# Adaptive archaic introgression related to cellular zinc homeostasis in humans

Ana Roca-Umbert[1,†], Jorge Garcia-Calleja[1,†], Marina Vogel-González[2], Alejandro Fierro-Villegas[2], Anja Bosnjak[2], Gerard Ill-Raga[2], Víctor Herrera-Fernández[2], Gerard Muntané[1,3,4], Felix Campelo[5], Rubén Vicente[2 *], Elena Bosch[1,4 *]

[1]Institut de Biologia Evolutiva (UPF-CSIC), Departament de Medicina i Ciències de la Vida, Universitat Pompeu Fabra, Parc de Recerca Biomèdica de Barcelona, Doctor Aiguader 88, 08003 Barcelona, Spain

[2]Laboratory of Molecular Physiology, Department of Medicine and Life Sciences (MELIS), Universitat Pompeu Fabra, 08003 Barcelona, Spain

[3]Hospital Universitari Institut Pere Mata, IISPV, Universitat Rovira i Virgili, 43206 Reus, Spain.

[4]Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), 43206 Reus, Spain

[5]ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, 08860 Barcelona, Spain

†Co-first authors

*Corresponding authors

Elena Bosch: elena.bosch@upf.edu

Ruben Vicente: ruben.vicente@upf.edu

**Abstract**

*SLC30A9* encodes a ubiquitously zinc transporter (ZnT9) and has been consistently suggested as a candidate for positive selection in humans. However, no direct adaptive molecular phenotype has been demonstrated to date. Our results provide evidence for directional selection operating in two major complementary haplotypes in Africa and East Asia. These haplotypes are associated with differential gene expression but also differ in the Met50Val substitution (rs1047626) in ZnT9, which we show was adaptively introgressed from archaic humans. Although we found no significant differences in systemic zinc content between individuals with different rs1047626 genotypes, we demonstrate that the expression of the derived isoform (ZnT9 50Val) in HEK293 cells shows a gain of function when compared with the ancestral (ZnT9 50Met) variant. In particular, the ZnT9 50Val variant was found to avoid zinc overload in the endoplasmic reticulum and mitochondria, with an impact on mitochondrial metabolism. Overall, our results show that this adaptatively introgressed variant, which is prevalent in East Asians and present at intermediate-high frequencies in other non-African populations, is associated with functional differences in zinc handling by the mitochondria and endoplasmic reticulum. Given the essential role of the mitochondria in both skeletal muscle and brain, and that of zinc in glutamatergic neurotransmission, we propose that archaic adaptation to cold may have driven this selection event, while also impacting susceptibility to neuropsychiatric disorders in modern humans.

## Introduction

How adaptation has shaped current genetic diversity in human populations is a long-standing question in evolutionary genetics. In turn, it has been shown that the identification and functional deciphering of adaptive variants in our genome can reveal biologically relevant variation with important phenotypic consequences in current populations (Vitti et al. 2013; Rees et al. 2020). Recent access to whole-genome sequencing data from diverse human populations coupled with improvements in methods that detect the molecular signatures of natural selection in the genome have facilitated the identification of hundreds of candidate genes for positive selection in humans (Akey 2009). As a result, several candidate regions for selection have been successfully associated with diverse adaptive phenotypes and their corresponding selective pressures, which were mostly related to our diet, immune response, high altitude environment, and UV radiation (Lachance and Tishkoff 2013; Fan et al. 2016; Rees et al. 2020). However, our understanding of the adaptive phenotypes and functional variants underlying most of the signatures of positive selection in the human genome is still limited.

Zinc is an essential micronutrient with different structural, catalytic, and regulatory roles in the human body (Vallee and Falchuk 1993; International Zinc Nutrition Consultative Group (IZiNCG) 2004; Kambe et al. 2015). As adequate zinc levels are fundamental for maintaining good health, the systemic and cellular homeostasis of zinc must be tightly regulated. The proteins controlling the import and export of zinc across cell membranes are known as Zinc Transporters (ZT) (Kambe et al. 2015; Hara et al. 2017). The 24 ZTs identified in humans can be classified in two families: the ZIP family, whose members import zinc into the cytosol, and the ZnT family, whose function is to export zinc outside the cell or into the cell organelles. The 14 ZIP transporters are encoded by the *SLC39A1-14* genes, whereas the 10 ZnT transporters are encoded by the *SLC30A1-10* genes (Fukada and Kambe 2011; Huang and Tepaamorndech 2013; Jeong and Eide 2013; Kambe et al. 2015). Variation in this set of 24 ZT encoding genes can lead to zinc dysregulation at a cellular or systemic level and, given the fundamental role of zinc in the human organism, such zinc imbalance might cause alterations in distinct phenotypes upon which adaptive selection could act (Zhang et al. 2015). In this context, a non-synonymous substitution at the *SLC39A4* gene, which encodes the most important intestinal zinc uptake transporter (ZIP4), has been shown to produce differential cellular zinc uptake and presents extreme population differentiation because of a local positive selection event in Sub-Saharan Africa (Barreiro et al. 2008; Engelken et al. 2014).

The *SLC30A9* gene has been repeatedly reported as a top candidate region for positive selection in several genome-wide scans of selection, mostly in East Asian populations (Carlson et al. 2005; Sabeti et al. 2007; Williamson et al. 2007; Barreiro et al. 2008; Pickrell et al. 2009; Grossman et al. 2013). This gene encodes the ZnT9 protein, a recently characterized ZT in the mitochondria (Deng et al. 2021; Ma et al. 2021). Moreover, the patterns of genetic variation and population differentiation at *SLC30A9* have also been investigated in candidate-based selection studies

analyzing human adaptation in relation to zinc homeostasis (Zhang et al. 2015; Engelken et al. 2016; Roca-Umbert et al. 2022). In most of these previous studies, the adaptive signals in *SLC30A9* have been attributed to a non-synonymous single nucleotide polymorphism (SNP) (rs1047626; c.148A>G) resulting in the substitution of a methionine by a valine in the N-terminus of the ZnT9 transporter (Met50Val) (Grossman et al. 2013; Zhang et al. 2015; Engelken et al. 2016; Roca-Umbert et al. 2022). In agreement with the strong signatures of recent positive selection detected in East Asia, this non-synonymous SNP presents extremely high levels of population differentiation, where the ancestral A-allele (encoding Met) is at high frequency in African populations, whereas the derived G-allele (encoding Val) is nearly fixed in East Asian populations and at intermediate frequencies in European, South Asian, and American populations. Moreover, contrasting extended haplotypes and selection signals along the *SLC30A9* region have been reported between Africans and East Asians (Zhang et al. 2015). Furthermore, the derived G-allele has been shown to correlate with the distribution of human zinc deficiency worldwide (Zhang et al. 2015) and to be in moderate linkage disequilibrium with three variants in the 3' flanking region of the *SLC30A9* gene, which are QTLs influencing zinc content in the liver (Engelken et al. 2016). However, to date, no experimental validation has been performed to demonstrate the functional relevance of this putative adaptive variant regarding zinc transport.

Here, we first compiled evidence for a selective sweep in the *SLC30A9* gene region, inferred the past allele trajectories and selection coefficients for two putative adaptive SNP variants characterizing the two major *SLC30A9* haplotypes in humans, traced back their genetic origins and then investigated the phenotypical relevance of the Met50Val substitution (rs1047626) in ZnT9. Accordingly, we overexpressed ZnT9 in HEK293 cells and explored whether the two variants of the Met50Val substitution induce differences in the expression levels and localization of ZnT9 and other ZTs, and whether they differentially affect zinc transport in the cytosol, mitochondria, and endoplasmic reticulum (ER). Our results show that the ZnT9 Met50Val substitution sits in an unusual haplotype, which was adaptatively introgressed from archaic humans, and demonstrate relevant differences in mitochondrial and ER zinc homeostasis directly influencing mitochondrial activity and cell viability.

## Results

### *Signals of positive selection outside Africa*

We first explored the genomic signatures of the selective sweep previously reported in the *SLC30A9* gene region (Carlson et al. 2005; Sabeti et al. 2006; Williamson et al. 2007; Barreiro et al. 2008; Pickrell et al. 2009; Grossman et al. 2013). Hence, we compiled evidence from different summary statistics and selection tests available at the PopHuman browser for the 26 human populations from the 1000 Genomes Project (1000 GP; http://pophuman.uab.cat) (Casillas et al. 2018). To summarize and visualize the continental patterns of positive selection along the

*SLC30A9* region, we extracted the $F_{ST}$, XP-EHH, and Fay and Wu's H values in three populations used as a reference for Africa (Yoruba: YRI), Europe (Residents of Utah with North and Western European ancestry: CEU), and Asia (Han Chinese: CHB) (Fig. 1A). Unusual high levels of population differentiation were confirmed along the *SLC30A9* region when comparing CEU and CHB with YRI. Similarly, when applying the XP-EHH statistic, we detected genome-wide departures for long haplotypes in CEU and above all in CHB (using YRI as the reference). Furthermore, a significant excess of derived alleles was detected around the *SLC30A9* region in both CHB and CEU populations when analyzing the site frequency spectrum with the Fay and Wu's H statistic. Notably, most of these signatures of selection available at the PopHuman browser replicate for the remaining 1000 GP populations in all continents except Africa.

We next investigated which variants along the *SLC30A9* gene could underlie the detected signals of selection (see complete list in Table S1). From a total of 354 biallelic SNPs with MAF ≥ 0.02 in *SLC30A9* gene, we initially selected as candidates those variants with high population differentiation when comparing YRI with CEU and CHB. From the resulting 235 variants, we further considered as putatively adaptive those that were either non-synonymous, present at the UTR regions of the *SLC30A9* gene, or had a CADD Phred score ≥ 10, as such a value predicts that the variant falls within the top 10% of the most deleterious (i.e., functional) variants of the genome (Kircher et al. 2014). A total of 18 SNPs matched our filtering criteria (Table 1). Among these, we identified a non-synonymous SNP, two synonymous variants, five variants at the 3' UTR, and ten intronic variants. According to the GTEX portal (https://www.gtexportal.org/home/; last accessed 13/04/2022), all 18 putative adaptive variants are eQTLs, with the most common allele in CHB and CEU associated with reduced *SLC30A9* expression mostly in brain (Fig. S1A and File S1).

We then took advantage of the previously inferred genome-wide genealogies available for the 1000 GP individuals and used Relate (Speidel et al. 2019) to explore whether the lineages carrying the derived allele of the Met50Val substitution had spread at an unusually fast rate in comparison with competing lineages in the YRI, CHB and CEU populations. The marginal tree corresponding to the surrounding region of rs1047626 displayed the Val-derived variant in multiple lineages concentrated in CHB and CEU, although it was also present in some YRI haplotypes (Fig. S2). In contrast, most of the lineages carrying the ancestral variant were mainly found in YRI. The subsequent analysis with CLUEs estimated a drastic allele frequency shift compatible with the action of positive selection acting on rs1047626 at ~56-77 kya (i.e., 2,000-2,750 generations before the present; using 28 years per generation) with selection coefficients s=0.0113 (logLR=5.68) and s=0.0017 (logLR=10.93) in CHB and CEU, respectively (Fig.1B). When analyzing other non-African populations such as FIN (European), JPT (East Asian), PJL (South Asian) and PEL (American), we obtained similar allelic trajectories and selection coefficients (Fig. S3). Thus, we seem to be detecting a strong signal of positive selection common to all non-African populations probably dating back before the expansion of modern humans though Europe and Asia.

*Archaic introgression*

We next explored whether such an early selection event could be explained by adaptive introgression. For that, we first investigated whether the Met50Val substitution was present in publicly available Neanderthal and Denisovan genomes. Remarkably, the Denisovan genome presents the derived allele at the rs1047626 position, whereas three high coverage Neanderthal genomes (Altai, Vindija, and Chagyrskaya) were found to carry the ancestral allele. We then analyzed the corresponding Neanderthal and Denisovan base positions for a total of 170 SNPs found in high LD ($r^2$>0.8) with rs1047626 in CHB and CEU and spanning a 98.74 kb region (Table S2). For 164 of these SNPs (96%), the most frequent allele in EAS, EUR, SAS, and AMR differs from the most prevalent in AFR, confirming the existence of two major contrasting Yin/Yang human haplotypes. On the corresponding 98 covered SNP positions, the Neanderthal and Denisovan genomes were found to display higher allelic similarity and greater sharing of derived alleles with the allelic combinations most frequently found in non-Africans (Fig. 1C; Table S2). Given these results, we next investigated whether this region had been identified in genome scans of Neanderthal and Denisovan introgression into modern humans. No hit was found within the most probable candidate genomic regions for Neanderthal introgression (Sankararaman et al. 2016; Browning et al. 2018; Skov et al. 2020). However, two independent studies recently reported two overlapping candidate regions for Denisovan introgression in Melanesians (Gower et al. 2021) and Papua New Guineans (Hubisz et al. 2020), respectively (Fig.1C; Table S2 and Fig. S4). The two most distant SNP positions where the archaic genomes were found to carry derived alleles shared with the most frequent allele combinations in present-day humans and coinciding with at least one of these introgression maps delimitate a region of 70.61 kb (hg38, chr4: 41,977,828-42,048,441). We next explored the haplotype structure of this putatively introgressed region using all the SNP variants in the genomes of the CHB, CEU and YRI populations from the 1000GP (The 1000 Genomes Project Consortium 2015) and the Oceanian individuals available in the HGDP dataset (Bergström et al. 2020) presenting a homozygous genotype in the Denisovan, together with that of the corresponding nucleotide positions in the three high-coverage Neanderthal genomes previously used (see details in Materials and Methods). The resulting set of 112 SNPs defined a total of 61 haplotypes among CHB, CEU and YRI, as well as 4 archaic haplotypes (Fig. S5 and Table S3). Up to 78 out of these 112 SNPs (including rs1047626) display a derived allele at high frequency in CHB, CEU and Oceanians shared with the Denisovan, while the total number of derived alleles shared between the three Neanderthal individuals (which all present the ancestral allele for rs1047626) and the Denisovan is 72 (Table S3). By contrast, the most frequent allelic configuration found in YRI (ht I) only presents 11 derived alleles shared with archaic humans. Accordingly, the haplotype network displays the Denisovan haplotype 12 mutational steps away from the most frequent haplotype in CHB and CEU (ht II, found at a haplotype frequency of 95% and 67% in

these populations, respectively) and between 11 to 13 steps to three major haplotypes in Oceanians (hts XXIII, LX, LXI, which represent 95% of the Oceanians chromosomes) (Fig. 1D and Table S4). Furthermore, two of the Neanderthal haplotypes are found less than 5 mutational steps from hts XL (present in 2 YRI chromosomes) and IX (present in 3 CEU, 3 YRI and 2 OCE chromosomes) while displaying less than 14 nucleotide differences to ht II and the three predominant haplotypes in Oceanians, and more than 16 nucleotide substitutions to the Denisovan (Table S4). By contrast, the most prevalent haplotype in YRI (ht I, present in up to 53% of the YRI chromosomes but also found in CEU and CHB at frequencies of 13% and 2%, respectively) lies in a long branch of the network more than 60 substitutions away from ht II and those of Neanderthals and the Denisovan (Table S4, Fig. 1D). Furthermore, up to 24 additional minor haplotypes, covering 26.85% of the remaining YRI chromosomes, were found less than 10 nucleotide differences away from ht I. Among these, we find the most ancestral allelic configuration in ht XXV, which only presents 2 derived alleles shared with archaic humans. Overall, the short nucleotide pairwise distances found between the Denisovan and several closely related haplotypes found at high frequencies among CHB, CEU and Oceanians (together with the high number of shared derived alleles among them) contrast with the long pairwise distances between most African and non-African human haplotype comparisons (and the lower number of shared archaic derived alleles with the most frequent allelic combinations in YRI). Similarly, the three Neanderthal haplotypes display unexpected high similarity to a few modern human haplotypes although in this case they were found at much lower frequencies and in all populations (including YRI). Altogether, this pattern clearly differs from the known phylogeny of archaic and modern humans but could be compatible with archaic introgression or incomplete lineage sorting (ILS). However, the probability of observing one haplotype of 70,614 bp in length shared by ancestry without being broken by recombination given the local recombination rate of the region (1.58 cM/Mb) and the time since divergence between modern and archaic humans (see details in Materials and Methods) is highly unlikely (p-value $< 2.07 \times 10^{-8}$ when considering a mutation rate $\mu = 1 \times 10^{-9}$; Table S5) and thus led us reject the ILS scenario. Therefore, the most likely scenario for the observed haplotype configuration is introgression.

*Signals of positive selection in Africa*

As signatures of positive selection had been previously reported for the most abundant haplotype around *SLC30A9* in YRI (Zhang et al. 2015), we also searched for putative candidate variants displaying unusual past allelic trajectories in Africa. Notably, several derived allele variants characterizing the African-like haplotype of *SLC30A9* were found to show marginal genome-wide p-values for positive selection with Relate (Table S2). After their corresponding functional annotation (Tables S6-9), we selected one of the strongest eQTLs (i.e., rs4861157) for further analysis with CLUEs, which estimated a selection coefficient s=-0.00425 and a logLR=16.2401 for

an adaptive event occurring within the last 1,000 generations (i.e., 28,000 years before present) in the YRI population (Fig. 1B, Figs. S6-S7). Although the genotype for rs4861157 was not available in the high coverage archaic genomes and therefore it was not included in the computation of the haplotype network, we checked that its derived allele (A) is found in all chromosomes presenting either the most prevalent haplotype in YRI (ht I) or their corresponding derived haplotypes in the network (Fig. 1D). These results confirm a pattern of contrasting directional selection between African and non-African populations acting on two major differentiated Yin/Yang haplotype configurations at *SLC30A9*. Moreover, as expected, the derived allele of rs4861157 has high frequencies in all African populations from the 1000GP and associates with higher *SLC30A9* expression (Figs. S1B and S8).

*Molecular phenotype for Met50Val*

We then investigated whether the Met50Val substitution in ZnT9 could determine any further differential molecular phenotype besides that of gene expression to explain this case of putative adaptative archaic introgression outside Africa. We first verified the predicted protein structure of ZnT9 using the AlphaFold algorithm. We observed that the polymorphic amino acid residue affected by rs1047626 is located at the cytosolic N-terminus domain of the ZnT9 transporter (Fig. 2A-B). It has been suggested that the N-terminus of the ZnT family modulates the transporter activity by facilitating either protein interactions or zinc binding (Eide 2006; Xue et al. 2020). Thus, we proceeded to analyze whether the Met50Val substitution could cause any differential phenotype on the ZnT9 protein expression levels, cellular location, or zinc homeostasis at cellular level. Accordingly, we used a heterologous expression system in HEK293 cells overexpressing each of the two ZnT9 forms resulting from the Met50Val substitution, carrying either the ancestral allele (ZnT9-50Met) or the derived allele (ZnT9-50Val). Our experiments showed no difference in protein expression between the ZnT9-50Met and ZnT9-50Val variants after their corresponding 24h transient expression (Fig. 2C). We next characterized the localization of the two ZnT9-50Met and ZnT9-50Val variants. For that, we performed immunostaining in non-permeabilizing and permeabilizing conditions. The results, shown in Fig. 2D, indicate that the ZnT9 N-terminus is located intracellularly and follows a reticular pattern independently of the overexpressed variant. We further explored the subcellular distribution of the variants co-transfecting the ZnT9 constructs with the FEMP probe, a double fluorescent reporter plasmid for the ER (in CFP) and the mitochondria (in YFP). Our colocalization analysis gave similar results for the two ZnT9 variants, both showing a higher colocalization degree with the ER compartment (Fig. S9). We then used super-resolution STED microscopy to confirm the presence of ZnT9 in mitochondria as described by others (Deng et al. 2021; Ma et al. 2021) (Fig. 3). ZnT9-HA was co-stained with anti-TOM20 (Fig. 3A), a mitochondrial outer membrane resident protein. Both ZnT9-50Met and ZnT9-50Val variants showed co-localization with mitochondria. Remarkably, by analyzing the mean distance

between mitochondria (Fig. 3B) and the relative area occupied by mitochondria in the cell (Fig. 3C), we observed that the overexpression of ZnT9 generated mitochondrial clustering in comparison with non-transfected cells. To check whether the distance between ER and mitochondria might be altered we carried out FRET analysis using the FEMP probe together with the ZnT9 variants or empty vector. Our results showed no differences in any of the tested conditions (Fig. 3E). Overall, this initial characterization revealed no differences in cellular localization for either variant.

We then studied the impact of ZnT9 overexpression on cellular zinc homeostasis, comparing cells transiently transfected with either of the ZnT9 variants or an empty vector. First, the mRNA expression of the different ZIP transporters was characterized to analyze whether the overexpression of ZnT9, a transporter reported to promote cytosolic zinc efflux, is somehow compensated for by a modification in a zinc importer expression (Fig. 4A). Our results showed alterations in ZIP12 and ZIP14 expression, although in the case of ZIP12 CT values were above 30 and ZIP14 displayed a reduction in cells overexpressing either of the two ZnT9 variants. We then used flow cytometry with Zinquin to measure the cytosolic zinc content in basal conditions or with excess zinc for 30 min and observed no differences in cytosolic zinc handling between vectors in the two tested conditions (Fig. 4B).

Given the presence of the ZnT9 transporter in ER membranes, we measured the endoplasmic zinc content using ER-ZapCY1 probe. No major alterations were observed in cells transfected with ZnT9-50Val, ZnT9-50Met, or an empty vector in basal conditions. However, upon incubation with 100 µM external zinc medium, only the cells transfected with ZnT9-50Met, or an empty vector had an increased ER zinc content, indicating that the overexpression of ZnT9-50Val was modifying the ER zinc homeostasis (Fig. 4C). When analyzing the RNA expression of the known ER zinc importers ZnT5, ZnT6 and ZnT7, we found that the overexpression of ZnT9-50Val causes a downregulation of ZnT5 and ZnT6 expression (Fig. 4D). Considering that ZIP7 is the major zinc exporter in the ER, we further explored its activity in conditions of ZnT9 overexpression. After 24h of transient expression, the results showed no difference in ZIP7 protein expression between the ZnT9 variants (Fig. S10A). We also measured the ER zinc content in the presence of a ZIP7 blocker, finding a strong increase in the fluorescence of the ER zinc reporter. However, differences between basal and 100 µM external zinc conditions were no longer observed, nor between ZnT9 variants and the empty vector (Fig. 4E), indicating that in the absence of ZIP7 activity the ER suffers a zinc overload.

As ZnT9 has been shown to regulate mitochondrial zinc homeostasis (Deng et al. 2021; Ma et al. 2021), we also interrogated the functional relevance of the Met50Val substitution by measuring the zinc content in the intermembrane space and the mitochondrial matrix, using SMAC-mCherry-GZnP2 and mito-Cherry-GZnP2 probes, respectively. In the intermembrane space, overexpression of ZnT9-50Val resulted in a slightly higher zinc content than cells overexpressing ZnT9-50Met or

an empty vector. However, zinc increased equally in all conditions upon incubation with 100 µM zinc (Fig. 4F). In the mitochondrial matrix, ZnT9-50Val-transfected cells had a higher zinc content in basal conditions as well but that was not further affected by incubating with 100 µM zinc, contrary to cells overexpressing ZnT9-50Met, or an empty vector (Fig. 4G). Given that mitochondrial zinc content is reported to depend on ER zinc transport (Deng et al. 2021; Ma et al. 2021), similar experiments were carried out, but this time blocking the ER zinc exporter ZIP7 for 30 min to accumulate zinc in this organelle. The result was an increase of zinc levels in the mitochondrial matrix of control cells but not in ZnT9 overexpressing cells (Fig. S10B). To gain further insight into the ZnT9 transport activity in the mitochondria, ZnT9 was silenced with siRNA and the zinc content of the mitochondrial matrix was measured. As the data showed that incubation with 100 µM zinc increases zinc levels despite silencing ZnT9, we discarded the possibility that ZnT9 is a major zinc importer in the mitochondria (Fig. S11A). Our expression data confirmed the silencing strategy and showed no alterations in the expression of ZnT5, the zinc importer in the ER, when ZnT9 was silenced (Fig. S11B). Finally, we also investigated the impact of the ZnT9 variants overexpression on mitochondrial metabolism. MTT experiments were carried out in cells incubated for 24h with 0, basal and 100 µM zinc content. Higher metabolic activity signal was found in ZnT9-50Val-overexpressing cells both in basal and zinc excess conditions compared with the control (Fig. 4H). Considering that at basal conditions cell number was not altered after 24h transfection with different plasmids (Fig. S12), MTT assays are revealing differences in mitochondrial metabolism. The absence of difference in the 0 zinc media supports the idea that the impact on metabolism is due to a zinc-dependent process.

*Exploring an adaptive phenotype for Met50Val*

As three SNPs in the 3' flanking region of the *SLC30A9* gene (rs2880666, rs6447133, rs7659700) have been previously identified as QTLs influencing zinc content in the liver (Engelken et al. 2016) and they are found in moderate linkage disequilibrium ($0.66 < r^2 < 0.76$) with rs1047626 (Table S10), we examined whether rs1047626 could be associated with such a systemic phenotype. No significant differences in liver zinc content were detected between individuals carrying A- or G-alleles (p-value=0.107), nor between the AA, AG, and GG genotypes at rs1047626 (p-value = 0.322). However, in agreement with the known effects of the surrounding nutriQTLs with which the alleles at rs1047626 associate, the homozygous individuals for the G-allele showed a tendency towards higher zinc concentrations in the liver than those homozygous for the A-allele, whereas the heterozygotes AG were associated with an intermediate zinc content (Table S11, Fig. S13). Additionally, we searched for GWAS annotations linked to any of the 170 SNPs in high linkage disequilibrium ($r^2 > 0.8$) with rs1047626. We found reported GWAS hits for helping behavior (rs1507086, rs2660319, rs11051), total PHF-tau SNP interaction (rs4861153), and one SNP (rs34215985) associated with multiple psychiatric phenotypes, such as anorexia nervosa,

hyperactivity disorder, autism spectrum disorder, bipolar disorder, major depression, obsessive compulsive disorder and schizophrenia (Table S2). A phenome-wide association study (PheWAS) for rs1047626 retrieved seven entries with a corrected p-value $\leq 1.5 \times 10^{-5}$ (0.05/3,302 unique traits), including traits within the psychiatric (i.e., neuroticism and major depressive disorder, being the derived G-allele of the rs1047626 polymorphism the allele that increases the risk of both phenotypes), metabolic (i.e., impedance, with the G-allele associated with lower impedance measures), activities (i.e., fish oil, and glucosamine uptake, with the G-allele increasing the phenotype value) and skeletal (i.e., height, with the G-allele decreasing the phenotype value) domains, respectively (Table S12).

**Discussion**

The motivation for this study was the previous detection of strong signatures of positive selection in the *SLC30A9* region (Carlson et al. 2005; Sabeti et al. 2006; Williamson et al. 2007; Barreiro et al. 2008; Grossman et al. 2013; Zhang et al. 2015). In this context, we first re-examined the genetic evidence for a hard-selective sweep using the $F_{ST}$, XP-EHH, and Fay and Wu's H selection statistics and found several functional variants on *SLC30A9* that could explain the strong adaptive signals detected. In accordance with Zhang et al. (2015), we confirmed contrasting signals of positive selection between two major haplotypes in Africa and East Asia, the latter also being present at intermediate-high frequencies in South Asia, Europe, and America. The major African haplotype carries alleles at several eQTLs associated with higher *SLC30A9* expression. In contrast, the most prevalent allelic combinations outside Africa are associated with lower gene expression and are linked to the derived allele of the rs1047626 polymorphism, which causes a Met to Val substitution at ZnT9 codon position 50. Although no experimental characterization of its corresponding molecular phenotype has been reported to date, previous studies have attributed the strong signals of adaptation found in East Asians to this non-synonymous variant (Grossman et al. 2013; Zhang et al. 2015; Engelken et al. 2016). In this work, we describe the presence of the derived Val allele in the Denisovan genome and an unusually high sharing of derived alleles between the major haplotype outside Africa and archaic humans (both Neanderthals and Denisovans). Moreover, we demonstrate that the persistence of such similar allelic configuration around the Met50Val substitution comprising a segment of 70,614 bp in length between modern and archaic humans cannot be explained by ILS when considering the recombination rate of the *SLC30A9* region. These results are compatible with two independent studies reporting a Denisovan adaptive introgressed region in modern Melanesians (Gower et al. 2021) and Papua New Guineans (Hubisz et al. 2020) overlapping with the *SLC30A9* gene and comprising the rs1047626 polymorphism. However, we also describe high similarity between (at least two) Neanderthal haplotypes and other much less frequent modern human haplotypes found in all four populations interrogated (including YRI). The presence of such Neanderthal-like haplotypes in Africa could

result from a back migration event (Chen et al. 2020) and it might have easily difficulted the detection of the *SLC30A9* gene in previous maps of Neandertal introgression in modern humans. The protein structure of ZnT9 predicted by AlphaFold (Jumper et al. 2021) shows a total of 6 transmembrane domains and a cytosolic N-terminus for ZnT9, a structure similar to that predicted for the other members of the ZnT family. The location of the Met50Val substitution in the N-terminus may indicate a putative effect of this non-synonymous SNP on zinc homeostasis through a possible accessory role of the N-terminus domain of ZnT9 in facilitating zinc transport or protein interactions. When exploring the consequences of the Met50Val substitution at the molecular level, the protein expression analyses revealed a similar partial localization of both ZnT9 variants in the ER and mitochondria. Moreover, we observed mitochondrial clustering near the ER when ZnT9 was overexpressed. The presence of the transporter in both cellular compartments has been previously described (Kowalczyk et al. 2021), and the ionic fluxes between the two organelles are highly interconnected. In this context, it has been recently described that the ZnT9KO phenotype can be rescued by targeting transporters in the ER and mitochondria (Ma et al. 2021). The ER plays a major role in protein folding, maturation, quality control, and trafficking. Many of these processes require zinc to function correctly, and its deficit or excess often leads to protein malformations that induce ER stress. The most well-known ZTs implicated in ER zinc homeostasis are the ZnT5-ZnT6 heterodimer and ZnT7 homodimer, which are involved in zinc importation (Suzuki et al. 2005; Ishihara et al. 2006), and ZIP7, which is the main zinc-releasing transporter in the ER (Taylor et al. 2012). Considering the effect on ER zinc homeostasis of the overexpression of either ZnT9 isoform, we propose that the 50Val variant avoids ER zinc accumulation due to the reduced ZnT5-ZnT6 heterodimer expression. Additionally, we observed that ZIP7 inhibition generates an excess of zinc in the ER, supporting the idea that the zinc exporting activity of ZIP7 is key for maintaining endoplasmic zinc homeostasis. Therefore, it is likely that the observed absence of differences in ER zinc content in basal conditions is due to the master role of ZIP7 in fine-tuning ER zinc content.

It has been proposed that ZnT9 in the mitochondria together with ZnT5 in the ER determine the zinc content in the mitochondrial matrix (Ma et al. 2021). ZnT9 has been identified as a zinc exporter from the mitochondria based on the zinc accumulation observed in the mitochondria in ZnT9KO models (Deng et al. 2021; Kowalczyk et al. 2021; Ma et al. 2021). We only detected a mild increase in zinc in ZnT9-silenced cells, which was not statistically significant. Probably, the zinc accumulation observed in KO models is the result of a complete and chronic effect. Nevertheless, this transport model implies that ZnT9 works in the opposite direction from the rest of the ZnT transport family. Our results showed that ZnT9 50Val expression increases the zinc content of the mitochondrial matrix but avoids zinc overload. The intermembrane space, however, showed no differences at high zinc concentrations between the tested conditions. Zinc homeostasis in this space has been previously reported to follow cytosolic zinc concentrations (Fudge et al.

2018). On the other hand, it is likely that the repression of ZnT5 expression observed in ZnT9 50Val cells might be a negative feedback to avoid a further zinc increase in the mitochondrial matrix. Having ruled out that ZnT9 is the main mitochondria zinc importer, we speculate that it might influence a wider machinery of proteins involved in mitochondrial zinc homeostasis. In this respect, ZnT9 has been recently connected to the expression of mitochondrial ribosomes and OxPhos components (Rensvold et al. 2022). Importantly, we show that the ZnT9 50Val variant increases mitochondrial metabolism in a zinc-dependent manner. ZnT9 has been previously reported to affect mitochondrial function (Deng et al. 2021; Ma et al. 2021) and in this work we show that it also affects zinc homeostasis in the ER. Both organelles are interconnected and essential for cell fate and metabolism (Prinz et al. 2020).

Previously, a mutation in the fourth TMD in *SLC30A9* was found to cause acute dysregulation of zinc homeostasis in all family members affected by a novel cerebro-renal syndrome with important early neurological deterioration (Perez et al. 2017). Given that the Met50Val substitution sits in what is expected to be an auxiliary region for the activity of the transporter, we hypothesize that the effects of rs1047626 on zinc homeostasis are milder, but somehow became adaptive in the past. Accordingly, when analyzing whether the genotype for rs1047626 influenced zinc content in the liver as a proxy for systemic zinc homeostasis, we only detected a tendency towards increased zinc in the liver when the derived G-allele (50Val) was present. Even if larger sample sizes might produce results with significant differences, any putative effect of rs1047626 on liver zinc content would be lower than that of the QTLs previously identified in the 3´region of the *SLC30A9* gene. Although *SLC30A9* is moderately expressed in most tissues, its expression is higher in the fetal brain, cerebellum, skeletal muscle, pituitary, thyroid, and kidneys (Perez et al. 2017). Hence, greater effects of the Met50Val substitution on other tissues or organs cannot be ruled out, for instance, related to adjustments in renal excretion, brain function or even skeletal muscle metabolism.

The major haplotype outside Africa was found significantly associated with greater susceptibility to major depression and other related psychiatric disorders (Wray et al. 2018; Coleman et al. 2020) as well as with higher values for self-reported helping behavior (Primes and Fieder 2018). Moreover, our PheWAS analysis confirms that the derived G-allele of rs1047626 contributes not only to a greater risk for both major depressive disorder and neuroticism but also lower impedance (i.e., a higher lean body mass), among others. In the last years it is emerging the crucial function of ZnT9 in core mitochondrial processes. The function of the mitochondrial ribosomes (mitoribosomes) and, consequently, the expression of proteins encoded by mitochondrial DNA, such as several OxPhos subunits, has been shown to depend on ZnT9 expression (Rensvold et al. 2022). In this context, our data proves that the Met50Val substitution influences mitochondrial metabolism. This effect might have an impact on systemic metabolism that could explain the association of rs1047626 with higher lean body mass. On the other hand, zinc accumulates in glutamate synaptic vesicles and modulates

excitability (Takeda et al. 2014). Glutamate excitotoxicity has been shown to be partially mediated by zinc overload in the postsynaptic neuron (Kim et al. 2020), whereas dysfunctional glutamatergic signaling has been associated with neuropsychiatric conditions such as schizophrenia, major depression, and anxiety among others (Gillespie and Ressler 2005; Duman et al. 2019; McGrath et al. 2022). Thus, we do not discard a potential role of the Met50Val substitution in modulating zinc overload derived from excitotoxicity, ischemia, or exacerbated ROS production (Kim et al. 2020), which may help to provide mechanistic links to such associated neurological traits. However, despite all the data we have compiled, it is not obvious how the observed differential molecular phenotype for the Met50Val substitution can be translated into an adaptive phenotype at the organism level or might relate to environmental zinc (Zhang et al. 2015). Given the essential role of the mitochondria in both skeletal muscle and brown adipose tissue thermogenesis, we propose that the derived G-allele, which we have shown provides better protection to the mitochondria, may have been positively selected to facilitate adaptation to cold. This is a plausible scenario for a case of archaic adaptive introgression, as Denisovans and Neanderthals were probably already well adapted to the local environmental conditions modern humans encountered when expanding across Eurasia.

In conclusion, we report that the signatures of positive selection recognized in the *SLC30A9* gene region outside Africa are associated with differential zinc homeostasis in the mitochondria and endoplasmic reticulum mediated by the Met50Val substitution. Although several promising traits have been associated with rs1047626, additional work is required to understand how such a molecular phenotype is refined into an adaptive phenotype at the organism level. Moreover, it cannot be ruled out the possible adaptive role of regulatory variants, influencing both *SLC30A9* expression and liver zinc content (and thus systemic homeostasis of zinc), also contributing to the adaptive phenotype we proposed for the Met50Val substitution. Finally, further work is necessary to ascertain the exact timing, place and admixture events experienced with archaic humans, which could explain the current patterns of variation and contrasting signatures of positive selection along the *SLC30A9* gene among modern human populations.

## Materials and Methods

### *Signals of positive selection in the SLC30A9 region*

The PopHuman genome browser (http://pophuman.uab.cat) (Casillas et al. 2018) was consulted to search for signatures of recent positive selection along the *SLC30A9* region. This browser contains a broad range of genomic metrics and neutrality tests for all 26 human populations of the 1000 Genomes Project (1000 GP) (The 1000 Genomes Project Consortium 2015). Window-based computed values for $F_{ST}$, XP-EHH, and Fay and Wu's H statistics were extracted from three geographically differentiated populations from Africa, Europe, and Asia: the Yoruba from Ibadan (YRI), Utah residents with Northern and Western European ancestry (CEU), and Han Chinese from

Bejing (CHB). Gene annotations surrounding the *SLC30A9* region were obtained from the Ensembl genome browser (http://grch37.ensembl.org/). The collected information was then represented using the Gviz package (Hahne and Ivanek 2016). For each selection statistic and population, deviations from neutrality were considered as those greater than two standard deviations from the genomic mean.

*Analysis of putative adaptive variants in SLC30A9*

All SNPs for the longest *SLC30A9* gene transcript (ENST00000264451) were extracted from the Ensembl Genome Browser (GRCh37 assembly). Only biallelic variants with a MAF ≥ 0.02 were kept for subsequent analysis and annotated using ANNOVAR (Wang et al. 2010) to obtain their corresponding genomic location, SNP type classification (i.e., coding, non-coding, synonymous, non-synonymous), and allele frequency information. All compiled variants were explored for associations in the GWAS Catalog (v1.0) and eQTLS in the GTEx Portal Dataset (V7) and they were further annotated with different *in silico* function predictors such as the score of Combined Annotation-Dependent Depletion (CADD score) (Kircher et al. 2014), the Eigen score (Ionita-Laza et al. 2016), and the FitCons score (Gulko et al. 2015). SNPs in high linkage disequilibrium ($r^2$>0.8) with rs1047626 in CEU or CHB and their corresponding allele frequencies by continental region (AMR, EAS, EUR, SAS and AFR) were extracted from Ensembl.

*Analysis of positive selection on candidate variants*

We used Relate (https://myersgroup.github.io/relate) (Speidel et al. 2019) to test whether the derived alleles at rs1047626 and rs4861157 were detected as fast spreading variants along multiple lineages when compared to competing non-carriers. For that, we first obtained the Relate-estimated coalescence rates, haplotypes and genealogies available for the 1000 GP (kindly provided by Leo Speidel), removed all sample haplotypes not belonging to CEU, CHB or YRI (with RelateFileFormats --mode RemoveSamples) and annotated SNPs as recommended in order to extract trees (with RelateFileFormats --mode GenerateSNPAnnotations) and the inferred genealogies for rs1047626 and rs4861157 in CHB, CEU and YRI (with RelateExtract --mode SubTreesForSubpopulation --years_per_gen 28), which were then visualized (with RelateTreeViewMutation -m 1.25E-8). Subsequently, we ran CLUES (Stern et al. 2019) to estimate the corresponding selection coefficients and allelic trajectories in either CEU and CHB or YRI using the SampleBranchLengths output from Relate as obtained with the Inference.py with N 30.000. The same procedure was repeated to run CLUES on FIN, PEL, JPT, and PJL.

*Introgression*

Allelic states in the Neanderthal and Denisovan genomes for those SNPs displaying high linkage disequilibrium ($r^2$>0.8) with rs1047626 in CEU and CHB were queried from the genotypes available in four high coverage archaic genomes (https://cdna.eva.mpg.de/neanderthal/). We

initially inferred as putatively introgressed the genomic region comprised between the two more distant SNP positions found at high linkage disequilibrium ($r^2$>0.8) with rs1047626 in CEU and CHB where the archaic genomes carried derived alleles shared with the most frequent allele combinations in present-day humans while also coinciding with at least one of the Denisovan introgression maps available for Melanesians (Gower et al. 2021) and Papua New Guineans (Hubisz et al. 2020). To further explore a scenario of archaic introgression, all SNP genotypes along this putatively introgressed region (hg 38, chr4: 41977828-42048441) were extracted for all the Oceanians and for the CHB, CEU, and YRI populations from the corresponding VCF files of the HGDP (Bergström et al. 2020) and 1000GP (Auton et al. 2015) sequencing projects, merged, and phased with Shapeit4 (Delaneau et al. 2019) using the recombination maps available (https://github.com/eyherabh/genetic_map_comparisons). In turn, VCF files for the Denisova, AltaiNeandertal, Vindija33.19, and Chagyrskaya-Phalanx archaic genomes were downloaded (https://cdna.eva.mpg.de/neanderthal/), lifted over to hg38, and merged. We then crossed all the Denisovan positions found with a homozygous genotype in the putatively introgressed region with the corresponding Neanderthal positions and phased genotypes across the YRI, CHB, CEU and Oceanian populations. Only three positions in the Altai Neanderthal were found as heterozygotic positions and were kept in the analysis. From this complete set of inferred modern and archaic haplotypes (65 in total defined by 112 polymorphic positions, see Table S3), we then computed nucleotide pairwise distances and plotted the haplotype network using the ape and pegas R packages (Paradis 2010). Finally, the probability of observing a segment of 70,614 pb in-length shared with archaic humans due to incomplete lineage sorting (ILS) was calculated following the same procedure as in Huerta-Sánchez et al. (2014) considering the generation (g=25 years) and split times as in Dannemann et al. (2016) and the average recombination rate of the putatively introgressed *SLC30A9* segment ($r = 1.58 \times 10^{-8}$; extracted from https://github.com/eyherabh/genetic_map_comparisons). Briefly, we first calculated which is the expected length (L) for a segment shared by ILS using the formula L = 1/(r(t1+t2)/g) considering the split times for the divergence between humans and hominids (t1) and between Denisovan and Neanderthals (t2) given in Prüfer et al. (2014) when considering two different mutation rates. Subsequently, the probability of observing an archaic segment of 70,614 bp in-length that persists due to ILS was computed as 1- GammaCDF (observed length, k, I/L), being 1/L the rate of the gamma distribution, L the expected length and k, the shape (k=2) (see details in Table S5).

*GWAS analysis*

We used the GWAS catalog (https://www.ebi.ac.uk/gwas/) (Buniello et al. 2019) and the PheWAS option in the GWAS atlas (https://atlas.ctglab.nl/) (Watanabe et al. 2019) to look for trait associations with rs1047626 and SNPs in high LD ($r^2$ >0.8). In the PheWAS analysis, a total of

3,302 unique traits as available at the GWAS Atlas database (last accession on 16/05/2022) were considered for Bonferroni multiple test correction (FDR=1.5 x $10^{-5}$).

*Generation of polymorphic sites*

The plasmid containing the human *SLC30A9* gene codifying for the ZnT9 transporter with an N-HA tag was obtained from Sino Biological Inc. (Catalog nr. HG22621-NY). The Met50Val polymorphism was generated via site-directed mutagenesis following standard conditions (QuikChange Lightning; Agilent) with the following primers: forward primer (5'-GACATTTGGAAGCTTTTCAAACGTGGTTCCCTGTAGTCA-3') and reverse primer (5'-TGACTACAGGGAACCACGTTTGAAAAGCTTCCAAATGTC-3'). Before being used for cell transfection, both ZnT9-50Met and ZnT9-50Val isoforms were confirmed by sequencing. The LightRun sequencing service of Eurofins Genomics was used following the standard conditions for purified plasmid DNA samples with the following primers: forward primer 5'-GGCACAGAACTCAAAGCT-3' and reverse primer 5'-TCTTCATGGGGACTTCGT-3'.

*Genotyping of DNA liver samples*

DNA samples and information regarding zinc concentration in the liver were obtained from the project of Engelken et al. (2016), which determined micronutrient concentrations in liver samples from healthy individuals of western European origin, originally gathered by the UKHTB (for details, see Engelken et al. (2016)). DNA concentrations for 143 available samples from the original study were quantified using a NanoDrop One spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). A genomic region of 226 bp comprising the rs1047626 position was amplified, purified, and sequenced using the LightRun sequencing service (Eurofins Genomics) following standard conditions for purified PCR products. The primers used for amplification were: forward primer 5'-GAAGCAGTGAAACACCTCTGG-3', reverse primer 5'-TGTTTGTGATCCCTGTCCTTC-3'. For sequencing we used primer 5'-TGCAGCTAGGACTTGGTTTG-3'.

*Cell culture and transfection procedure*

HEK293 cells were cultured in DMEM supplemented with 10% FBS, 1% L-glutamine, and 1% penicillin and streptomycin (Basal medium) at 37 °C in a humidified 5% $CO_2$ atmosphere. $ZnSO_4$ was added as needed to the final medium to generate specific $Zn^{2+}$ concentration conditions. Cells were transiently transfected with ZnT9-50Met, ZnT9-50Val, pEGFP, or empty pCDNA3 vectors depending on the experiment. Polyethyleneimine (PEI) was used as the transfection reagent using 3 µg DNA for a 6-well plate or 1 µg DNA for a 24-well plate. In RNA interference experiments we used siRNA control (1027310, Qiagen) and siRNA ZnT9 (114789, Eupheria Biotech), and the transfection reagent was Lipofectamine 3000. Cells were incubated with the transfection solution for 3 hours, then washed and replaced with normal media. The experiments were performed 24h

after transfection. To characterize the genotype of cells used for all the experiments, the DNA of HEK293 cells was extracted and purified following the NucleoSpin Tissue protocol (Macherey-Nagel, Düren, Germany). The genotype at rs1047626 was verified to be heterozygote (AG) in HEK293 cells with the same sequencing procedure previously described for the liver DNA samples.

*Western Blotting*

Transfected cells were grown in 6-well plates and incubated for 24 hours. Cell lysis to detect ZnT9 protein was performed with 30 µL of lysis buffer containing 50 mM Tris-HCl at pH 7.4, 150 mM NaCl, 0.5% Nonidet P-40, and EDTA-free protease inhibition cocktail (Roche). Cell lysates were vortexed for 30 min at 4 °C and centrifuged at 10,000× *g* to remove aggregates and boiled for 5 min at 95 °C to lastly be placed on ice for 1 min. After electrophoresis in 12.5% polyacrylamide gel, proteins were transferred to nitrocellulose membranes using the iBlot system (Invitrogen, Waltham, MA, USA). Membranes were blocked with either 5% milk in TBS-Tween 0.1%, for GAPDH, or 5% BSA for HA-ZnT9, for 1 hour at room temperature. Primary antibodies were diluted in the corresponding blocking solution anti-HA for ZnT9 (H3663, Sigma-Merck), and anti-GAPDH (ab8245, Abcam). HRP secondary antibodies (1:1000; GE Healthcare) were used depending on the primary antibody. The ChemiDoc XRS+ system (Bio-Rad, Hercules, CA, USA) was used to obtain high-quality images. Quantity One Software (Bio-Rad) was used to analyze the results.

*Immunostaining*

For immunodetection of the two isoforms of ZnT9, for both cell surface and total cell expression, immunostaining was performed 24 hours after transfection with ZnT9-50Met or ZnT9-50Val plasmids in 24-well plates with coverslips coated with collagen. For cell surface expression, cells were incubated with anti HA (1:1000) in DMEM for 1h at 37°C. All samples were fixed with 4% paraformaldehyde (PFA) and only for total cell expression experiments were cells permeabilized with 0.1% Triton x-100 in PBS for 10 minutes. For all experiments, samples were blocked overnight with 1% BSA and 2% FBS in PBS. For total cell expression experiments, samples were incubated for 1h with an anti-HA antibody (1:1000 in blocking solution). For all experiments, cells were incubated with the secondary antibody (1:2000 in blocking solution), a goat anti-mouse Alexa Fluor 488 (Molecular Probes). Images were acquired using an inverted Leica SP8 Confocal Microscope with a 63× Oil objective and analyzed using ImageJ software.

In the colocalization analysis, to determine the subcellular distribution of the variants, immunostaining was performed 24 hours after transfection in 24-well plates with coverslips. Cells were co-transfected with ZnT9-50Met or ZnT9-50Val plasmids and a modified FRET-based ER-mitochondria proximity probe named FEMP (Naon et al. 2016). The immunostaining procedure

was the same as the one described above, except for the secondary antibody, which in this case was a goat anti-mouse Alexa Fluor 555 (Molecular Probes). The analysis of colocalization of ZnT9 with either the ER or mitochondria was performed with both Pearson's correlation coefficient and Mander's overlap coefficient using ImageJ software.

For STED imaging, cells were grown and seeded on 1.5H-thickness cover glasses. Primary antibodies were diluted in 2% BSA blocking buffer reagent. We used the rabbit anti-human TOM20 and the rabbit anti-human Calreticulin antibodies (1:1000) and secondary antibodies Abberior STAR RED or ORANGE (1:350). STED images were taken with a commercial Leica TCS SP8 STED 3× microscope equipped with a pulsed supercontinuum white light laser excitation source, using a 100× 1.4 NA oil HC PL APO CS2 objective. Post-analysis was performed on ImageJ. To measure the area occupied by the mitochondria with the Freehanded selection tool, we measured the cell region on the HA channel, subtracted the nucleus area and selected the area occupied by these organelles. The distribution of the mitochondria was quantified tracing a segmented line that connected all single mitochondria in one cell and then compare the length of these lines against the length in control cells.

*Real-Time RT PCR*

Cells transfected with either ZnT9-50Met, ZnT9-50Val, or empty pCDNA3 vectors were grown in 6-well plates and incubated for 24 hours. Total RNA was extracted from cells using Macherey-Nagel total RNA extraction kit. RNA was measured using a NanoDrop 1000 spectrophotometer (Thermo Fisher). cDNA was generated using a SuperScript Reverse Transcriptase system (Invitrogen). Quantitative PCR was performed using SYBR Green (Applied Biosystems) in the QuantStudio 12K system (Applied Biosystems). Primers are listed in Table S13.

*Zinc Measurements*

For the detection of zinc at the cytosolic level, cells were co-transfected with ZnT9-50Met or ZnT9-50Val vectors and pEGFP in 24-well plates. 24 hours after transfection, cells were incubated with 25 µM of Zinquin (Sigma-Aldrich, Darmstadt, Germany) for 30 min at 37 °C (5% $CO_2$) in isotonic solution (ISO) containing 140 mM NaCl, 2.5 mM KCl, 1.2 mM $CaCl_2$, 0.5 mM $MgCl_2$, 5 mM glucose, and 10 mM Hepes (300 milliosmoles/liter, pH 7.4), and different concentrations of $Zn^{2+}$. Cells were then dissociated with Trypsin 0.05% in 0.53 mM EDTA and washed with PBS. Fluorescence was quantified using an LSRII flow cytometer. Further analysis was performed using Flowing software to quantify Zinquin fluorescence of live transfected cells (Perttu Terho, Turun yliopisto, Turku, Finland).

To determine ER zinc levels, *in vivo* confocal imaging was used in cells co-transfected with the FRET-based ER-ZapCY1 probe (Addgene, Catalog nr. 36321)(Qin et al. 2011) and ZnT9-50Met, ZnT9-50Val, or empty pCDNA3 vectors in 6-well plates with 22 mm coverslips. 24 hours after

transfection, cells were incubated with basal medium or supplemented with 100 µM $ZnSO_4$ for 3 hours. Then, samples were placed under the microscope in ISO for imaging with an SP8 Leica microscope (Wetzlar, Germany). When blocking ZIP7 activity, 1 µM of the ZIP7 inhibitor NVS-ZP7-4; HY-114395, MedChemExpress (Quimigen) or 1 µM of DMSO used as a control were added at 1:1000 in the treatment 3 hours before imaging. The measured zinc content is dependent on the FRET signal, which is expressed as the mean of YFP/CFP ratios from each condition normalized across experiments to an empty vector in basal zinc conditions. Images were analyzed using ImageJ software.

Mitochondria zinc concentration was measured *in vivo* using a plate reader (VICTOR Nivo, Perkin Elmer). Cells were co-transfected with Mito-cCherry-Gn2Zn or SMAC-Gn2Zn probes (Fudge et al. 2018) and ZnT9-50Met, ZnT9-50Val, or empty pCDNA3 vectors in 24-well plates. 24 hours after transfection, cells were incubated for 40 min with basal medium or supplemented with 100 µM $ZnSO_4$. Then the media was changed to ISO solution and samples were placed in the plate reader for measuring. When blocking the ZIP7 activity, 1µM of the ZIP7 inhibitor NVS-ZP7-4 or DMSO was used as a control. Zinc content is expressed as the mean of GFP/Cherry ratios from each condition normalized to an empty vector in basal zinc conditions.

*MTT assays*

Cells were transfected with ZnT9-50Met, ZnT9-50Val, or empty pCDNA3 vectors in 24-well plates and incubated for 24h with normal medium, supplemented with 100 µM ZnSO4 or $Zn^{2+}$-free growth medium generated with Chelex 100 resin. Then, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) reagent was added (0.5 mg/mL) for 2 h at 37 °C. After that, the supernatant was removed, and cells were resuspended in 400 µl of DMSO. The absorbance was read at 570 nm.

**Data availability**

All genotyping data used in the manuscript is provided in the Supplemental Information or publicly available from the indicated public resources. Cell lines, probes and clones used in the reported experiments will be available upon request.

**Acknowledgments**

We thank Martin Kuhlwilm for his comments and help. This work was supported by Agencia Estatal de Investigación (AEI) and Fondo Europeo de Desarrollo Regional (FEDER) (project grants PID2019-110933GB-I00/AEI/10.13039/501100011033, PID2019-106755RB-I00/AEI/10.13039/501100011033); Unidad de Excelencia María de Maeztu funded by the Ministerio de Ciencia e Innovación and the Agencia Estatal de Investigación (DOI: 10.13039/501100011033, Ref: CEX2018-000792-M); and by Direcció General de Recerca,

Generalitat de Catalunya (2017SGR00702).

**Competing interests**

The authors declare that they have no competing interests.

**Author contributions**

Conceptualization: RV, EB; Methodology: FC, GM, RV, EB; Formal analysis and investigation: ARU, MVG, AF, JGC, AB, GIR, VHF, FC, GM, RV, EB; Supervision: RV, EB; Writing original draft: ARU and EB; Writing review & editing: GM, RV, EB; Funding acquisition: RV, EB.

**References**

Akey JM. 2009. Constructing genomic maps of positive selection in humans: Where do we go from here? *Genome Res* 19:711–722.

Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE, et al. 2015. A global reference for human genetic variation. *Nature* 526:68–74.

Barreiro LB, Laval G, Ne Quach H, Patin E, Quintana-Murci L. 2008. Natural selection has driven population differentiation in modern humans. *Nat Genet* 40:340–345.

Bergström A, McCarthy SA, Hui R, Almarri MA, Ayub Q, Danecek P, Chen Y, Felkel S, Hallast P, Kamm J, et al. 2020. Insights into human genetic variation and population history from 929 diverse genomes. *Science* 367 (6484).

Browning SR, Browning BL, Zhou Y, Tucci S, Akey JM. 2018. Analysis of Human Sequence Data Reveals Two Pulses of Archaic Denisovan Admixture. *Cell* 173:53-61.e9.

Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, McMahon A, Morales J, Mountjoy E, Sollis E, et al. 2019. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 47:D1005–D1012.

Carlson CS, Thomas DJ, Eberle MA, Swanson JE, Livingston RJ, Rieder MJ, Nickerson DA. 2005. Genomic regions exhibiting positive selection identified from dense genotype data. *Genome Res* 15:1553–1565.

Casillas S, Mulet R, Villegas-Mirón P, Hervas S, Sanz E, Velasco D, Bertranpetit J, Laayouni H, Barbadilla A. 2018. PopHuman: the human population genomics browser. *Nucleic Acids Res* 46:D1003–D1010.

Chen L, Wolf AB, Fu W, Li L, Akey JM. 2020. Identifying and Interpreting Apparent Neanderthal Ancestry in African Individuals. *Cell* 180:677-687.

Coleman JRI, Gaspar HA, Bryois J, Byrne EM, Forstner AJ, Holmans PA, de Leeuw CA, Mattheisen M, McQuillin A, Whitehead Pavlides JM, et al. 2020. The Genetics of the Mood

Disorder Spectrum: Genome-wide Association Analyses of More Than 185,000 Cases and 439,000 Controls. *Biol Psychiatry* 88:169–184.

Dannemann M, Andrés AM, Kelso J. 2016. Introgression of Neandertal- and Denisovan-like Haplotypes Contributes to Adaptive Variation in Human Toll-like Receptors. *Am J Hum Genet* 98:22–33.

Delaneau O, Zagury J-F, Robinson MR, Marchini JL, Dermitzakis ET. 2019. Accurate, scalable and integrative haplotype estimation. *Nat Commun* 10:5436.

Deng H, Qiao X, Xie T, Fu W, Li H, Zhao Yanmei, Guo M, Feng Y, Chen L, Zhao Yan, et al. 2021. SLC-30A9 is required for Zn2+ homeostasis, Zn2+ mobilization, and mitochondrial health. *Proc Natl Acad Sci U S A* 118.

Duman RS, Sanacora G, Krystal JH. 2019. Altered Connectivity in Depression: GABA and Glutamate Neurotransmitter Deficits and Reversal by Novel Treatments. *Neuron* 102:75–90.

Eide DJ. 2006. Zinc transporters and the cellular trafficking of zinc. *BBA* 1763:711–722.

Engelken J, Carnero-Montoro E, Pybus M, Andrews GK, Lalueza-Fox C, Comas D, Sekler I, de la Rasilla M, Rosas A, Stoneking M, et al. 2014. Extreme population differences in the human zinc transporter ZIP4 (*SLC39A4*) are explained by positive selection in Sub-Saharan Africa. *PLoS Genet* 10:e1004128.

Engelken J, Espadas G, Mancuso FM, Bonet N, Scherr AL, Jimenez-Alvarez V, Codina-Sola M, Medina-Stacey D, Spataro N, Stoneking M, et al. 2016. Signatures of evolutionary adaptation in quantitative trait loci influencing trace element homeostasis in liver. *Mol Biol Evol* 33:738–754.

Fan S, Hansen MEB, Lo Y, Tishkoff SA. 2016. Going global by adapting local: A review of recent human adaptation. *Science* 354:54–59.

Fudge DH, Black R, Son L, Lejeune K, Qin Y. 2018. Optical Recording of Zn 2+ Dynamics in the Mitochondrial Matrix and Intermembrane Space with the GZnP2 Sensor. *ACS Chem Biol* 13:1897–1905.

Fukada T, Kambe T. 2011. Molecular and genetic features of zinc transporters in physiology and pathogenesis. *Metallomics* 3:662–674.

Gillespie CF, Ressler KJ. 2005. Emotional learning and glutamate: translational perspectives. *CNS Spectr* 10:831–839.

Gower G, Picazo PI, Fumagalli M, Racimo F. 2021. Detecting adaptive introgression in human evolution using convolutional neural networks. *Elife* 10.

Grossman SR, Andersen KG, Shlyakhter I, Tabrizi S, Winnicki S, Yen A, Park DJ, Griesemer D, Karlsson EK, Wong SH, et al. 2013. Identifying Recent Adaptations in Large-Scale Genomic Data. *Cell* 152:703–713.

Gulko B, Hubisz MJ, Gronau I, Siepel A. 2015. A method for calculating probabilities of fitness consequences for point mutations across the human genome. *Nat Genet* 47:276–283.

Hahne F, Ivanek R. 2016. Visualizing Genomic Data Using Gviz and Bioconductor. *Method Mol Biol* 1418:335–351.

Hara T, Takeda T-A, Takagishi T, Fukue K, Kambe T, Fukada T. 2017. Physiological roles of zinc transporters: molecular and genetic importance in zinc homeostasis. *J Physiol Sci* 67:283–301.

Huang L, Tepaamorndech S. 2013. The SLC30 family of zinc transporters-A review of current understanding of their biological and pathophysiological roles. *Mol Aspects Med* 34:548–560.

Hubisz MJ, Williams AL, Siepel A. 2020. Mapping gene flow between ancient hominins through demography-aware inference of the ancestral recombination graph. *PLoS Genet* 16:e1008895.

Huerta-Sánchez E, Jin X, Asan, Bianba Z, Peter BM, Vinckenbosch N, Liang Y, Yi X, He M, Somel M, et al. 2014. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* 512:194–197.

International Zinc Nutrition Consultative Group (IZiNCG). 2004. Assessment of the Risk of Zinc Deficiency in Populations and Options for Its Control. (Hotz C, Brown KH, editors.).

Ionita-Laza I, McCallum K, Xu B, Buxbaum JD. 2016. A spectral approach integrating functional genomic annotations for coding and noncoding variants. *Nat Genet* 48(2):214-20.

Ishihara K, Yamazaki T, Ishida Y, Suzuki T, Oda K, Nagao M, Yamaguchi-Iwai Y, Kambe T. 2006. Zinc Transport Complexes Contribute to the Homeostatic Maintenance of Secretory Pathway Function in Vertebrate Cells. *J Biol Chem* 281:17743–17750.

Jeong J, Eide DJ. 2013. The SLC39 family of zinc transporters. *Mol Aspects Med* 34:612–619.

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.

Kambe T, Tsuji T, Hashimoto A, Itsumura N. 2015. The Physiological, Biochemical, and Molecular Roles of Zinc Transporters in Zinc Homeostasis and Metabolism. *Physiol Rev* 95:749–784.

Kim Y-H, Eom J-W, Koh J-Y. 2020. Mechanism of Zinc Excitotoxicity: A Focus on AMPK. *Front Neurosci* 14:577958.

Kircher M, Witten DM, Jain P, O 'roak BJ, Cooper GM, Shendure J, Edu M, Edu DW, Org P. 2014. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 46:310–315.

Kowalczyk A, Gbadamosi O, Kolor K, Sosa J, Andrzejczuk L, Gibson G, St Croix C, Chikina M, Aizenman E, Clark N, et al. 2021. Evolutionary rate covariation identifies SLC30A9 (ZnT9) as a mitochondrial zinc transporter. *Biochemical Journal* 478:3205–3220.

Lachance J, Tishkoff SA. 2013. Population Genomics of Human Adaptation. *Annu Rev Ecol Evol Syst* 44:123–143.

Ma T, Zhao L, Zhang J, Tang R, Wang X, Liu N, Zhang Q, Wang F, Li M, Shan Q, et al. 2021. A pair of transporters controls mitochondrial Zn 2+ levels to maintain mitochondrial homeostasis. *Protein Cell* 13(3):180-202.

McGrath T, Baskerville R, Rogero M, Castell L. 2022. Emerging Evidence for the Widespread Role of Glutamatergic Dysfunction in Neuropsychiatric Diseases. *Nutrients* 14(5):917.

Naon D, Zaninello M, Giacomello M, Varanita T, Grespi F, Lakshminaranayan S, Serafini A, Semenzato M, Herkenne S, Hernández-Alvarez MI, et al. 2016. Critical reappraisal confirms that Mitofusin 2 is an endoplasmic reticulum–mitochondria tether. *Proc Nat Acad Sci* 113:11249–11254.

Paradis E. 2010. pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26:419–420.

Perez Y, Shorer Z, Liani-Leibson K, Chabosseau P, Kadir R, Volodarsky M, Halperin D, Barber-Zucker S, Shalev H, Schreiber R, et al. 2017. *SLC30A9* mutation affecting intracellular zinc homeostasis causes a novel cerebro-renal syndrome. *Brain* 140:928–939.

Pickrell JK, Coop G, Novembre J, Kudaravalli S, Li JZ, Absher D, Srinivasan BS, Barsh GS, Myers RM, Feldman MW, et al. 2009. Signals of recent positive selection in a worldwide sample of human populations. *Genome Res*:826–837.

Primes G, Fieder M. 2018. Real-life helping behaviours in North America: A genome-wide association approach.Novelli G, editor. *PLoS One* 13:e0190950.

Prinz WA, Toulmay A, Balla T. 2020. The functional universe of membrane contact sites. *Nat Rev Mol Cell Biol* 21:7–24.

Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, Heinze A, Renaud G, Sudmant PH, de Filippo C, et al. 2014. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505:43–49.

Qin Y, Dittmer PJ, Park JG, Jansen KB, Palmer AE, Valentine JS. 2011. Measuring steady-state and dynamic endoplasmic reticulum and Golgi Zn 2þ with genetically encoded sensors. *Proc Nat Acad Sci* 108:7351–7356.

Rees Jasmin S., Castellano S, Andrés AM. 2020. The Genomics of Human Local Adaptation. *Trends Genet*. 36:415–428.

Rensvold JW, Shishkova E, Sverchkov Y, Miller IJ, Cetinkaya A, Pyle A, Manicki M, Brademan DR, Alanay Y, Raiman J, et al. 2022. Defining mitochondrial protein functions through deep multiomic profiling. *Nature* 606:382–388.

Roca-Umbert A, Caro-Consuegra R, Londono-Correa D, Rodriguez-Lozano GF, Vicente R, Bosch E. 2022. Understanding signatures of positive natural selection in human zinc transporter genes. *Sci Rep* 12:4320.

Sabeti PC, Schaffner SF, Lohmueller J, Varilly P, Shamovsky O, Palma A, Mikkelsen TS, Altshuler D, Lander ES. 2006. Positive Natural Selection in the Human Lineage. *Science* 312:1614–1620.

Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, et al. 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature* 449:913–918.

Sankararaman S, Mallick S, Patterson N, Reich D. 2016. The Combined Landscape of Denisovan and Neanderthal Ancestry in Present-Day Humans. *Curr Biol* 26:1241–1247.

Skov L, Coll Macià M, Sveinbjörnsson G, Mafessoni F, Lucotte EA, Einarsdóttir MS, Jonsson H, Halldorsson B, Gudbjartsson DF, Helgason A, et al. 2020. The nature of Neanderthal introgression revealed by 27,566 Icelandic genomes. *Nature* 582:78–83.

Speidel L, Forest M, Shi S, Myers SR. 2019. A method for genome-wide genealogy estimation for thousands of samples. *Nat Genet* 51:1321–1329.

Stern AJ, Wilton PR, Nielsen R. 2019. An approximate full-likelihood method for inferring selection and allele frequency trajectories from DNA sequence data.Hernandez RD, editor. *PLoS Genet* 15:e1008384.

Suzuki T, Ishihara Kaori, Migaki H, Ishihara Kengo, Nagao M, Yamaguchi-Iwai Y, Kambe T. 2005. Two Different Zinc Transport Complexes of Cation Diffusion Facilitator Proteins Localized in the Secretory Pathway Operate to Activate Alkaline Phosphatases in Vertebrate Cells. *J Biol Chem* 280:30956–30962.

Takeda A, Fujii H, Minamino T, Tamano H. 2014. Intracellular Zn(2+) signaling in cognition. *J Neurosci Res* 92:819–824.

Taylor KM, Hiscox S, Nicholson RI, Hogstrand C, Kille P. 2012. Protein kinase CK2 triggers cytosolic zinc signaling pathways by phosphorylation of zinc channel ZIP7. *Sci Signal* 5.

The 1000 Genomes Project Consortium. 2015. A global reference for human genetic variation. *Nature* 526:68–74.

Vallee BL, Falchuk KH. 1993. The Biochemical Basis of Zinc Physiology. *Physiol Rev* 73:79–118.

Vitti JJ, Grossman SR, Sabeti PC. 2013. Detecting natural selection in genomic data. *Annu Rev Genet* 47:97–120.

Wang K, Li M, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38.

Watanabe K, Stringer S, Frei O, Umićević Mirkov M, de Leeuw C, Polderman TJC, van der Sluis S, Andreassen OA, Neale BM, Posthuma D. 2019. A global overview of pleiotropy and genetic architecture in complex traits. *Nat Genet* 51:1339–1348.

Williamson SH, Hubisz MJ, Clark AG, Payseur BA, Bustamante CD, Nielsen R. 2007. Localizing Recent Adaptive Evolution in the Human Genome. *PLoS Genet* 3:e90.

Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, Adams MJ, Agerbo E, Air TM, Andlauer TMF, et al. 2018. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet* 50:668–681.

Xue J, Xie T, Zeng W, Jiang Y, Bai XC. 2020. Cryo-EM structures of human ZnT8 in both outward-and inward-facing conformations. *Elife* 9:1–32.

Zhang C, Li J, Tian L, Lu D, Yuan K, Yuan Y, Xu S. 2015. Differential natural selection of human zinc transporter genes between african and non-African populations. *Sci Rep* 5:9658.

**Table 1.** Putative adaptive variants in the *SLC30A9* region.

| SNP | Ref | Alt | SNP type | AFR | EUR | EAS | SAS | AMR | CADD Phred score | eQTLs [a] (Tissue; NES; p-value) |
|---|---|---|---|---|---|---|---|---|---|---|
| rs2581434 | C | G | synonymous | 0.246 | 0.815 | 0.959 | 0.860 | 0.840 | 12.00 | Brain-Nucleus accumbens (basal ganglia); -0.23; 7.9 x $10^{-10}$ |
| rs1047626 | A | G | non-synonymous | 0.105 | 0.757 | 0.948 | 0.780 | 0.768 | 0.001 | Brain-Cortex; -0.24; 4.4 x $10^{-10}$ |
| rs2581452 | C | T | intronic | 0.193 | 0.776 | 0.958 | 0.800 | 0.803 | 11.38 | Brain-Nucleus accumbens (basal ganglia); -0.24; 2.5 x $10^{-11}$ |
| rs2660319 | G | A | intronic | 0.146 | 0.756 | 0.949 | 0.780 | 0.770 | 12.93 | Brain-Nucleus accumbens (basal ganglia); -0.22; 5.0 x $10^{-10}$ |
| rs1848182 | A | G | intronic | 0.193 | 0.776 | 0.958 | 0.800 | 0.803 | 10.10 | Brain-Nucleus accumbens (basal ganglia); -0.24; 2.5 x $10^{-11}$ |
| rs2581424 | G | A | intronic | 0.188 | 0.776 | 0.958 | 0.800 | 0.804 | 10.78 | Brain-Nucleus accumbens (basal ganglia); -0.23; 3.1 x $10^{-10}$ |
| rs15857 | C | A | synonymous | 0.188 | 0.776 | 0.958 | 0.800 | 0.804 | 14.93 | Brain-Nucleus accumbens (basal ganglia); -0.23; 3.1 x $10^{-10}$ |
| rs7439806 | G | A | intronic | 0.188 | 0.776 | 0.958 | 0.800 | 0.804 | 10.63 | Brain-Nucleus accumbens (basal ganglia); -0.23; 3.1 x $10^{-10}$ |
| rs55835604 | T | A | intronic | 0.191 | 0.776 | 0.958 | 0.800 | 0.804 | 13.81 | Brain-Nucleus accumbens (basal ganglia); -0.23; 3.1 x $10^{-10}$ |
| rs12510574 | G | A | intronic | 0.393 | 0.123 | 0.037 | 0.100 | 0.134 | 12.13 | Adipose - Subcutaneous; -0.16; 1.9 x $10^{-6}$ |
| rs4861014 | A | G | intronic | 0.067 | 0.752 | 0.951 | 0.780 | 0.759 | 12.23 | Brain-Nucleus accumbens (basal ganglia); -0.25; 1.2 x $10^{-12}$ |
| rs10019356 | A | G | intronic | 0.176 | 0.789 | 0.960 | 0.820 | 0.811 | 15.41 | Brain-Nucleus accumbens (basal ganglia); -0.24; 4.2 x $10^{-10}$ |
| rs7660223 | C | T | intronic | 0.153 | 0.789 | 0.960 | 0.820 | 0.808 | 15.33 | Brain-Nucleus accumbens (basal ganglia); -0.24; 4.2 x $10^{-10}$ |
| rs11051 | G | A | 3' UTR | 0.051 | 0.751 | 0.951 | 0.780 | 0.758 | 10.81 | Brain-Nucleus accumbens (basal ganglia); -0.25; 1.2 x $10^{-12}$ |
| rs11935648 | G | A | 3' UTR | 0.176 | 0.788 | 0.960 | 0.820 | 0.808 | 1.44 | Brain-Nucleus accumbens (basal ganglia); -0.24; 4.2 x $10^{-10}$ |
| rs12511999 | C | T | 3' UTR | 0.087 | 0.751 | 0.951 | 0.780 | 0.764 | 0.24 | Brain-Nucleus accumbens (basal ganglia); -0.26; 9.9 x $10^{-14}$ |
| rs10938178 | G | C | 3' UTR | 0.067 | 0.751 | 0.951 | 0.780 | 0.759 | 1.32 | Brain-Nucleus accumbens (basal ganglia); -0.25; 1.2 x $10^{-12}$ |
| rs12512101 | C | T | 3' UTR | 0.086 | 0.751 | 0.951 | 0.780 | 0.764 | 4.31 | Brain-Nucleus accumbens (basal ganglia); -0.26; 9.9 x $10^{-14}$ |

[a] eQTLs described for *SLC30A9*. Extracted from GTEx database V7. NES: Normalized effect size.

## Figure Legends

**Figure 1. Signals of positive selection and introgression at *SLC30A9*.** (A) Signatures of adaptation in the *SLC30A9* region. Chromosomal location and genes are displayed at the top. In the first selection panel, window-based $F_{ST}$ values when comparing CEU versus YRI and CHB versus YRI are represented by cyan and pink bars, respectively. Similarly, window-based Fay and Wu's H values and XP-EHH scores (using YRI as the reference population) in CEU and CHB are shown with the same color scheme in the middle and bottom selection panels, respectively. Overlapping population scores can result in blue shades. Horizontal cyan and pink lines represent the genomic average of each test for each population, whereas 2 standard deviations are represented by the background shade in each case as a mean to interpret the genome-wide significance of each selection signature. All selection test values were extracted from the corresponding 1000 GP populations as available at the PopHuman browser (http://pophuman.uab.cat) (Casillas et al. 2018) and represented using the Gviz package (Hahne and Ivanek 2016). (B) Allele frequency trajectories and selection coefficients inferred with CLUES for rs1047626 in CHB and CEU and for rs4861157 in YRI. (C) Schematic representation of the two major human haplotypes as defined by 84 SNPs in high linkage disequilibrium ($r^2$>0.8) with rs1047626 in CEU and CHB along the putatively inferred introgressed region (chr4: 41,977,828-42,048,441) and the corresponding allele states found in four high coverage Denisovan and Neanderthal genomes. Derived states are indicated in black, and ancestral alleles in white. Den to Papuan and Den to Melanesian represent Denisovan indicate introgression segments previously described in Melanesians (Gower et al. 2021) and Papua New Guineans (Hubisz et al. 2020). The black arrow point to the rs1047626 position. GWAS hits are indicated with an asterisk (see Table S2 for details). Note that a further detailed representation of modern and archaic haplotypes as defined by all SNPs with a homozygous genotype in Denisova along the same putatively introgressed region is available in Fig. S5. (D) Network of modern and archaic human haplotypes as defined by the complete set of SNPs available with a homozygous genotype in the Denisovan genome along the putatively introgressed region (chr4: 41,977,828-42,048,441). Pie charts show the frequency of each haplotype among YRI, CEU, CHB, and Oceanians (OCE). Denisovan and Neanderthal haplotypes are show in green and yellow, respectively. Arrow in black, position of the rs1047626 substitution; arrow in grey, position of the rs4861157 polymorphism. Haplotype frequencies, and polymorphic positions defining each haplotype are available in Table S3. For nucleotide pairwise distances between haplotypes, see Table S4.

**Figure 2. Characterization of the structure and overexpression of ZnT9 in HEK293 cells.** (A) ZnT9 protein structure prediction by AlphaFold, showing 6 transmembrane domains and a cytosolic N-terminus, where the Met50Val substitution is indicated by a red arrow. (B) Protein

sequence of the ZnT9 transporter with the Met50Val substitution in red and the predicted alpha-helices corresponding to transmembrane domains in blue. (C) Representative Western blot against HA and GAPDH in cells transfected with ZnT9-50Met, ZnT9-50Val, or an empty vector. (D) Representative pictures of the membrane and intracellular immunostaining of HEK293 cells expressing ZnT9-50Met and ZnT9-50Val isoforms (both in green). Nucleae were stained with DAPI (blue). Scale bar = 10 µm.

**Figure 3. Subcellular localization of ZnT9 variants.** (A) Superresolution STED microscopy in cells transfected with ZnT9-50Met and ZnT9-50Val immunoassayed with anti-HA (green) and with the mitochondrial marker anti-TOM20. Scale bar = 10 µm. (B-C) Bar graph representing mean inter-mitochondrial distance (B) and relative mitochondria area (C) in cells transfected with ZnT9-50Met or ZnT9-50Val (n=10). *** $p<0.001$ using Bonferroni test between conditions. (D) Bar graph measuring FRET between the endoplasmic reticulum and mitochondria transfected with an empty vector (control), ZnT9-50Met or ZnT9-50Val together with the FEMP probe (n=46-65).

**Figure 4. Characterization of the impact on cellular zinc homeostasis of ZnT9 variant overexpression in HEK293 cells.** (A) RT-PCR comparing the expression of several ZTs in basal conditions in cells transfected with ZnT9-50Met, ZnT9-50Val, or an empty vector. $2^{-(DDCT)}$ plotted using GAPDH as the housekeeping gene (ZIP1, ZIP6, and ZIP7 n = 6; Rest of ZTs n = 3); * $p<0.05$ versus control using Bonferroni, # $p<0.05$ versus control using t-test. (B) Evaluation of zinc content by flow cytometry using Zinquin in 10 µM and 100 µM of $ZnSO_4$ (n = 9); *** $p<0.001$ using t-test. (C) Evaluation of endoplasmic zinc content using an endoplasmic reticulum fluorescent zinc sensor (ER-ZapCY1) in basal and 100 µM $ZnSO_4$ conditions (n = 14-20). *** $p<0.001$ using t-test; * $p<0.05$ using Mann-Whitney test. (D) RT-PCR comparing the expression of several ZnT transporters in basal conditions in cells transfected with ZnT9-50Met, ZnT9-50Val, or an empty vector. $2^{-(DDCT)}$ plotted using GAPDH as the housekeeping gene (n = 5-6) ** $p<0.01$, *** $p<0.001$ using Bonferroni. (E) Evaluation of endoplasmic zinc content using an ER fluorescent zinc sensor (ER-ZapCY1) in basal and 100 µM $ZnSO_4$ in the presence of DMSO and Zip7 blocker (n = 12-19). (F) Evaluation of zinc mitochondrial intermembrane space content using SMAC-Gn2Zn probe in conditions of 10 µM and 100 µM $ZnSO_4$ (n = 8-9); # $p<0.05$ using t-test between transfection conditions. (G) Evaluation of zinc mitochondrial matrix content using Mito-cCherry-Gn2Zn incubating 40min with basal and 100 µM $ZnSO_4$ conditions (n = 9-12); ** $p<0.01$ using t-test between basal and 100 µM zinc conditions, # $p<0.05$ using t-test between transfection conditions. (H) Viability MTT assay in cells transfected with ZnT9-50Met, ZnT9-50Val, or an empty vector incubated

for 24h at 0, basal and 100 µM $ZnSO_4$ conditions (n=8-15) * $p<0.05$, ** $p<0.01$ using Bonferroni between transfection conditions, # $p<0.05$ using Bonferroni between zinc conditions.

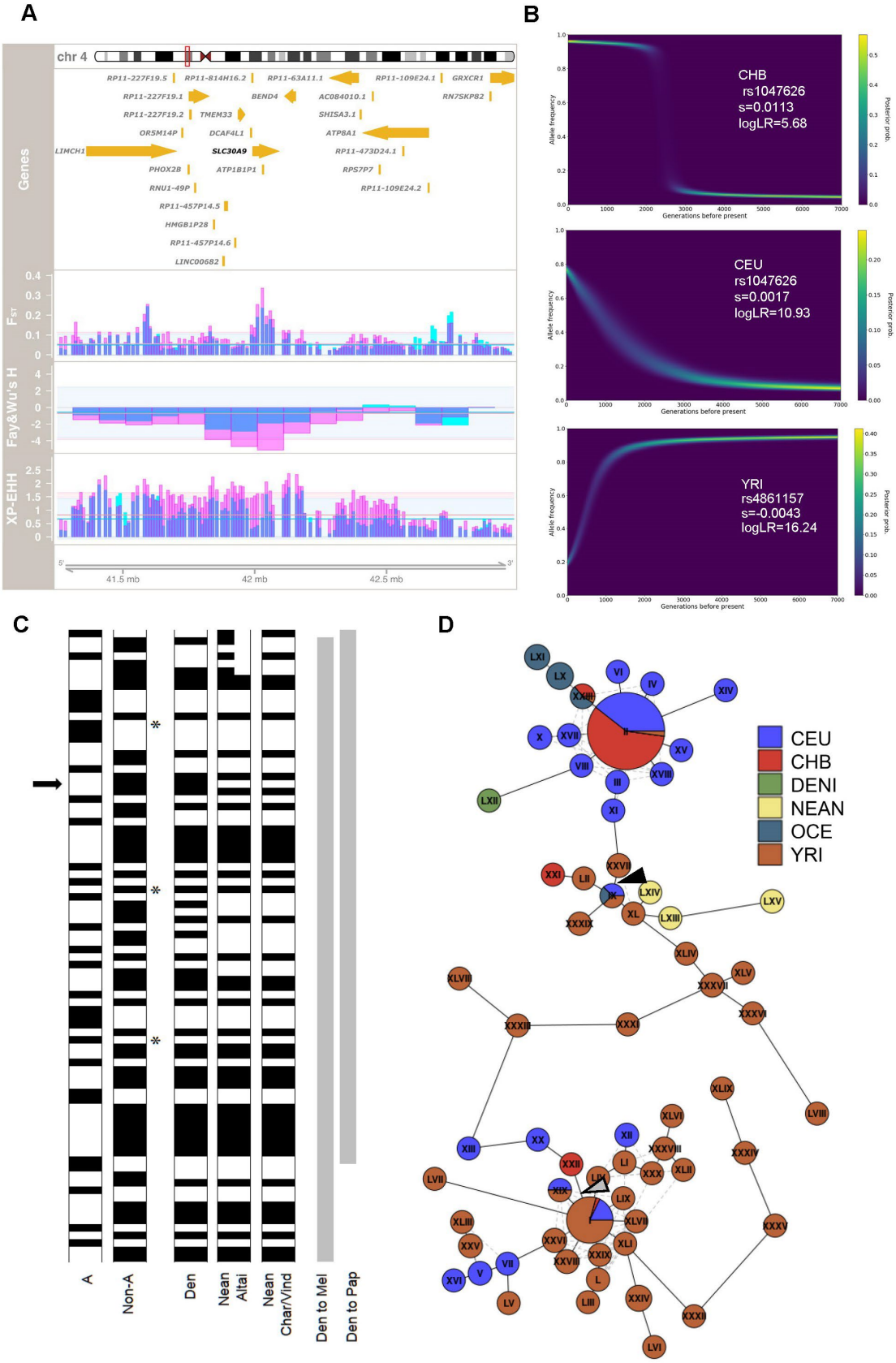
A
B
C
D
chr 4
Genes
RP11-227F19.5
RP11-814H16.2
RP11-63A11.1
RP11-109E24.1
GRXCR1
RP11-227F19.1
BEND4
AC084010.1
RN7SKP82
RP11-227F19.2
TMEM33
SHISA3.1
OR5M14P
DCAF4L1
ATP8A1
LIMCH1
SLC30A9
RP11-473D24.1
PHOX2B
ATP1B1P1
RPS7P7
RNU1-49P
RP11-109E24.2
RP11-457P14.5
HMGB1P28
RP11-457P14.6
LINC00682
FST
Fay&Wu's H
XP-EHH
41.5 mb
42 mb
42.5 mb
CHB
rs1047626
s=0.0113
logLR=5.68
CEU
rs1047626
s=0.0017
logLR=10.93
YRI
rs4861157
s=-0.0043
logLR=16.24
Allele frequency
Generations before present
Posterior prob.
A
Non-A
Den
Nean Altai
Nean Char/Vind
Den to Mel
Den to Pap
CEU
CHB
DENI
NEAN
OCE
YRI

**Figure 1**

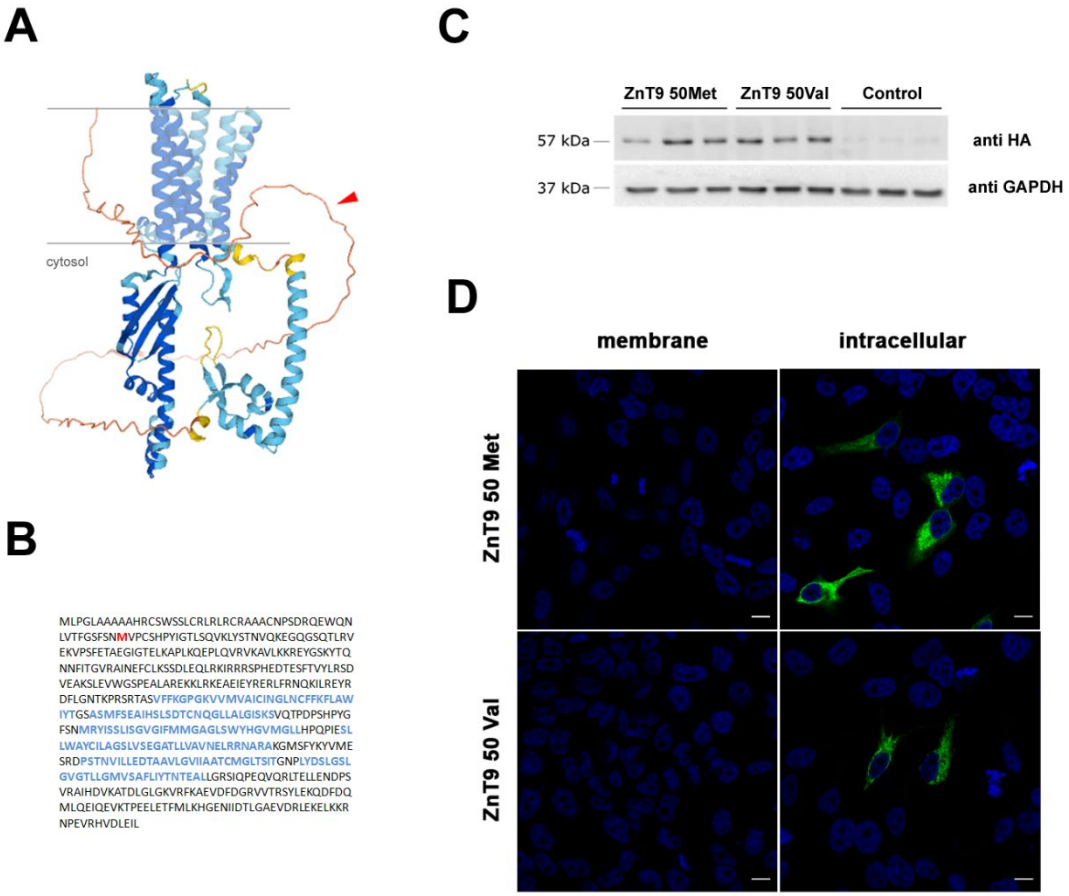
A
cytosol
B
MLPGLAAAAAHRCSWSSLCRLRLRCRAAACNPSDRQEWQN
LVTFGSFSNMVPCSHPYIGTLSQVKLYSTNVQKEGQGSQTLRV
EKVPSFETAEGIGTELKAPLKQEPLQVRVKAVLKKREYGSKYTQ
NNFITGVRAINEFCLKSSDLEQLRKIRRRSPHEDTESFTVYLRSD
VEAKSLEVWGSPEALAREKKLRKEAEIEYRERLFRNQKILREYR
DFLGNTKPRSRTASVFFKGPGKVVMVAICINGLNCFFKFLAW
IYTGSASMFSEAIHSLSDTCNQGLLALGISKSVQTPDPSHPYG
FSNMRYISSLISGVGIFMMGAGLSWYHGVMGLLHPQPIESL
LWAYCILAGSLVSEGATLLVAVNELRRNARAKGMSFYKYVME
SRDPSTNVILLEDTAAVLGVIIAATCMGLTSITGNPLYDSLGSL
GVGTLLGMVSAFLIYTNTEALLGRSIQPEQVQRLTELLENDPS
VRAIHDVKATDLGLGKVRFKAEVDFDGRVVTRSYLEKQDFDQ
MLQEIQEVKTPEELETFMLKHGENIIDTLGAEVDRLEKELKKR
NPEVRHVDLEIL
C
ZnT9 50Met
ZnT9 50Val
Control
57 kDa
anti HA
37 kDa
anti GAPDH
D
membrane
intracellular
ZnT9 50 Met
ZnT9 50 Val

**Figure 2**

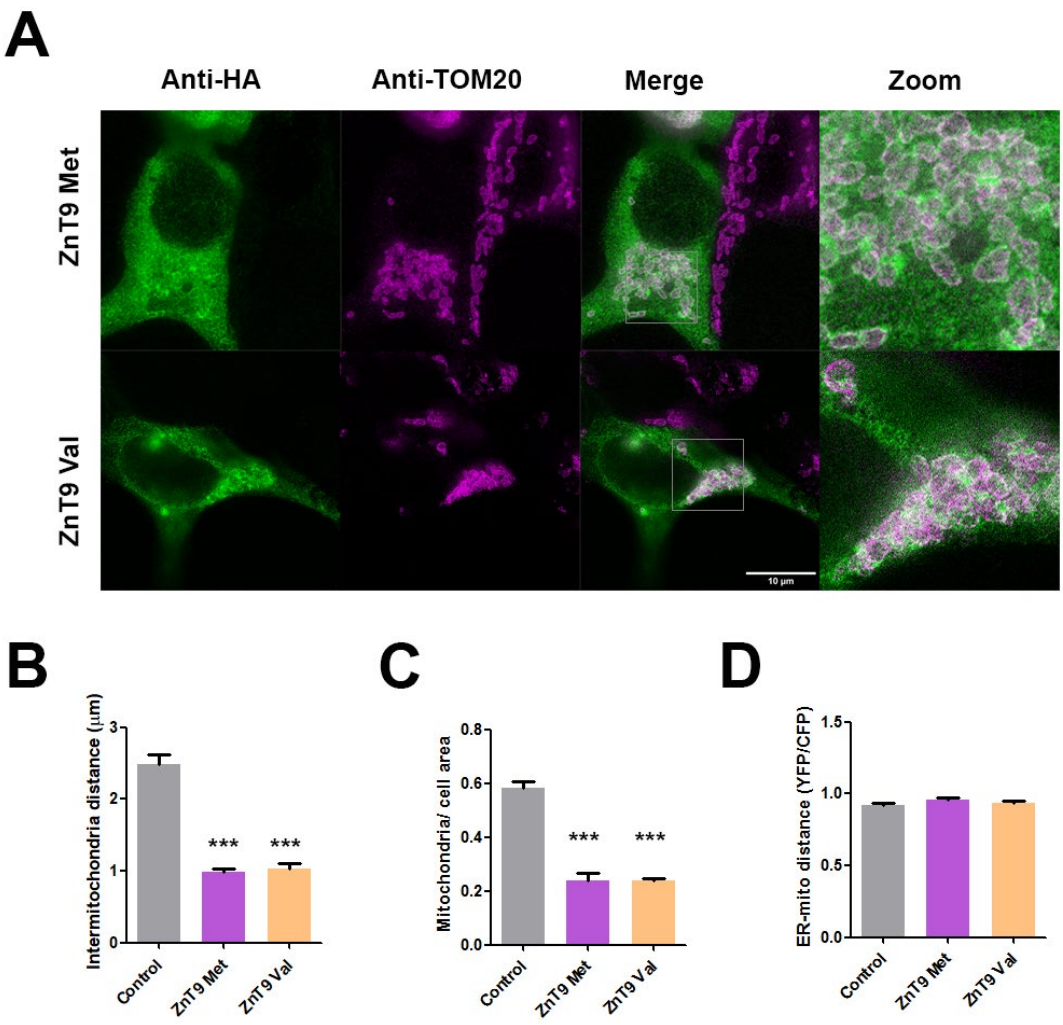
A
Anti-HA
Anti-TOM20
Merge
Zoom
ZnT9 Met
ZnT9 Val
10 µm
B
Intermitochondria distance (µm)
3
2
1
0
***
***
Control
ZnT9 Met
ZnT9 Val
C
Mitochondria/ cell area
0.8
0.6
0.4
0.2
0.0
***
***
Control
ZnT9 Met
ZnT9 Val
D
ER-mito distance (YFP/CFP)
1.5
1.0
0.5
0.0
Control
ZnT9 Met
ZnT9 Val

**Figure 3**

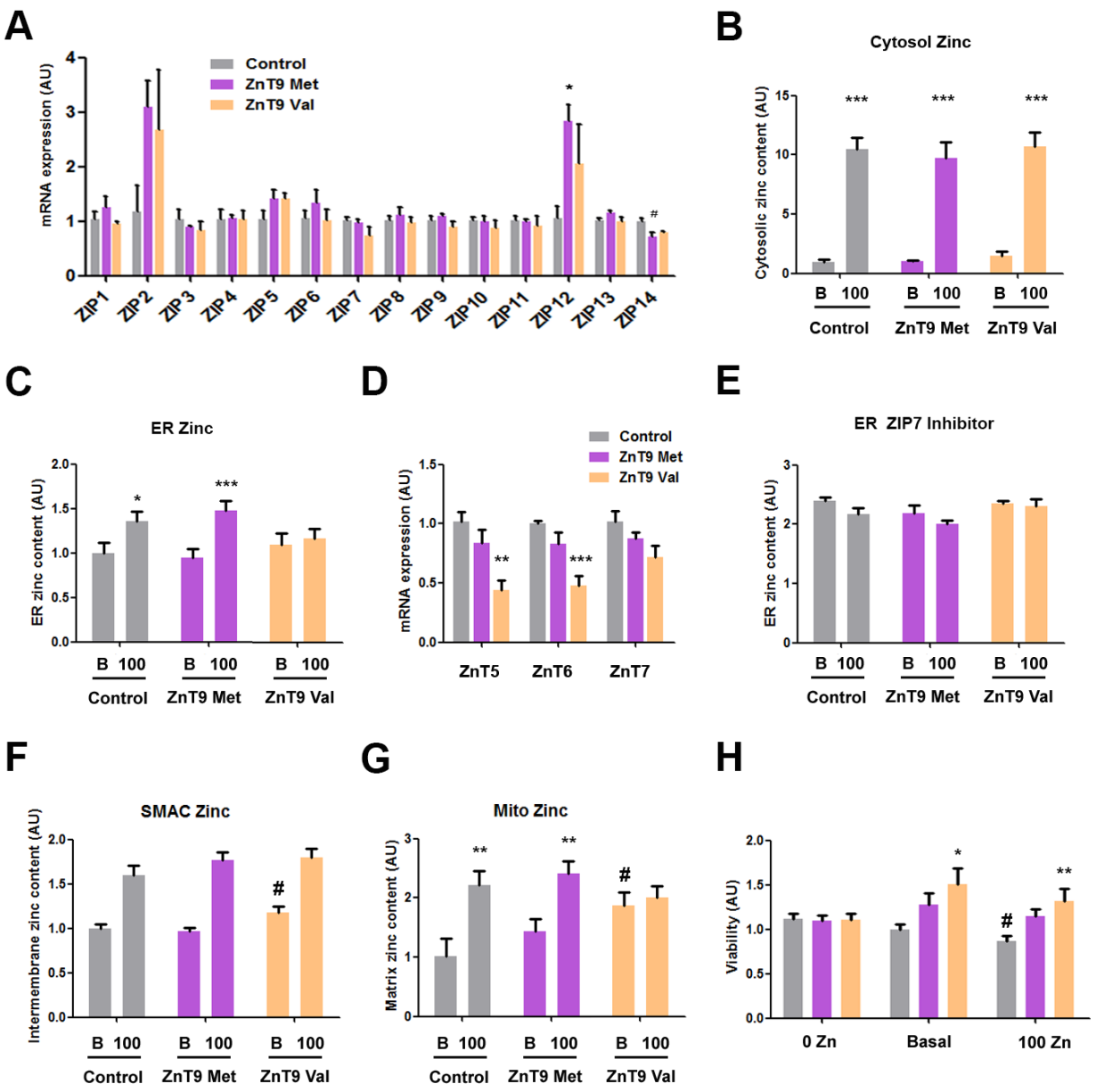
A
mRNA expression (AU)
Control
ZnT9 Met
ZnT9 Val
ZIP1
ZIP2
ZIP3
ZIP4
ZIP5
ZIP6
ZIP7
ZIP8
ZIP9
ZIP10
ZIP11
ZIP12
ZIP13
ZIP14
B
Cytosol Zinc
Cytosolic zinc content (AU)
B 100
Control
ZnT9 Met
ZnT9 Val
C
ER Zinc
ER zinc content (AU)
D
mRNA expression (AU)
ZnT5
ZnT6
ZnT7
E
ER ZIP7 Inhibitor
ER zinc content (AU)
F
SMAC Zinc
Intermembrane zinc content (AU)
G
Mito Zinc
Matrix zinc content (AU)
H
Viability (AU)
0 Zn
Basal
100 Zn

**Figure 4**