Interplay of human *ABCC11* transporter gene variants with axillary skin microbiome functional genomics

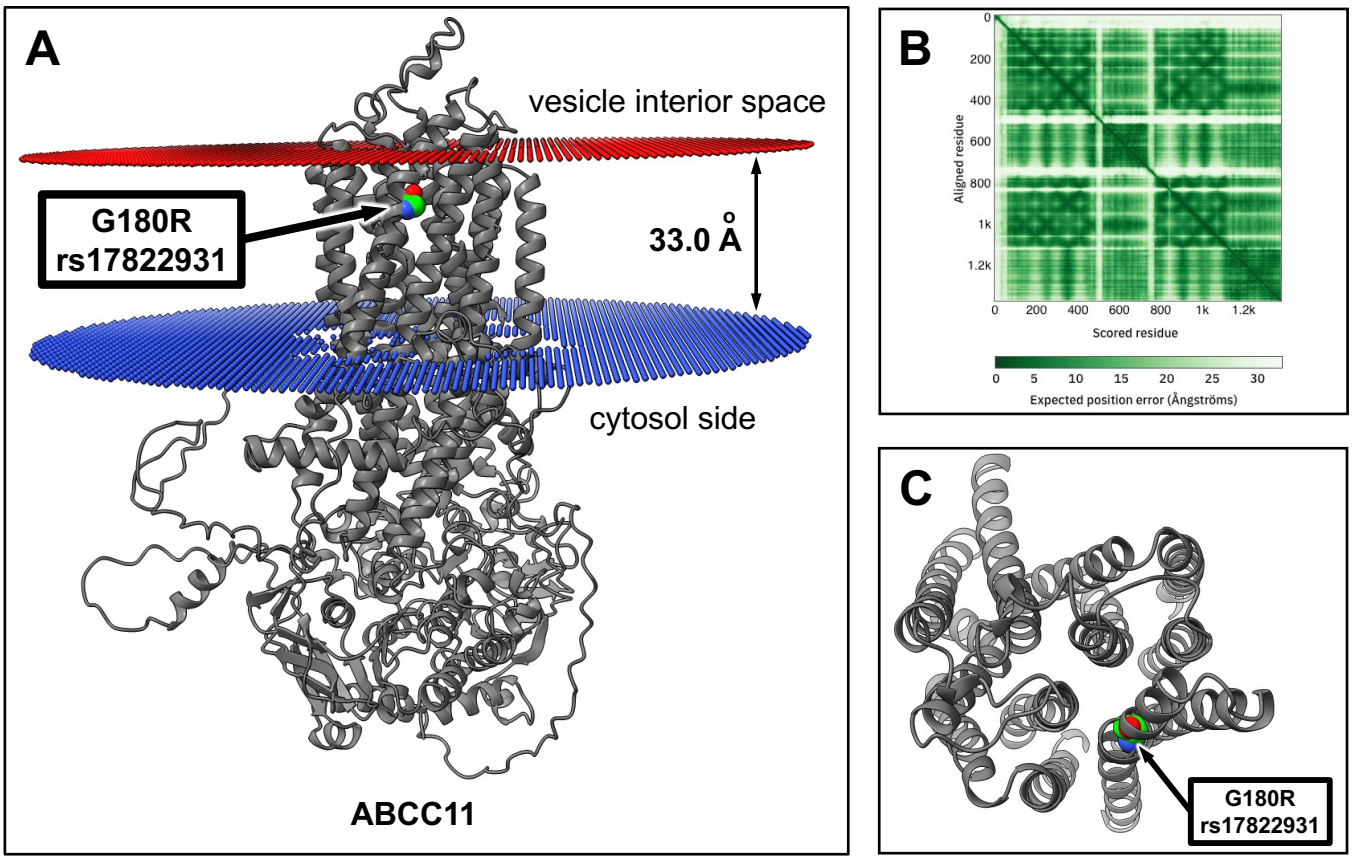
Bruce R. Stevens and Luiz F.W. Roesch

Supplementary Discussion text.

Alternative theoretical explanations for the global patterns of ABCC11 allele distributions.

Among alternative theoretical explanations for the global patterns of ABCC11 distributions, is the notion of genetic convergence. In this scenario, spontaneous *ABCC11* mutations may have arisen sporadically and independently in multiple regions in human populations. For animal orthologs of *ABCC11*, convergence may have played a role in a number of *ABCC11* loss-of-function non-synonymous variant mutations other than rs17822931. For example, ancient Siberian woolly mammoth *ABCC11* genes exhibit five SNP non-synonymous missense variants and one stop-gain variant, each putatively rendering non-functional ABCC11 transporter activity, and these do not occur in modern tropical elephants[1,2]. This was assessed based on DNA extracted from 700,000 year old woolly mammoths that lived in Pleistocene cold (-50^{o}C) winter tundra environments[1,2]. It has been speculated that ABCC11 loss of function putatively rendering dry ear wax phenotype in woolly mammoths may have been a cold climate advantage, complementing concomitant heat conserving DNA variants such as loss of function of certain antecedent mammalian genes resulting in unusually small ears of mammoths in contrast to elephants[2]. In concert with modern elephant *ABCC11*, equatorial non-human primates including chimpanzees, gorillas, and baboons favor the C allele encoding wildtype ABCC11 activity[3,4]. There is some support for this type of convergent selective adaptation in humans, such that CC and CT haplotypes express a wet type earwax with attending clinical complications of impacted auditory canal, middle ear cholesteatoma and attenuated hearing, while TT subjects have dry earwax and lack certain disadvantages of wet cerumen[4-9]. It can be extrapolated that adaptive physiological advantages may be enjoyed by various other tissue functions of the body that express ABCC11, as described above in the Introduction, although this remains to be investigated.

Supplementary Information Figure S1.
Human *ABCC11* SNP residue position p.G180R affecting S-glutathione conjugate transport in intracellular vesicle membrane of apocrine gland cells.



The 3D atomic structure and SNP variant residue positions of human ABCC11 transporter polypeptide have not been reported in the literature. Therefore, we deployed DeepMind AlphaFold[10] to predict the structure based on Uniprot Q96J66 amino acid sequence of human ABCC11 post-translational mature polypeptide[11] (https://www.uniprot.org/uniprotkb/Q96J66/entry). The resulting atomic structure model coordinates (https://alphafold.ebi.ac.uk/entry/Q96J66, [10]) were visualized using ChimeraX 1.6.1 [12]. ABCC11 polypeptide positioning within the membrane of apocrine intracellular vesicles was computed using Orientation of Proteins in Membranes PPM 3.0 Web Server[13]. Residue position p.G180R placement was assessed in the polypeptide structure. Free energy minimizations using TMPfold Server[13] were used to compute the 3D structure of the 12 transmembrane alpha helices engaging p.G180R that form a S-glutathione conjugate transporting conduit pore of ABCC11 spanning the membrane bilayer.

(A) Structure of human ABCC11 post-translational mature polypeptide was predicted based on Uniprot Q96J66 using DeepMind AlphaFold[10]. Residue position p.G180R represents the alternative protein expressions of SNP rs17822931 at allele locus c.C538T, and is positioned 10 Å within the transmembrane hydrophobic region embedded in apocrine gland intracellular vesicle membrane, oriented with respect to membrane interfacing with cytosol side (blue) and intravesicular interior space side (red). Membrane thickness 33.0 $\pm$ 0.6 Å and protein positioning were computed using Orientation of Proteins in Membranes PPM 3.0 Web Server[13].

(B) Predicted aligned error of AlphaFold model indicating high degree of interdomain accuracy of the model of Panel ***A***.

(C) Transmembrane helices' assembly forming the ABCC11 transmembrane pore for substrate transport. This is an Alt perspective view of *Panel **A*** looking from the top downward onto membrane surface from the entry location on the cytosol side. The 12 transmembrane alpha helices form a transporting conduit pore that spans the lipid bilayer, with the key p.G180R residue residing on the inner rim surface of helix #1 approximately 10 Å inside this channel mouth from the extracellular membrane surface. The membrane surface and ABCC11 intracellular residues have been masked in this panel for clarity, revealing the conduit pore formed in the center of the structure. The 3D structure of the pore accommodates apocrine gland metabolite S-glutathione-conjugate metabolite molecules as a ABCC11 transporter substrate, as computed based on free energy minimizations using TMPfold Server[13].

Supplementary Information Table S1.

Rarefaction of microbiome DADA2 amplicon ASVs in human subject samples. Subject wildtype allele is 'C', while mutant non-synonymous *ABCC11* allele 'T' is identified as SNP rs17822931.

Metadata human subjects sample codes	*ABCC11* gene haplotype alleles	ASV Total Reads	Rarefied R Reads	Coverage after rarefaction
Arm11	CT	323943	85486	0.999918115
Arm12	CT	285068	85486	0.999766044
Arm13	CC	256378	85486	0.999883022
Arm14	CC	219993	85486	0.99989472
Arm15	TT	268386	85486	0.999976604
Arm16	TT	279584	85486	0.999976604
Arm17	CC	272100	85486	0.999988302
Arm18	CC	208033	85486	1
Arm19	CT	401091	85486	0.999988302
Arm20	CT	214649	85486	0.999871324
Arm21	TT	307714	85486	0.999976604
Arm22	TT	243553	85486	0.99983623
Arm23	CT	317421	85486	0.999929813
Arm24	CT	378540	85486	0.999941511
Arm25	CT	344620	85486	0.999941511
Arm26	CT	212507	85486	0.999988302
Arm27	TT	298663	85486	0.999964907
Arm28	TT	237601	85486	0.999988302
Arm29	TT	282614	85486	0.999672461
Arm30	TT	194250	85486	0.999742648
Arm31	TT	327192	85486	0.999976604
Arm32	TT	376695	85486	0.999929813
Arm33	TT	309911	85486	0.999918115
Arm34	TT	256614	85486	0.999929813
Arm35	TT	225117	85486	0.999976604
Arm36	TT	321024	85486	1
Arm37	TT	369530	85486	0.999976604
Arm38	TT	345498	85486	0.999847928
Arm39	CT	219399	85486	0.999777741
Arm40	CT	328154	85486	0.99919285
Arm41	CC	356734	85486	0.999824533
Arm42	CC	270128	85486	0.999918115
Arm43	CC	258426	85486	0.999988302
Arm44	CC	286894	85486	0.999684159
Arm45	TT	213161	85486	0.999964907
Arm46	TT	178441	85486	0.999988302
Arm47	TT	251541	85486	0.999941511
Arm48	TT	267637	85486	0.999929813
Arm49	CT	319205	85486	0.999976604
Arm50	CT	366030	85486	0.999976604
Arm51	TT	355781	85486	0.999964907
Arm52	TT	222236	85486	0.999976604
Arm53	CT	193153	85486	0.999988302
Arm54	CT	274269	85486	1
Arm55	CT	213776	85486	1
Arm56	CT	247281	85486	1
Arm57	TT	302315	85486	0.999941511
Arm58	TT	85486	85486	0.999906417
Arm60	CC	321985	85486	0.999883022

References Cited in Supplementary Information.

1. Smith, S.D., Kawash, J.K., Karaiskos, S., Biluck, I., and Grigoriev, A. (2017). Evolutionary adaptation revealed by comparative genome analysis of woolly mammoths and elephants. DNA Res *24*, 359-369. 10.1093/dnares/dsx007.

2. Diez-Del-Molino, D., Dehasque, M., Chacon-Duque, J.C., Pecnerova, P., Tikhonov, A., Protopopov, A., Plotnikov, V., Kanellidou, F., Nikolskiy, P., Mortensen, P., et al. (2023). Genomics of adaptive evolution in the woolly mammoth. Curr Biol *33*, 1753-1764 e1754. 10.1016/j.cub.2023.03.084.

3. Council, S.E., Savage, A.M., Urban, J.M., Ehlers, M.E., Skene, J.H., Platt, M.L., Dunn, R.R., and Horvath, J.E. (2016). Diversity and evolution of the primate skin microbiome. Proc Biol Sci *283*. 10.1098/rspb.2015.2586.

4. Ohashi, J., Naka, I., and Tsuchiya, N. (2011). The impact of natural selection on an ABCC11 SNP determining earwax type. Mol Biol Evol *28*, 849-857. 10.1093/molbev/msq264.

5. Toyoda, Y., Sakurai, A., Mitani, Y., Nakashima, M., Yoshiura, K., Nakagawa, H., Sakai, Y., Ota, I., Lezhava, A., Hayashizaki, Y., et al. (2009). Earwax, osmidrosis, and breast cancer: why does one SNP (538G>A) in the human ABC transporter ABCC11 gene determine earwax type? FASEB J *23*, 2001-2013. 10.1096/fj.09-129098.

6. Yoshiura, K., Kinoshita, A., Ishida, T., Ninokata, A., Ishikawa, T., Kaname, T., Bannai, M., Tokunaga, K., Sonoda, S., Komaki, R., et al. (2006). A SNP in the ABCC11 gene is the determinant of human earwax type. Nat Genet *38*, 324-330. 10.1038/ng1733.

7. Toyoda, Y., Gomi, T., Nakagawa, H., Nagakura, M., and Ishikawa, T. (2016). Diagnosis of Human Axillary Osmidrosis by Genotyping of the Human ABCC11 Gene: Clinical Practice and Basic Scientific Evidence. Biomed Res Int *2016*, 7670483. 10.1155/2016/7670483.

8. Nakagawa, H., Toyoda, Y., Albrecht, T., Tsukamoto, M., Praetorius, M., Ishikawa, T., Kamiya, K., Kusunoki, T., Ikeda, K., and Sertel, S. (2018). Are human ATP-binding cassette transporter C11 and earwax associated with the incidence of cholesteatoma? Med Hypotheses *114*, 19-22. 10.1016/j.mehy.2018.02.030.

9. Nakano, M., Miwa, N., Hirano, A., Yoshiura, K., and Niikawa, N. (2009). A strong association of axillary osmidrosis with the wet earwax type determined by genotyping of the ABCC11 gene. BMC Genet *10*, 42. 10.1186/1471-2156-10-42.

10. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583-589. 10.1038/s41586-021-03819-2.

11. The_UniProt_Consortium (2022). UniProt: the Universal Protein Knowledgebase in 2023, https://www.uniprot.org/uniprotkb/Q96J66/entry . Nucleic Acids Research *51*, D523-D531. 10.1093/nar/gkac1052.

12. Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H., and Ferrin, T.E. (2023). UCSF ChimeraX: Tools for structure building and analysis. Protein Science *32*, e4792. https://doi.org/10.1002/pro.4792.

13. Lomize, M.A., Pogozheva, I.D., Joo, H., Mosberg, H.I., and Lomize, A.L. (2011). OPM database and PPM web server: resources for positioning of proteins in membranes. Nucleic Acids Research *40*, D370-D376. 10.1093/nar/gkr703.