded in a 1:12 dilution in 500 cm$^2$ plates and transfection was performed the following day. Two days after transfection, cells were harvested to obtain crude mitochondrial pellets, following the protocol described in[15]. The mitochondrial pellets were resuspended in a mitochondrial uptake buffer, which was composed of KCl (120 mM), sucrose (25 mM), HEPES (10 mM), EGTA (1 mM), KH$_2$PO$_4$ (1 mM), MgCl$_2$ (5 mM), glutamate (15 mM), and malate (7.5 mM), adjusted to a pH 7.2 [6].

**$^3$H choline Transport studies**

A total of 25 µg of crude mitochondria was suspended in 50 µl of mitochondrial uptake buffer at room temperature. Subsequently, 50 µl of 2x choline buffer (mitochondrial uptake buffer supplemented with 20 µM choline chloride (Merck C7017-5G) (cold choline)), and 10 nM Choline Chloride, [Methyl-3H]-, 1 mCi (revvity, NET109001MC) (hot choline)), resulting in a final concentration of 10 µM cold choline and 5 nM hot choline. The mixture was incubated at room temperature for 5 minutes to facilitate radiolabeled transport.

Upon completion of the incubation, 1 ml of ice-cold mitochondria washing buffer (mitochondrial uptake buffer containing an additional 20 µM choline chloride) was added, and the samples were promptly placed on ice. Following this, centrifugation was carried out at 10,000 g for 5 minutes at 4°C to remove the radiolabeled choline. The resulting pellet was subjected to an additional wash with 300 µl of mitochondria washing buffer through centrifugation at 10,000 g for 5 minutes at 4°C. The final mitochondrial pellet was resuspended in 4 ml of Ultima Gold scintillation cocktail (revvity 6013329), and the $^3$H-choline content was quantified using liquid scintillation counting.

279 For time-course experiments, the incubation duration for radiolabeled transport was varied:

280 30 seconds, 1 minute, 2 minutes, 3 minutes, 5 minutes, 15 minutes, and 30 minutes. For

281 concentration dependence experiments, the choline buffer exclusively contained hot choline

282 (no cold choline) at four different concentrations: 0.05 nM, 0.5 nM, 5 nM, and 50 nM. For

283 time- or concentration-dependent choline uptake experiments, data were normalized to the

284 30 s timepoint or the 0.5 nM measurement from control samples respectively (Fig. 1e, f). For

285 the quantification of relative choline transport (Fig. 2f), the transport activity of mock-

286 transfected mitochondria was subtracted from the transport activity observed in

287 mitochondria over-expressing wild type and mutant SLC25A48, and normalized to the

288 respective wild type transport activity.

289

290 **Statistics**

291 GraphPad Prism® 9.5.1 software was used to graph, analyze and present the obtained data.

292 All results are expressed as mean ± SEM. All experiments were independently performed at

293 least three times and two-tailed Mann-Whitney test was used to calculate p values. A p-

294 value < 0.05 was considered significant.

295

296 **Study population**

297 Genetic associations with choline levels were evaluated in the German Chronic Kidney Disease

298 (GCKD) study, an ongoing prospective cohort study of 5,217 participants with chronic kidney

299 disease stages G3; A1-3 or G1-2; A3 at inclusion that has been described in detail before[16,17].

300 Plasma and urine samples originate from the study's baseline visit, where biosamples were

301 collected, processed, and shipped frozen to a central biobank for storage at -80 degrees

302 Celsius[18]. The GCKD study was registered in the national registry for clinical studies (DRKS

303    00003971) and approved by local ethics committees of the participating institutions[16]. All

304    participants provided written informed consent.

305

306    **Measurement of choline levels**

307    Choline was measured as part of the non-targeted mass spectrometry-based Metabolon HD4

308    Global Discovery panel at Metabolon, Inc. Generation and processing of metabolomics data

309    from stored plasma and spot urine samples of the GCKD study has been described previously[2].

310    Briefly, metabolites were identified by automated comparison of the ion features in the

311    experimental sample to a reference library of chemical standards. Peak quantification was

312    based on the area under the curve, followed by normalization to account for inter-day

313    instrument variation. After quality control, choline levels in urine were normalized for inter-

314    individual dilution using the probabilistic quotient method[19]. Prior to gene-based aggregation

315    testing, choline levels were inverse normal transformed.

316

317    **Whole Exome Sequencing (WES)**

318    The *SLC25A48* gene was investigated using WES data of GCKD participants, which has been

319    described previously[20]. Briefly, extracted genomic DNA underwent paired-end 100-bp WES at

320    Human Longevity Inc, using the IDT xGen v1 capture kit on the Illumina NovaSeq 6000

321    platform. The average coverage of the consensus coding sequence (CCDS) release 22 was 141-

322    fold read depth. Exomes were processed in a custom-built cloud compute platform using the

323    Illumina DRAGEN Bio-IT Platform Germline Pipeline v3.0.7 at Astra Zeneca's Centre for

324    Genomics Research, including alignment to the GRCh38 reference genome and variant

325    calling[21].

326    Sample quality control comprised removal of duplicates, samples with withdrawn

327    consent, mismatch of genetic and reported sex, gonosomal aneuploidies, estimated

328    VerifyBamID contamination level >4%, <94.5% of CCDS bases covered with ≥10-fold coverage,

329    highly related samples (kinship >0.884 by KING --kinship v2.2.3)[21], and missing sample call rate

330    >0.03. Only samples with available high-quality DNA microarray genotype data and without

331    outlying values (>8 SD) along any of the first 10 genetic principal components from a PCA were

332    retained, for a final sample size of 4,779 samples.

333    Variant quality control comprised exclusion for coverage of <10x, genotype quality

334    score (GQ) <30, mapping quality score (MQ) <40, quality score (QUAL) <30, read position rank

335    sum score (RPRS) <-2, mapping quality rank sum score (MQRS) <-8, heterozygous variants with

336    a one-sided binomial exact test P-value for Hardy-Weinberg equilibrium of <1e-6 or genotype

337    called based on an alternative allele read ratio <0.2 or >0.8, single nucleotide variants with a

338    Fisher's strand bias score (FS) >60 and insertions and deletions (indel) with a FS >200, variants

339    that did not pass the DRAGEN calling algorithm filters, and variants with a missing call rate

340    >10% among all remaining samples.

341

342    **Variant annotation and rare variant aggregation (burden) testing**

343    Called variants were annotated using Variant Effect Predictor (VEP)[22] version 109 with

344    standard settings. Predicted deleteriousness of variants was added via REVEL[23] (version 2020-

345    5) and CADD[24] (version 3.0) VEP plugins and via dbNSFP version 4.1a for additional prediction

346    scores[25]. The LoFtee VEP plugin[26] (version 2020-8) was used to downgrade loss-of-function

347    variants. Only variants in transcript ENST00000681962.1, annotated as the MANE Select and

348    Ensembl Canonical transcript for *SLC25A48* by VEP, were considered.

349        The burden test is appropriate when all variants are assumed to affect the trait in the

350        same direction, as is the case for pLoF variants[27]. Qualifying variants in *SLC25A48* for

351        aggregation in burden tests were selected based on variant frequency and VEP annotations.

352        Two complementary masks of selected variants were evaluated: while both masks only

353        assessed rare variants with MAF of <1% in the MANE Select transcript ENST00000681962.1,

354        they differed in the selection of predicted variant effect. The "LoF_mis" mask contained all

355        variants that were predicted to be either high-confidence loss-of-function variants or

356        missense variants with a MetaSVM score >0 or in-frame non-synonymous variants with a

357        fathmm-XF-coding score >0.5, whereas the "HI_mis" mask contained all variants that were

358        predicted either to have a high-impact consequence defined by VEP (transcript ablation, splice

359        acceptor variant, splice donor variant, stop gained, frameshift variant, stop lost, start lost, and

360        transcript amplification) or to be missense variants with either a REVEL score >0.5, a CADD

361        PHRED score >20, or a M-CAP score >0.025. Burden tests were carried out as implemented in

362        the seqMeta R-package version 1.6.7[28], adjusting for age, sex, ln(eGFR), the first three genetic

363        principal components as well as serum albumin for plasma choline and ln(UACR) for urine

364        choline, respectively. Genotypes were coded as number of copies of the rare allele (0, 1, 2).

365        Statistical significance was defined as p<0.05. Single-variant association tests between each

366        selected variant and choline levels were performed under additive modeling, adjusting for the

367        same covariates. To prioritize selected rare putatively damaging variants according to their

368        contribution to the gene signal, we used a previously described forward selection procedure[29]:

369        variants $v$ were ranked by the magnitude of the difference $\Delta_v = P_v - P$, where $P_v$ corresponds to

370        the p-value of the burden test aggregating all variants except for the variant $v$, and $P$ to the

371        total P-value of the burden test including all selected variants. Variants with the greatest $\Delta_v$

372        providing the lowest p-value when aggregated are referred to as "driver variants". Lower

373  association p-values were observed with the "HI_mis" mask as compared to the "LoF_mis"

374  mask, suggesting that the former better captures the genetic architecture of *SLC25A48* with

375  respect to choline levels. Therefore, results are shown for the "HI_mis" mask.

376
377  **Associations with human traits and diseases**

378

379  **Colocalization analysis**

380  We performed colocalization analysis to investigate whether genetic variants associated with

381  choline levels were also associated with gene expression levels (i.e., if these variants act as

382  expression quantitative trait loci, eQTLs). Summary statistics for eQTLs were obtained from

383  the Genotype-Tissue Expression Project (GTEx, V8)[30]. We focused on *SLC25A48* expression in

384  tissues where *SLC25A48* was tissue-enhanced according to the Human Protein Atlas[31]: brain,

385  kidney, and skeletal muscle. For colocalization we used the "coloc" R package and applied the

386  enumeration approach implemented in the "coloc.abf" function[32]. This approach tests

387  whether two traits share the same causal variant in a region by estimating posterior

388  probabilities for five hypotheses: both traits have no causal variant in a region (hypothesis

389  H0), either the first or the second trait has a causal variant (H1 and H2, respectively), both

390  traits have causal variants but they are distinct (H3), and both traits share a common causal

391  variant (H4). Posterior probabilities are assigned to each hypothesis and their sum is equal to

392  1. Two traits show evidence for genetic colocalization when hypothesis H4 obtains the highest

393  probability (0.5 or higher).

394

395

396

**Rare variant collapsing analysis**

We performed phenome-wide association analysis of rare variants in *SLC25A48* with binary traits (clinical diagnoses based on ICD-10 codes) and quantitative traits in the UK Biobank, which comprises up to 500,000 individuals with WES and comprehensive phenotypic data[33] (UK Biobank application ID 64806). We excluded strongly related individuals (for whom ten or more third-degree relatives were identified) and individuals not included in the kinship inference process. Participants of all ancestries were included into analysis. To adjust for population stratification, we included the first 10 principal components based on genotype array data (UKB Data-Field 22009) when performing all downstream regression analyses for both binary and quantitative traits. In addition to principal components, we included sex, age at recruitment, and interaction between sex and age (sex*age) as covariates.

After individual-level filtering, a total of 468,292 individuals were available for association analyses. Both masks were tested: "LoF_mis" (94 variants) and "HI_mis" (191 variants)[20]. All variants in each mask passed the "90pct10dp" quality filter, defined as a read depth of at least 10 in at least 90% of all genotypes for a given variant independent of variant allele zygosity. Carriers and non-carriers of variants in each mask were compared using the collapsing approach, in which carriers are defined as individuals with either heterozygous or alternative homozygous genotype for at least one of the variants in the mask. The remaining individuals were defined as non-carriers. All analyses were performed on the UK Biobank Research Analysis Platform.

**Analysis of binary traits and diseases**

Traits and diseases were defined based on ICD-10 codes (International Classification of Diseases, 10th Revision). We queried all distinct diagnosis codes from the UK Biobank

18

421    database for each participant across all hospital inpatient records in either the primary or

422    secondary position (UK Biobank Data-Field 41270). Next, we grouped ICD-10 codes using

423    phecode system[34]. The phecode system maps related ICD-10 codes to a larger group of codes

424    known as "phecode". In total, there are approximately 1,500 phecodes. Each phecode is

425    assigned to a clinically meaningful category (e.g., neurological diseases, musculoskeletal

426    diseases) to facilitate interpretation. We analyzed 658 phecodes belonging to the following

427    categories: circulatory system, congenital anomalies, digestive, endocrine/metabolic,

428    genitourinary, mental disorders, musculoskeletal, neurological, pregnancy complications, and

429    retinal traits from sense organs. For each phecode, we defined case-control status of

430    participants based on the presence or absence of the respective phecode. Since we focused

431    on rare variants, we applied Firth's logistic regression for association testing (mean bias-

432    reducing adjusted scores approach) as implemented in the "brglm2" R package[35]. Firth's

433    regression provides bias-reduction in case of rare events and also allows for inclusion of

434    covariates. We report 83 phecodes from both masks (73 unique phecodes) with nominally

435    significant associations (P-value <0.05) in **Supplementary Table 2**.

436

437    **Analysis of quantitative traits**

438    Among quantitative traits available in the UK Biobank, we selected a subset of 1,143 physical

439    measures, blood assays, and imaging traits that belonged to the following categories: Chapter

440    V Mental and behavioural disorders, Chapter VI Diseases of the nervous system, Chapter VII

441    Diseases of the eye and adnexa, Chapter IX Diseases of the circulatory system, Chapter XI

442    Diseases of the digestive system, Chapter XIII Diseases of the musculoskeletal system and

443    connective tissue, Chapter XIV Diseases of the genitourinary system[21], as well as all available

444    blood biochemistry measurements from all chapters. Values of these traits were inverse-

19

445 normal transformed prior to analyses. We applied linear regression for association testing

446 between quantitative traits and collapsed variants adjusting for covariates as mentioned

447 above. We report 106 quantitative traits from both masks (103 unique traits) with nominally

448 significant associations (P-value <0.05) in **Supplementary Table 3**.

449

450 **Structural model generation and analysis**

451 The SLC25A48 AlphaFold[36] model was retrieved from the AlphaFold Protein Structure

452 Database[37] hosted by EMBL-EBI (https://alphafold.ebi.ac.uk, accession code Q6ZT89). To

453 identify potential homologous proteins with experimental structures available, a sequence-

454 based BLAST[38] search of the Protein Data Bank as well as a structural homology based search

455 using the DALI Server[39] were performed. Structures of other SLC25 family members with

456 accession codes 1ock (ADP/ATP carrier from bovine[40]), 2lck (UCP2 from mouse[41]), 4c9g

457 (ADP/ATP carrier 2 from yeast[42]), 4c9j (ADP/ATP carrier 3 from yeast[42]), 6gci (ADP/ATP carrier

458 from *Thermothelomyces* thermophilus[43]), 8g8w (UCP1 from human, GTP bound[12]) and 8hbv

459 (UCP1 from human, nucleotide-free state[44]) were obtained from the Protein Data Bank.

460 Structures and model were superimposed using Coot[45]. As the SLC25A48 AlphaFold model is

461 in the intermembrane space open conformation, Swiss-Model[46] was used to model the matrix

462 open conformation using the structure of ADP/ATP carrier from *T. thermophilus*[43] in matrix

463 open conformation as template. Figures were prepared using the PyMOL Molecular Graphics

464 System, Schrödinger, LLC.

465

**References**

1.  Meixner, E. *et al.* A substrate-based ontology for human solute carriers. *Mol Syst Biol* **16**, e9652 (2020).
2.  Schlosser, P. *et al.* Genetic studies of paired metabolomes reveal enzymatic and transport processes at the interface of plasma and urine. *Nat Genet* **55**, 995-1008 (2023).
3.  Suhre, K. *et al.* Human metabolic individuality in biomedical and pharmaceutical research. *Nature* **477**, 54-60 (2011).
4.  Ueland, P.M. Choline and betaine in health and disease. *J Inherit Metab Dis* **34**, 3-15 (2011).
5.  Apparsundaram, S., Ferguson, S.M., George, A.L., Jr. & Blakely, R.D. Molecular cloning of a human, hemicholinium-3-sensitive choline transporter. *Biochem Biophys Res Commun* **276**, 862-7 (2000).
6.  Bennett, J.A. *et al.* The choline transporter Slc44a2 controls platelet activation and thrombosis by regulating mitochondrial function. *Nat Commun* **11**, 3479 (2020).
7.  Cater, R.J. *et al.* Structural and molecular basis of choline uptake into the brain by FLVCR2. *bioRxiv* (2023).
8.  Kenny, T.C. *et al.* Integrative genetic analysis identifies FLVCR1 as a plasma-membrane choline transporter in mammals. *Cell Metab* **35**, 1057-1071 e12 (2023).
9.  Michel, V. & Bakovic, M. The solute carrier 44A1 is a mitochondrial protein and mediates choline transport. *FASEB J* **23**, 2749-58 (2009).
10. Traiffort, E., O'Regan, S. & Ruat, M. The choline transporter-like family SLC44: properties and roles in human diseases. *Mol Aspects Med* **34**, 646-54 (2013).
11. Garcia, E. *et al.* Quantification of choline in serum and plasma using a clinical nuclear magnetic resonance analyzer. *Clin Chim Acta* **524**, 106-112 (2022).
12. Jones, S.A. *et al.* Structural basis of purine nucleotide inhibition of human uncoupling protein 1. *Sci Adv* **9**, eadh4251 (2023).
13. Ruprecht, J.J. & Kunji, E.R.S. The SLC25 Mitochondrial Carrier Family: Structure and Mechanism. *Trends Biochem Sci* **45**, 244-258 (2020).
14. Hofherr, A. *et al.* The mitochondrial transporter SLC25A25 links ciliary TRPP2 signaling and cellular metabolism. *PLoS Biol* **16**, e2005651 (2018).
15. Wieckowski, M.R., Giorgi, C., Lebiedzinska, M., Duszynski, J. & Pinton, P. Isolation of mitochondria-associated membranes and mitochondria from animal tissues and cells. *Nat Protoc* **4**, 1582-90 (2009).
16. Eckardt, K.U. *et al.* The German Chronic Kidney Disease (GCKD) study: design and methods. *Nephrol Dial Transplant* **27**, 1454-60 (2012).
17. Titze, S. *et al.* Disease burden and risk profile in referred patients with moderate chronic kidney disease: composition of the German Chronic Kidney Disease (GCKD) cohort. *Nephrol Dial Transplant* **30**, 441-51 (2015).
18. Prokosch, H.U. *et al.* Designing and implementing a biobanking IT framework for multiple research scenarios. *Stud Health Technol Inform* **180**, 559-63 (2012).
19. Dieterle, F., Ross, A., Schlotterbeck, G. & Senn, H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal Chem* **78**, 4281-90 (2006).

535    20.    Pfau, A. *et al.* SLC26A1 is a major determinant of sulfate homeostasis in humans. *J*
536            *Clin Invest* **133**(2023).
537    21.    Wang, Q. *et al.* Rare variant contribution to human disease in 281,104 UK Biobank
538            exomes. *Nature* **597**, 527-532 (2021).
539    22.    McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* **17**, 122 (2016).
540    23.    Ioannidis, N.M. *et al.* REVEL: An Ensemble Method for Predicting the Pathogenicity of
541            Rare Missense Variants. *Am J Hum Genet* **99**, 877-885 (2016).
542    24.    Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J. & Kircher, M. CADD: predicting
543            the deleteriousness of variants throughout the human genome. *Nucleic Acids Res* **47**,
544            D886-D894 (2019).
545    25.    Liu, X., Li, C., Mou, C., Dong, Y. & Tu, Y. dbNSFP v4: a comprehensive database of
546            transcript-specific functional predictions and annotations for human nonsynonymous
547            and splice-site SNVs. *Genome Med* **12**, 103 (2020).
548    26.    Karczewski, K.J. *et al.* The mutational constraint spectrum quantified from variation in
549            141,456 humans. *Nature* **581**, 434-443 (2020).
550    27.    Lee, S., Abecasis, G.R., Boehnke, M. & Lin, X. Rare-variant association analysis: study
551            designs and statistical tests. *Am J Hum Genet* **95**, 5-23 (2014).
552    28.    Voorman, A., Brody, J., Chen, H., Lumley, T. & Davis, B. . seqMeta: Meta-Analysis of
553            Region-Based Tests of Rare DNA Variants. (2017).
554    29.    Bomba, L. *et al.* Whole-exome sequencing identifies rare genetic variants associated
555            with human plasma metabolites. *Am J Hum Genet* **109**, 1038-1054 (2022).
556    30.    Consortium, G.T. The GTEx Consortium atlas of genetic regulatory effects across
557            human tissues. *Science* **369**, 1318-1330 (2020).
558    31.    Uhlen, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**,
559            1260419 (2015).
560    32.    Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic
561            association studies using summary statistics. *PLoS Genet* **10**, e1004383 (2014).
562    33.    Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data.
563            *Nature* **562**, 203-209 (2018).
564    34.    Wu, P. *et al.* Mapping ICD-10 and ICD-10-CM Codes to Phecodes: Workflow
565            Development and Initial Evaluation. *JMIR Med Inform* **7**, e14325 (2019).
566    35.    Kosmidis, I., Kenne Pagui, E.C. & Sartori, N. Mean and median bias reduction in
567            generalized linear models. . *Stat Comput* 43–59 (2020).
568    36.    Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature*
569            **596**, 583-589 (2021).
570    37.    Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the
571            structural coverage of protein-sequence space with high-accuracy models. *Nucleic*
572            *Acids Res* **50**, D439-D444 (2022).
573    38.    Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment
574            search tool. *J Mol Biol* **215**, 403-10 (1990).
575    39.    Holm, L., Laiho, A., Toronen, P. & Salgado, M. DALI shines a light on remote
576            homologs: One hundred discoveries. *Protein Sci* **32**, e4519 (2023).
577    40.    Pebay-Peyroula, E. *et al.* Structure of mitochondrial ADP/ATP carrier in complex with
578            carboxyatractyloside. *Nature* **426**, 39-44 (2003).
579    41.    Berardi, M.J., Shih, W.M., Harrison, S.C. & Chou, J.J. Mitochondrial uncoupling protein
580            2 structure determined by NMR molecular fragment searching. *Nature* **476**, 109-13
581            (2011).

582    42.    Ruprecht, J.J. *et al.* Structures of yeast mitochondrial ADP/ATP carriers support a
583           domain-based alternating-access transport mechanism. *Proc Natl Acad Sci U S A* **111**,
584           E426-34 (2014).
585    43.    Ruprecht, J.J. *et al.* The Molecular Mechanism of Transport by the Mitochondrial
586           ADP/ATP Carrier. *Cell* **176**, 435-447 e15 (2019).
587    44.    Kang, Y. & Chen, L. Structural basis for the binding of DNP and purine nucleotides
588           onto UCP1. *Nature* **620**, 226-231 (2023).
589    45.    Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot.
590           *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
591    46.    Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and
592           complexes. *Nucleic Acids Res* **46**, W296-W303 (2018).
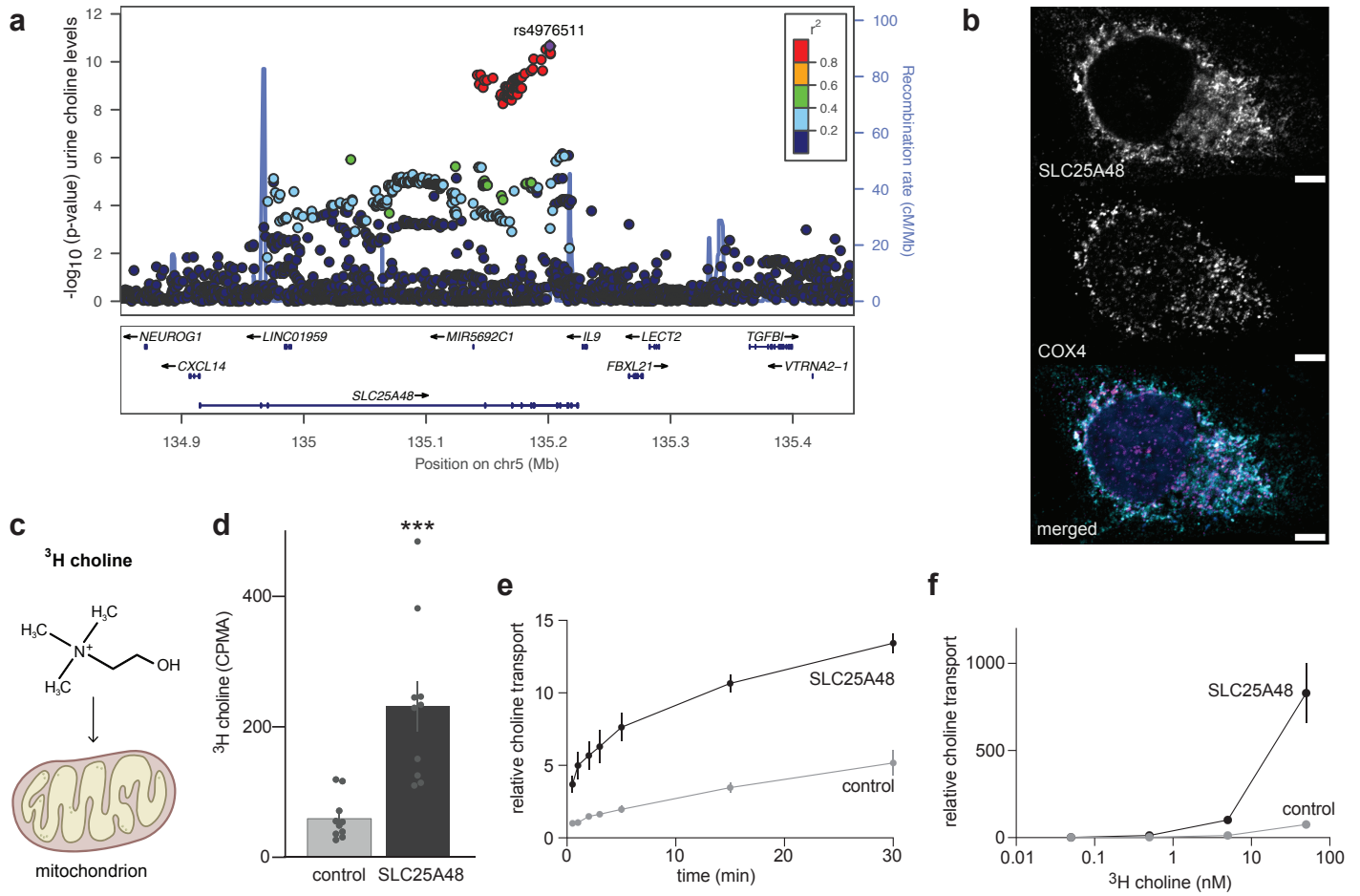593

**Figure 1**



a



b



c

$^3$H choline



mitochondrion

d



e



f

# Figure 2



**a**

choline levels (SD)

Matrix: plasma, urine

Carriers of driver variants in SLC25A48: no, yes

**b**

frameshift, missense, splice_acceptor, stop_gained

−log10(p−value)

p.D27G, p.F64L, p.R179P, p.R243*

Mito_carr, SOLCAR

**c**

kDa

wild type, p.F64L, p.R243*, p.D27G, p.R179P

37
25

SLC25A48-flag

37

actin

**d**

wild type: SLC25A48, COX4, merged

p.R179P: SLC25A48, COX4, merged

**e**

Pearson correlation coefficient

wild type, p.R179P

****

**f**

relative choline transport

wild type, p.F64L, p.R243*, p.D27G, p.R179P

*, ****, ****, ****

**g**

IMS, MATRIX

N, C

F64, R179, D27

~90°

K128, D27, D125, K241, K30, D238