

## Did variants in inborn errors of immunity genes contribute to the extinction of Neanderthals?

Zijun Zhou,<sup>1,2,5</sup> Sigrid M.A Swagemakers,<sup>3</sup> Mirthe S. Lourens,<sup>1,5</sup> Narissara Suratannon,<sup>1,4</sup> Peter J. van der Spek,<sup>3</sup> Virgil A.S.H. Dalm,<sup>1,2,5</sup> Willem A. Dik,<sup>1,5</sup> Hanna Ijspeert,<sup>1,5</sup> P. Martin van Hagen<sup>1,2,4,5</sup>

### Abstract

**Background:** Neanderthals were a species of archaic humans that became extinct around 40,000 years ago. Modern humans have inherited 1-6% of Neanderthal DNA as a result of interbreeding. These inherited Neanderthal genes have paradoxical influences, while some can provide protection to viral infections, some others are associated with autoimmune/auto-inflammatory diseases.

**Objective:** We aim to investigate whether genetic variants with strong detrimental effects on the function of the immune system could have potentially contributed to the extinction of the Neanderthal population.

**Methods:** We used the publically available genome information from an Altai Neanderthal and filtered for potentially damaging variants present in genes associated with inborn errors of immunity (IEI) and checked whether these variants were present in the genomes of the Denisovan, Vindija and Chagyrskaya Neanderthals.

**Results:** We identified 24 homozygous variants and 15 heterozygous variants in IEI-related genes in the Altai Neanderthal. Two homozygous variants in the *UNC13D* gene and one variant in the *MOGS* gene were present in all archaic genomes. Defects in the *UNC13D* gene are known to cause a severe and often fatal disease called hemophagocytic lymphohistiocytosis (HLH). One of these variants p.(N943S) has been reported in patients with HLH. Variants in *MOGS* are associated with glycosylation defects in the immune system affecting the susceptibility for infections.

**Conclusion:** Although the exact functional impact of these three variants needs further elucidation, we speculate that they could have resulted in an increased susceptibility to severe diseases and may have contributed to the extinction of Neanderthals after exposure to specific infections.

**Key words:** Neanderthals, Denisovan, inborn errors of immunity (IEI), innate immunity, Hemophagocytic Lymphohistiocytosis (HLH), HLA, *MOGS*, *UNC13D*, MHC I presentation, glycosylation, cytotoxicity, viral infection, SARS-Cov2

#### Citation:

Zhou, Z., Swagemakers, S. M. A., Lourens, M. S., Suratannon, N., van der Spek, P. J., Dalm, V. A. S. H., Dik, W. A., Ijspeert, H., van Hagen, P. M. (2022). Did variants in inborn errors of immunity genes contribute to the extinction of Neanderthals? *Asian Pac J Allergy Immunol*, 40(4), 422-434. <https://doi.org/10.12932/ap-251022-1489>

<sup>3</sup> Erasmus MC, University Medical Center Rotterdam, department of Pathology, Division of Clinical Bioinformatics, the Netherlands

<sup>4</sup> Center of Excellence for Allergy and Clinical Immunology, Division of Allergy, Immunology and Rheumatology, Department of Pediatrics, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

<sup>5</sup> Erasmus MC, University Medical Center Rotterdam, Academic Center for Rare Immunological Diseases (RIDC), the Netherlands

#### Affiliations:

<sup>1</sup> Erasmus MC, University Medical Center Rotterdam, Laboratory Medical Immunology, department of Immunology, the Netherlands

<sup>2</sup> Erasmus MC, University Medical Center Rotterdam, department of Internal Medicine, Division of Clinical Immunology, the Netherlands

#### Corresponding author:

P. M. van Hagen  
Erasmus MC, Dept. of Internal Medicine  
Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands  
E-mail: p.m.vanhagen@erasmusmc.nl

## Introduction

Neanderthals were a group of archaic humans that used to inhabit Eurasia. Neanderthals lived in relatively small social groups (10-30 individuals),<sup>1</sup> probably composed of individuals of both sexes and different age classes.<sup>1</sup> A related but distinctive sister group of Neanderthals known as the Denisovans lived at the same time in Siberia and eastern Asia.<sup>2</sup> Ancestors of the Neanderthals and Denisovans diverged from the main human lineage at least 600,000 years ago, and then split from each other around 400,000 years ago.<sup>3-6</sup> Although they lived mainly in isolated groups, they were not completely isolated from the early modern humans that migrated from Africa.<sup>3,7</sup> Archaeological evidence supports that these groups of archaic humans and early modern humans shared the same habitat until, by some yet-unknown causes, Neanderthals and Denisovans disappeared around 40,000 years ago. Their extinction coincided with the migration of early modern humans outside of Africa starting around 50,000 years ago.<sup>8</sup> Early modern humans have co-existed with Neanderthals and Denisovans for at least 10,000 years which allowed for interbreeding, albeit a low magnitude between these two groups.<sup>9</sup> Gene flow analysis showed that genomes of modern Eurasians contain around 1-6% DNA inherited from interbreeding with Neanderthals and Denisovans.<sup>3,9,10</sup>

Although the exact function of the surviving archaic human genes in modern humans genomes remains unclear, several recent genomic studies suggested that the inherited archaic alleles may explain certain human clinical traits.<sup>11,12</sup> Genetic variants impacting immunity have been shown to be transferred from archaic humans to modern humans.<sup>13-15</sup> In particular, innate immunity regulatory genes as the *TLR6-1-10* gene cluster and the *MERTK* gene were introduced into modern European genomes from both Neanderthals and Denisovans.<sup>14</sup> Furthermore, it is suggested that alleles encoding for proteins which provide protection against RNA viruses in modern Europeans are most likely inherited from the Neanderthals.<sup>16</sup> The conservation of such alleles indicates that they could have been protective to both early modern humans and Neanderthals when they were exposed to pathogens, such as viruses and bacteria, and that such pathogenic exposure contributed significantly in the shaping of the modern human immune system.<sup>13,16</sup> Human leukocyte antigen (HLA) class I and class II molecules are involved in antigen presentation to T-lymphocytes, which initiates the adaptive immune responses. The heterogeneity of HLA genes (*HLA-A*, *-B*, *-C*) showed a strong correlation with SARS-CoV-2 infection susceptibility, severity and progression.<sup>17</sup> It is estimated that the contribution of Neanderthal *HLA-A* alleles to non-Africans is 50-95%.<sup>13,18</sup> These findings point to a vital role of HLA molecules in the host's defense against viral infection and that these inherited HLA alleles provide an evolutionary advantage to modern human survival.<sup>13</sup> A disadvantage of Neanderthal admixture may however be susceptibility to immune diseases associated with HLA. Archaic HLA alleles may have been positively selected and preserved in modern genomes because they were beneficial to the survival of Neanderthal and early modern humans encountering similar local pathogens at the time.

However, some of the archaic alleles have been associated with development of autoimmune and auto-inflammatory diseases such as Behçet's disease in modern humans.<sup>19</sup> Furthermore, genetic variants inherited from Neanderthals have been shown to contain risk factors for a range of diseases including systemic lupus erythematosus, biliary cirrhosis, Crohn's disease, and type 2 diabetes (T2D).<sup>20</sup> Recently, it was also shown that a gene cluster on chromosome 19, which introgressed from the Neanderthal lineage, is a major risk factor for severe symptoms upon SARS-CoV-2 infection.<sup>21</sup> These studies suggest that the gene flow from Neanderthals could also negatively affect modern humans.

Evidently, the gene flow between early modern humans and archaic humans has undergone strong balancing selection during the course of human evolution. Thus, we wondered whether genes that have strong detrimental consequences on the function of our immune system were present in Neanderthals and introgressed to the modern human genome. In modern humans the diseases associated with errors in the immune system are called "inborn errors of immunity" (IEI). These include primary immunodeficiencies, autoimmune diseases and auto-inflammatory diseases. Currently, more than 450 IEI-associated genes are known to cause a profound impairment of the immune system and therefore have a significant impact on survival and quality of life.<sup>22</sup> In this study we screened the genomes of three Neanderthals and one Denisovan for the presence of potentially damaging genetic variants in genes associated with IEI to explore our hypothesis that defects in the immune system might have contributed to the disappearance of the archaic humans.

## Methods

Variant call format files from the Neanderthal from the Denisova Cave in the Altai mountains (*Denisova 5*, referred to as Altai Neanderthal) were downloaded (<http://cdna.eva.mpg.de/neandertal/altai/AltaiNeandertal/VCF/>) per chromosome and uploaded in our local Galaxy cloud.<sup>3,23</sup> Files were filtered for variants (removing reference genotype 0/0) and merged. Variants (6,374,229) were annotated using ANNOVAR.<sup>24</sup> The annotated variants file was exported and uploaded into TIBCO Spotfire Analyst 12.0.0 (<https://www.tibco.com/>) for filtering. The variants were filtered for IEI candidate genes (448 unique genes) based on the 2022 updated classification from the international union of immunological societies (IUIS) resulting in 76,516 variants.<sup>22</sup> Subsequently, these variants were filtered for protein affecting variants (PAV): exonic and splicing variants (336 variants). Next, the variants that were absent or present in less than 0.01 in the control cohorts (GnomAD\_exome, GnomAD\_genome, 1K, ExAC, and GoNL) were selected (134 variants) and exported to excel. The variants were filtered on combined annotation dependent depletion (CADD v1.6) scores.<sup>25</sup> Fifty-three variants had a CADD (v1.6) score  $\geq 20$  and 39 of these variants with a QUAL score (phred-scaled quality score) above 100 were selected. Among them, 24 variants were homozygous, and 15 variants were heterozygous in the

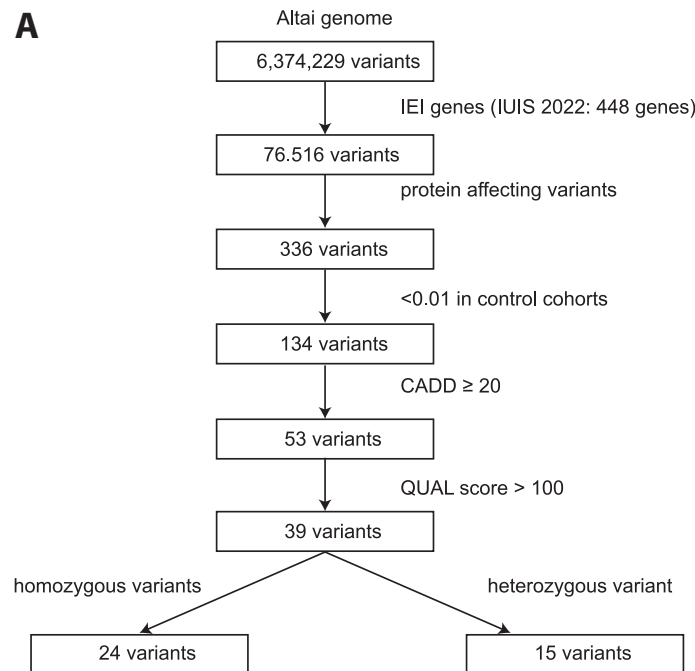
Altai Neanderthal genome. All variants were checked in the genome browser from the Max Planck Institute (<https://www.eva.mpg.de/genetics/genome-projects/> and <https://bioinf.eva.mpg.de/jbrowse/?loc=1%3A99708627..149558751&tracks=&highlight=>) for the presence of the variants in the genomes of the Neanderthal from the Vindija Cave in Croatia (*Vindija* 33.19 referred to as Vindija Neanderthal), and the Denisovan also from the Denisova Cave in the Altai mountains (*Denisova* 3 referred to as Denisovan).<sup>3,6,26</sup> In addition, the presence of all variants were checked in the genome from the Neanderthal found in the Chagyrskaya Cave in the Altai Mountains by uploading the BAM files (<http://ftp.eva.mpg.de/neandertal/Chagyrskaya/BAM/>)<sup>27</sup> into our Galaxy cloud, and visualized by integrative genome viewer (IGV) version 2.3.94.<sup>28</sup> Five variants (*MTHFD1*; *WRAP53*; *NOS2*; *PEPD*; *CFTR*) in our homozygous or heterozygous lists were not found in the Max Planck Institute genome browser therefore these genes were removed from further analysis. Finally presence of the variants were checked in modern human genomes (Dai (HGDP01308), French (HGDP00533), Han (HGDP00775), Mandenka (HGDP01286), Mbuti (HGDP00982), Papuan (HGDP00546), San (HGDP01036), Sardinian (HGDP01076), Yoruba (HGPD00936), Karitiana (HGDP01015), Australian (WON,M), Australian (BUR,E),

Mixe 0007 (MIXE 0007), and Dinka (DNK07)) from the Max Planck Institute genome browser and primate genomes (Chimp, Gorilla, Orangutan, Gibbon, Rhesus, Macaque, Baboon, Green Monkey, Marmoset, Squirrel Monkey, Bushbaby) from the UCSC genome browser. The human gene mutation database (HGMD) was used to check if the variants have been described before.

## Results

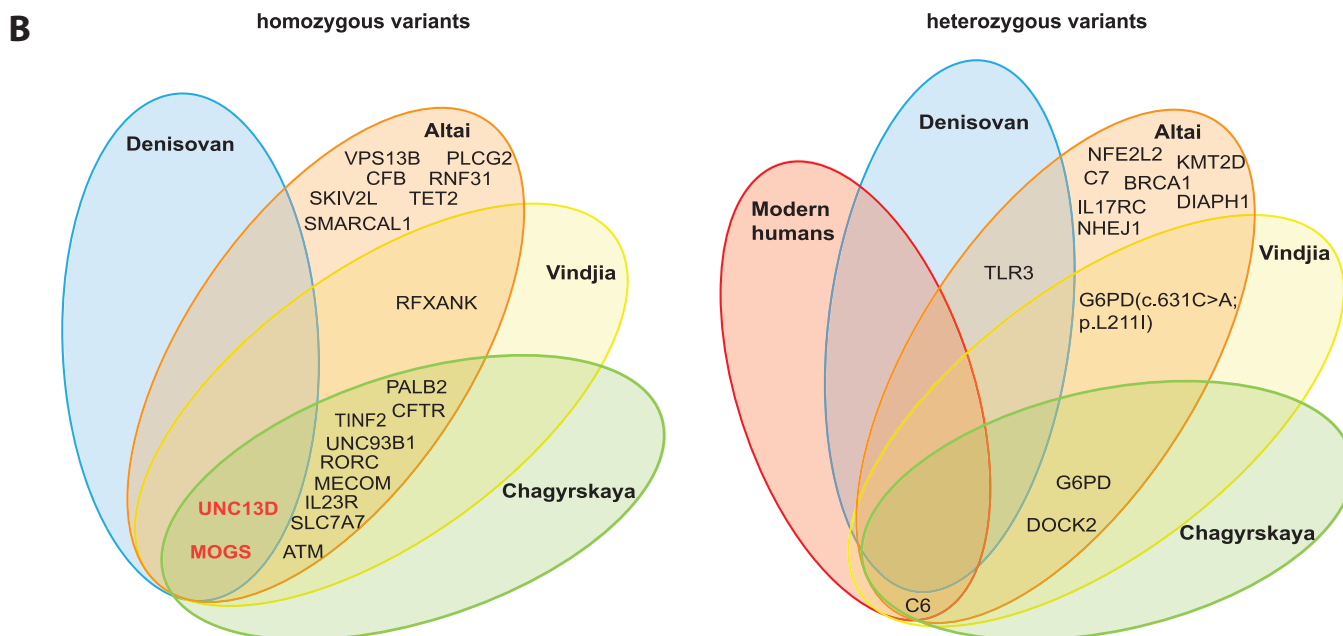
### Identification of variants in genes associated with IEI in Neanderthal and Denisovan genomes

To determine if Neanderthals and Denisovan genomes contained genetic variants in genes associated with IEI we used the variant call format files from a Neanderthal from the Altai mountains which contained 6,374,229 variants compared to the Homo sapiens genome assembly (hg19) (**Figure 1A**). From these variants, only the variants present in genes related with IEI were selected. These variants were filtered for protein affecting variants, the absence in control cohorts, with a CADD score equal or higher than 20, and with a QUAL score (phred-scaled quality score) above 100, leaving us with 39 variants in the Altai Neanderthal genome. Twenty-four of these variants were homozygous, and fifteen variants were heterozygous (**Figure 1A**).



**Figure 1. Genetic variants associated with IEIs.**

Filtering scheme for the identification of variants related to IEIs in archaic humans (Altai, Vindija, Chagyrskaya Neanderthals; Denisovan); modern humans and primate genomes (A).



**Figure 1. (Continued)**

Venn diagrams of the 24 IEI-related homozygous (left panel) and 15 heterozygous (right panel) variants found in Altai, Vindija, Chagyrskaya Neanderthals; Denisovan; modern humans and primate genomes (B). Genes with variants that are common to all archaic humans (Altai, Vindija, Chagyrskaya Neanderthals; Denisovan) are highlighted in red.

Subsequently, we checked if these 24 homozygous and 15 heterozygous variants from the Altai Neanderthal were present in the published genomes of the Vindija Neanderthal, Chagyrskaya Neanderthal, and Denisovan genome. In addition, we checked if the variants were present in modern human genomes or primates. While none of the homozygous and heterozygous variants were found in the primate genomes, one variant (NM\_001115131(C6):c.1848G>A,p.M616I) in the 15 heterozygous variants was also present in some modern genomes.

#### **Homozygous variants in IEI-related genes**

Interestingly, we identified three variants in two different genes that were homozygous in all archaic humans genomes (Table 1 and Supplementary Table 1). Both genes (*UNC13D* and *MOGS*) have been associated with an autosomal recessive form of IEI. Two of the variants were located in *UNC13D* gene NM\_199242(*UNC13D*):c.1889C>T,p.A630V (rs372467471, CADD:26.2) and NM\_199242(*UNC13D*):c.2828A>G,p.N943S (rs147748627, CADD:24.3). Both variants have very low allele frequency in the genome aggregation database (gnomAD) (0.000004063 and 0.005 respectively). Autosomal recessive variants in *UNC13D* cause hemophagocytic lymphohistiocytosis (HLH), a rare, life-threatening hyper-inflammatory immune disorder.<sup>29</sup> HLH is characterized by uncontrolled activation of macrophages, T lymphocytes, and increased inflammatory cytokine production (cytokine storm) that leads to multi-organ failure. The *UNC13D* variants are in the Munc13-homology domain 1 (MHD1) and the Ca<sup>2+</sup>-binding (C2B) domain of Munc13-4 protein respectively (Figure 2A and Figure 3A). The *UNC13D* variant p.(A630V) has not been previously reported in modern human genome, but the p.(N943S) variant has been described in 2014 in two patients presenting with HLH.<sup>30</sup>

The other variant that was present in all archaic genomes is located in the *MOGS* gene (Figure 2B and Figure 3B), which encodes for  $\alpha$ -glucosidase 1, also known as mannosyl-oligosaccharide glucosidase (MOGS). Loss-of-function defects in *MOGS* underlie MOGS-CDG, a group of disorders caused by inborn errors of glycosylation, previously known as congenital disorders of glycosylation type II (CDG-IIb; OMIM;606056). Pathogenic variants of *MOGS* influence the processing of protein-bound glycans as well as glycoprotein folding and quality control.<sup>31</sup> Additionally, *MOGS* is a known regulator of the HLA-class I antigen presentation pathway.<sup>32</sup> The variant in the *MOGS* gene (NM\_006302(*MOGS*):c.1484G>A,p.R495Q (rs34075781, CADD: 23.7)) is located in the C-terminal catalytic domain (Figure 2B). This variant has a 0.0001 allele frequency in controls and was reported as a variant of unknown significance (VUS) in 2021 in a patient with common variable immunodeficiency (CVID).<sup>33</sup>

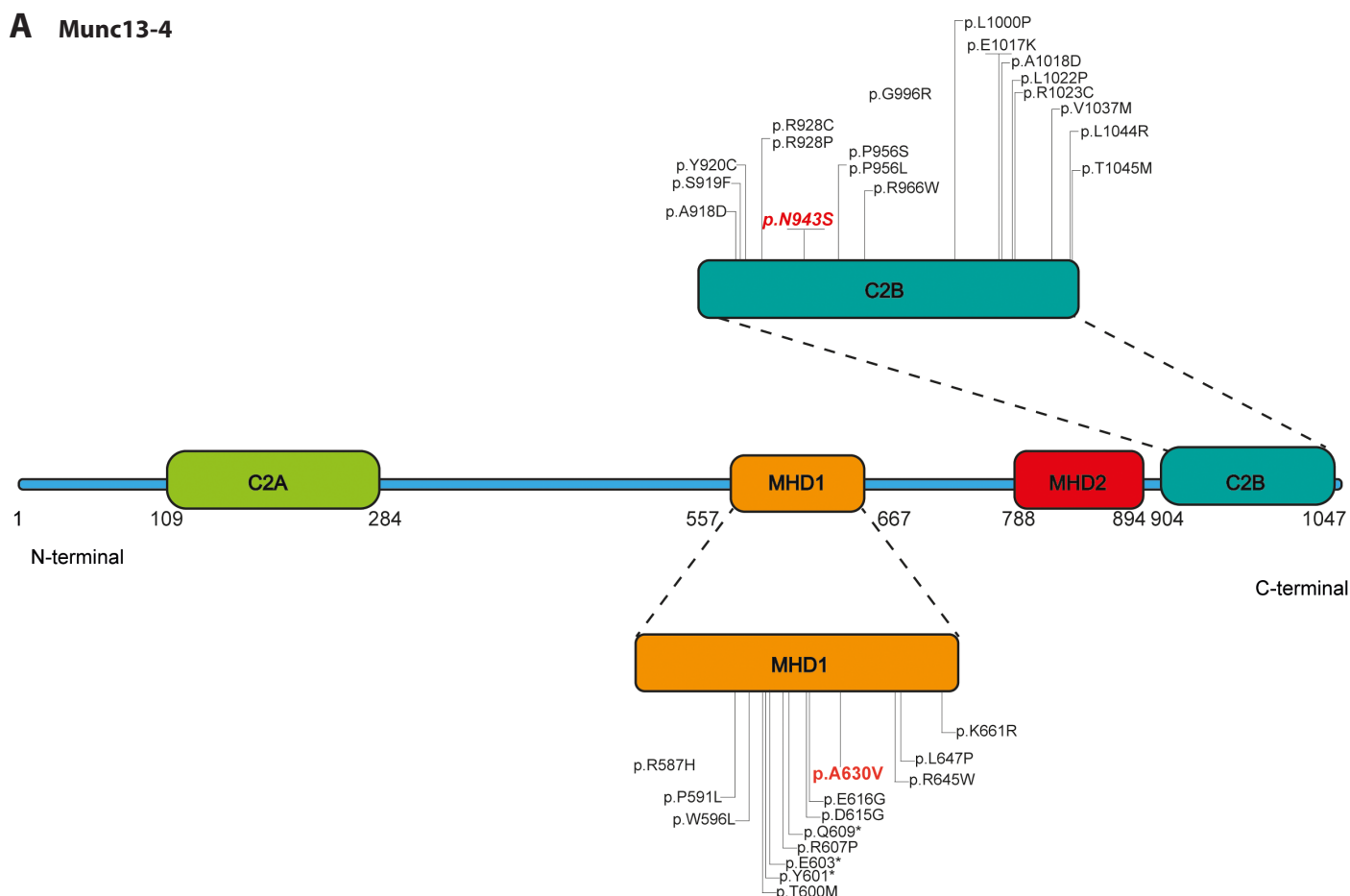
Nine homozygous variants were present in all Neanderthal genomes (Altai, Vindija, and Chagyrskaya), but not in the Denisovan genome. Most of these variants are of unknown significance because they have not been described before. However, the amino acid affected by the variant in the *ATM* gene (NM\_000051(*ATM*):c.1967C>A,p.T656N) has been previously described to be affected in patients with Ataxia Telangiectasia (p.T656P).<sup>34</sup> One variant on a gene which is associated with HLA-class II deficiency (NM\_001278727 (*RFXANK*):c.155C>A;p.P52H) was shared by the Altai and Vindija Neanderthals but absent in the Chagyrskaya Neanderthal genome. Eight variants were only present in the Altai and Vindija Neanderthals, or even only in the Altai genome, indicating differential expression in archaic humans. None of these have been described before.

**Table 1. Homozygous variants in IEI-related genes in the archaic genomes.**

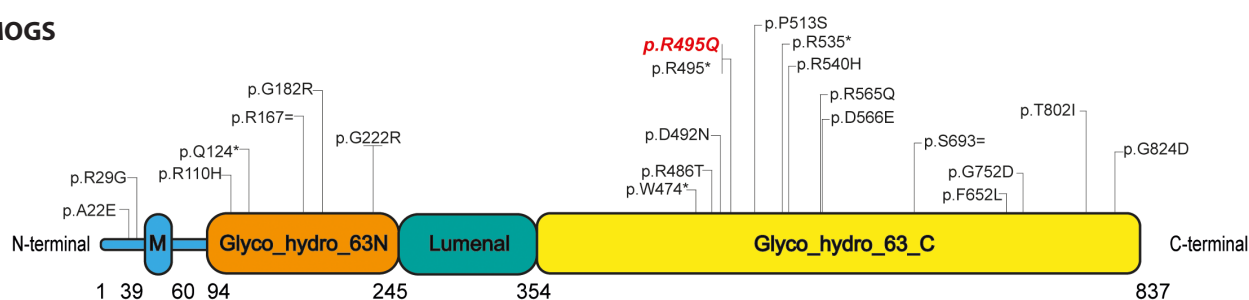
Gene	variant	CADD	OMIM	inheritance	Major category by IUIS related to the gene	Subcategory IUIS related to the gene
<b>Only in ancients (Altai, Vindija, Chagyrskaya, and Denisovan)</b>						
MOGS	NM_006302(MOGS):c.1484G>A,p.R495Q	23.7	601336	AR	Predominantly Antibody Deficiencies	CVID Phenotype
UNC13D	NM_199242(UNC13D):c.1889C>T,p.A630V	26.2	608897	AR	Diseases of Immune Dysregulation	HLH
UNC13D	NM_199242(UNC13D):c.2828A>G,p.N943S	24.3				
<b>Only in Neanderthal (Altai, Vindija, Chagyrskaya)</b>						
RORC	NM_005060(RORC):c.373C>T,p.R125W	25.5	602943	AR	Defects in intrinsic and innate immunity	MSMD
IL23R	NM_144701(IL23R):c.726G>T,p.W242C	25	607562	AR	Defects in intrinsic and innate immunity	MSMD
MECOM	NM_001105077(MECOM):c.88G>T,p.A30S	22.3	165215	AD	Bone marrow failure	
CFTR	NM_000492(CFTR):c.1963G>A,p.A655T	23.8	602421	AR	Congenital defects of phagocyte number or function	Defects of Motility
UNC93B1	NM_030930.4(UNG93B1):c.626C>T,p.P209L	23.3	608204	AR	Defects in intrinsic and innate immunity	Herpes Simplex Encephalitis (HSE)
ATM	NM_000051(ATM):c.1967C>A,p.T656N	23.5	607585	AR	Combined immunodeficiencies with associated or syndromic features	DNA Repair Defects
TINF2	NM_012461(TINF2):c.721C>T,p.P241S	22.9	604319	AD	Bone marrow failure	
SLC7A7	NM_001126106(SLC7A7):c.619G>A,p.G207S	21.5	603593	AR	Diseases of Immune Dysregulation	HLH
PALB2	NM_024675(PALB2):c.3068A>G,p.Q1023R	22.8	610355	AR	Bone marrow failure	
<b>Only in Altai and Vindija Neanderthals</b>						
RFXANK	NM_001278727(RFXANK):c.155C>A,p.P52H	21.5	603200	AR	Immunodeficiencies affecting cellular and humoral immunity	CID
<b>Only in Altai Neanderthal</b>						
SMARCAL1	NM_001127207(SMARCAL1):c.1880C>T,p.S627F	28	606622	AR	CID with associated or syndromic features	Immuno-osseous Dysplasias
TET2	NM_001127208(TET2):c.2960G>A,p.C987Y	22	612839	AR	Diseases of Immune Dysregulation	Susceptibility to EBV and Lymphoproliferative Conditions
CFB	NM_001710(CFB):c.2038T>C,p.F680L	28.9	138470	AR	Complement Deficiencies	
SKIV2L	NM_006929(SKIV2L):c.2347G>A,p.E783K	23.3	600478	AR	CID with associated or syndromic features	Other CID with syndromic features
VPS13B	NM_152564(VPS13B):c.8844C>G,p.I2948M	21	607817	AR	Congenital defects of phagocyte number or function	Congenital Neutropenias
RNF31	NM_001310332(RNF31):c.2759G>A,p.R920Q	23.2	612487	AR	CID with associated or syndromic features	Other CID with syndromic features
PLCG2	NM_002661(PLCG2):c.1526T>C,p.I509T	20.5	600220	AD GOF	Autoinflammatory Disorders	Non-Inflammasome Related Conditions

AR, autosomal recessive; AD, autosomal dominant; CVID, common variable immunodeficiency; MSMD, mendelian susceptibility to mycobacterial disease; HLH, hemophagocytic lymphohistiocytosis; CID, combined immunodeficiency; EBV, Epstein-Barr virus;

## A Munc13-4

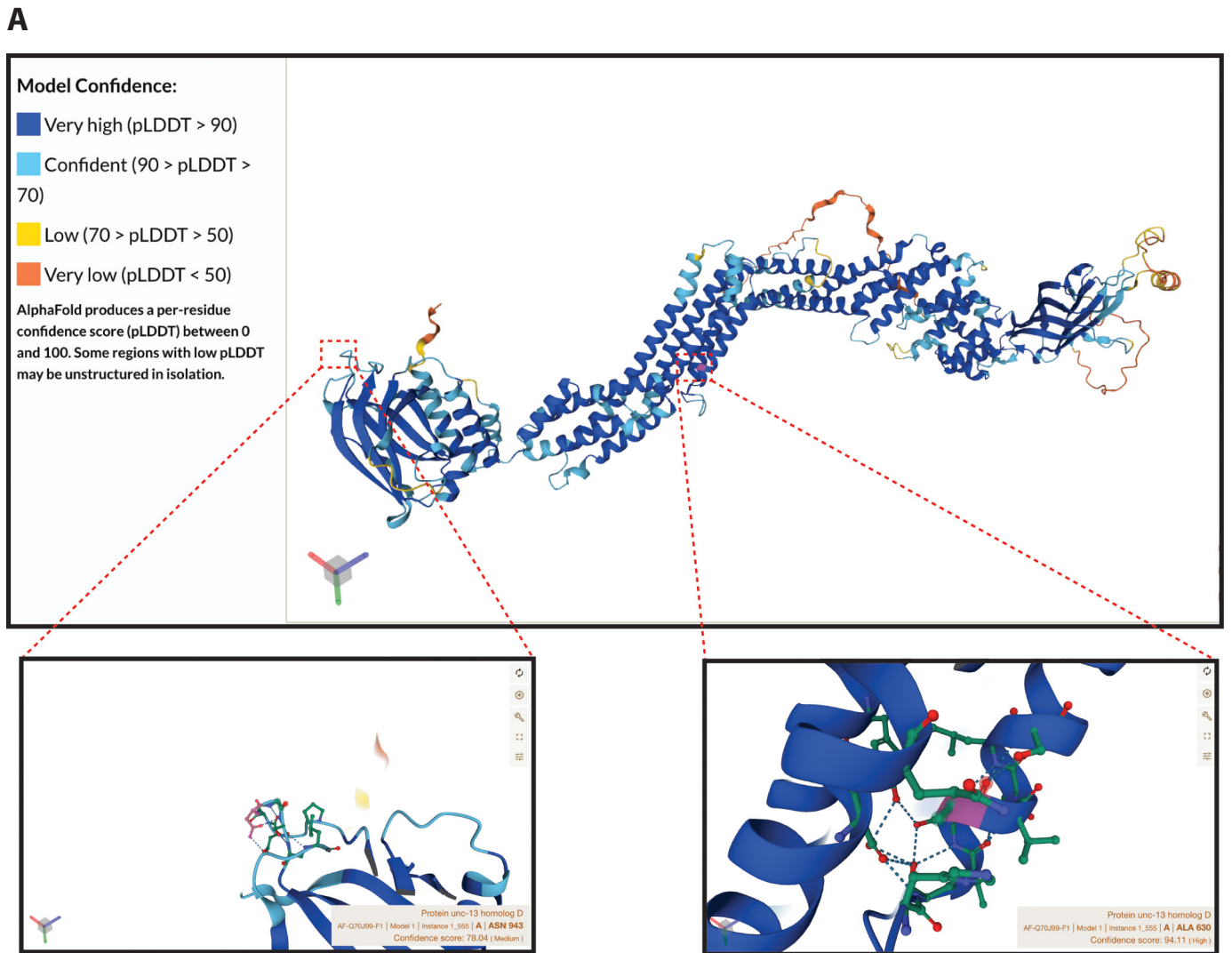


## B MOGS



**Figure 2. Schematic presentation of Munc13-4 and MOGS protein structure.**

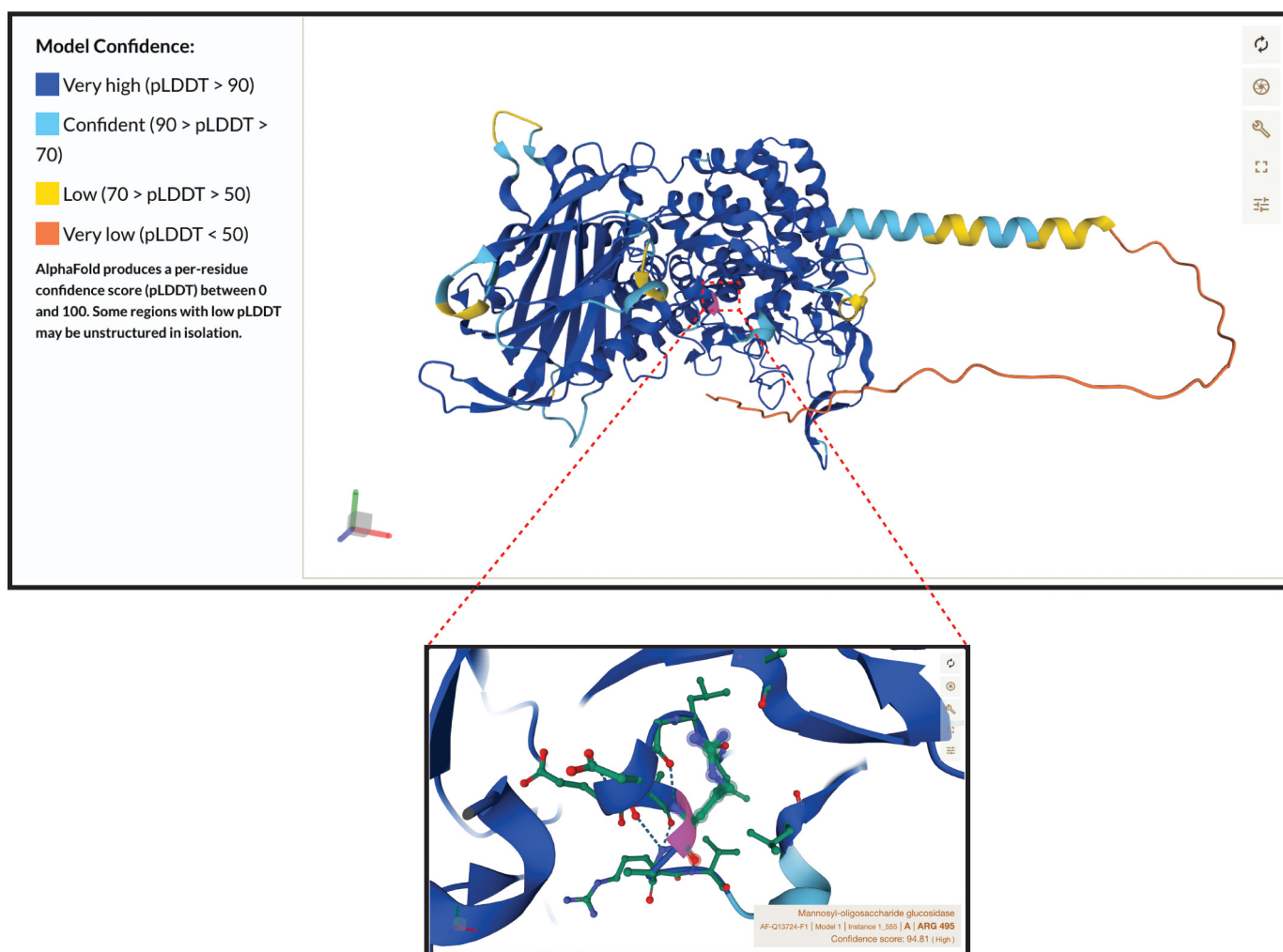
Linear representation of the human Munc13-4 protein. Variants associate with HLH in the Munc13-homology domain 1 (MHD1) and the  $\text{Ca}^{2+}$ -binding (C2B) domain are depicted (A). Linear representation of the human MOGS protein. Variants in the domain which provides Signal-anchor for type II membrane protein is marked M and in blue; the N-terminal beta sandwich domain found in glycosyl hydrolase family 63 proteins is marked Glyco\_hydro\_63N and in orange; the C-terminal catalytic domain found in glycosyl hydrolase family 63 proteins is marked Glyco\_hydro\_63C and in yellow (B). Variants described in modern humans are italicized; homozygous variants identified in archaic humans (Altai, Vindija, Chagyrskaya Neandertals; Denisovan) is coloured red.



**Figure 3. 3D structure of human Munc13-4 and MOGS proteins.**

Predicted 3D structure of the human Munc13-4 protein extracted from structural data in UniProtKB. A magnified view of specific residue sites of identified variants of *UNC13D* (p.A630V and p.N943S) are highlighted below (A).

B



**Figure 3. (Continued)**

Predicted 3D structure of the human MOGS protein extracted from structural data in UniProtKB. A magnified view of specific residue site of identified variant of MOGS (p.R495Q) is highlighted below (B).

### Heterozygous variants in IEI-related genes

Fifteen variants in IEI-related genes were heterozygous in the Altai genome, but none of these were present in all archaic humans (Table 2 and Supplementary Table 2). However, three variants were present in all the Neanderthals (Altai, Vindija and Chagyrskaya). Interestingly, two of three variants are in the *G6PD* gene. The NM\_001042351 (*G6PD*):c.620T>C:p.M207TT variant has not been reported yet. The NM\_001042351 (*G6PD*):c.590A>T:p.Y197F variant in *G6PD* has also not been described, but another variant on the same amino acid (p.Y197S) has been described in boys with glucose-6-phosphate dehydrogenase deficiency.<sup>35</sup> The third variant present in all Neanderthals is in the *DOCK2* gene (NM\_004946(*DOCK2*):c.1855G>T:p.G619C). This gene has been recently reported to be involved in immune response to SARS-Cov2 infection and the development of severe COVID-19.<sup>36</sup> The inheritance of *G6PD* is X-linked, so the variants in *G6PD* could have potentially affected the archaic males.

Most of the other heterozygous variants are related to autosomal recessive traits, so their clinical impact remains highly uncertain. However, the variant in Toll-like receptor 3 (*TLR3*) (NM\_003265(*TLR3*):c.2276A>T:p.Y759F) shared by the Altai Neanderthal and Denisovan has an autosomal dominant inheritance. Loss-of-function disease-causing variants in *TLR3* have also been described in patients with herpes encephalitis (autosomal dominant and recessive) and severe COVID-19 infection due to impaired type I IFN induction.<sup>37,38</sup> The other two genes that are associated with autosomal dominant inheritance are the *KMT2D* and *NFE2L2* genes, which were only found in the Altai Neanderthal, and both are associated with combined immunodeficiency with syndromic features (Table 2). The exact variants in these genes have not been described, but a variant on the same amino acid as the p.A5212T variant in the *KMT2D* gene has been described in a patient with Kabuki syndrome.<sup>39</sup>



**Table 2. Heterozygous variant in IEI-related genes in the archaic genomes.**

Gene	Variant	CADD	OMIM	Inheritance	Major category by IUIS related to the gene	Subcategory IUIS related to the gene
<b>Only in Neanderthals (Altai, Vindija, Chagyrskaya)</b>						
DOCK2	NM_004946(DOCK2):c.1855G>T,p.G619C	32	603122	AR	Immunodeficiencies affecting cellular and humoral immunity	CID Generally Less Profound than Severe Combined Immunodeficiency
G6PD	NM_001042351(G6PD):c.620T>C,p.M207T	25.7	305900	XL	Congenital defects of phagocyte number or function	Defects of Respiratory Burst
G6PD	NM_001042351(G6PD):c.590A>T,p.Y197F	23.8				
<b>Only in Altai and Vindija</b>						
G6PD	NM_001042351(G6PD):c.631C>A,p.L211I	24.7	305900	XL	Congenital defects of phagocyte number or function	Defects of Respiratory Burst
<b>Only in Altai Neanderthal and Denisovan</b>						
TLR3	NM_003265(TLR3):c.2276A>T,p.Y759F	23.4	603029	AD	Defects in intrinsic and innate immunity	Herpes Simplex Encephalitis (HSE)
<b>Only in Altai Neanderthal</b>						
DIAPH1	NM_001079812(DIAPH1):c.1337G>A,p.R446Q	27.3	602121	AR	CID with associated or syndromic features	Other CID with syndromic features
C7	NM_000587(C7):c.450G>C,p.Q150H	24.4	217070	AR	Complement Deficiencies	
IL17RC	NM_001203263(IL17RC):c.592G>A,p.E198K	23.3	610925	AR	Defects in intrinsic and innate immunity	Predisposition to Mucocutaneous Candidiasis
NHEJ1	NM_024782(NHEJ1):c.79G>A,p.V27I	22.4	611290	AR	Immunodeficiencies affecting cellular and humoral immunity	T-B- SCID
KMT2D	NM_003482(KMT2D):c.15634G>A,p.A5212T	27.8	602113	AD	CID with associated or syndromic features	Other CID with syndromic features
NFE2L2	NM_001313903(NFE2L2):c.887C>T,p.T296I	22	600492	AD	CID with associated or syndromic features	Other CID with syndromic features
BRCA1	NM_007299(BRCA1):c.1684T>G,p.Y562D	21.1	113705	AR	Bone marrow failure	
<b>Also present in modern human genomes</b>						
C6	NM_001115131(C6):c.1848G>A,p.M616I	23.2	217050	AR	Complement Deficiencies	

AR, autosomal recessive; XL, X-linked; AD, autosomal dominant; CID, combined immunodeficiency; SCID, severe combined immunodeficiency

## Discussion

In this study we investigated whether genetic variants that could have detrimental effects on the function of the immune system were present in the genomes of three Neanderthals and one Denisovan. We hypothesized that those variants, upon specific pathogenic exposure, could contribute to the extinction of the Neanderthal population. We first determined whether we could identify potentially damaging variants in the IEI-related genes in the genome of the Altai Neanderthal, and subsequently checked which of these variants were also present in the genomes of the Vindija Neanderthal, Chagyrskaya Neanderthal and the Denisovan. In total, we identified 24 homozygous and 15 heterozygous variants in the Altai Neanderthal from the human IEI gene panel.<sup>22</sup>

Most of these variants were not present in the other archaic human genomes studied, so it is unlikely that these variants were strong determinants in the extinction of Neanderthals. In contrast, we identified three homozygous variants in two different genes that were shared by all four archaic human genomes studied, while these variants were not present in the genomes of modern humans nor in the genomes of primates. This is interesting because the three Neanderthals and the Denisovan lived in different areas and time periods. The Altai Neanderthal and the Denisovan were found in the same Denisova Cave in the Altai Mountains in Siberia, but they were found in different sublayers. The Denisovan should be older than the Altai Neanderthal which was estimated to

live  $50,300 \pm 2,200$  years ago.<sup>3</sup> The Chagyrskaya Neanderthal was also found in a cave in the Altai mountains, but this cave was 100 km away from the Denisova cave and genetic analysis showed that she was more related to Neanderthals from western Eurasia.<sup>27</sup> The archaeological layer in which the Chagurskaya Neanderthal was found, was estimated to be 60,000 years old. The Vindija Neanderthal was found in a cave in Croatia and is estimated to be 45,500 years old.<sup>26</sup> The fact that all three variants were present in all archaic genomes suggest that they must have been present before the different lineages diverged.

*UNC13D* encodes for Munc13-4, which is a crucial protein for the function of NK cells and cytotoxic T lymphocytes ( $CD8^+$ ) to destruct virus-infected cells and tumor cells, but also for the regulation of activated macrophages when a pathogen is cleared.<sup>40</sup> Loss-of-function variants in *UNC13D* result in HLH, a complex inflammatory and often severe and fatal disease.<sup>29,40</sup> HLH is triggered by infections as NK-cell and cytotoxic T-cell dysfunction prevent efficient antigen clearance. Because of the defects in *UNC13D*, NK cells and cytotoxic T lymphocytes struggle with the successful elimination of pathogens, and in addition, fail to lyse the activated macrophages. This results in haemophagocytosis of host platelets, leucocytes and red blood cells. In addition, the (over)activation of the cytotoxic T lymphocytes, NK cells, and macrophages induces a cytokine storm in which large amounts of cytokines including IL-1 $\beta$ , IL-6, IL-18, IFN- $\gamma$ , and TNF- $\alpha$  are produced and released.<sup>29</sup> Patients with HLH usually have poor prognosis,<sup>29</sup> with the current treatments, mortality rates among HLH in adults are around 60%.<sup>41</sup>

The two homozygous *UNC13D* variants that were only found in the archaic human genomes were in the MHD1 region (c.1889C>T;p.A630V) and in the C2B region (c.2828A>G;p.N943S) in the Munc13-4 protein. Both variants are reported in control cohorts, however at a very low allele frequency (c.2828A>G;p.N943S; allele frequency: 0.0006630, c.1889C>T;p.A630V; allele frequency: 0.000007362). The c.1889C>T variant has not been described in patients with HLH before. Nonetheless, the c.2828A>G variant was described in two patients with HLH.<sup>30</sup> These patients were heterozygous for this variant, and expressed an additional heterozygous variant in another gene associated with HLH, suggesting a digenic mode of inheritance that results in a synergistic defect.<sup>30</sup> It is noteworthy that the two variants in *UNC13D* did not introgress into the modern humans, implying that these variants may have a negative effect on the immune system of archaic humans after exposure to certain microbes. Although there are no patients with HLH described who are homozygous for one of the two *UNC13D* variants identified in our study, the data so far suggests that it is very likely that both variants, and especially the c.2828A>G variant could have impaired the function of Munc13-4 in the archaic humans.

HLH in patients with a defect in *UNC13D* is usually triggered by an infection. Infections with *Mycobacterium tuberculosis* (*Mtb*), Epstein-Barr virus, Herpes simplex virus and Varicella zoster virus have been described to trigger HLH.<sup>42,43</sup> One of the well-known natural reservoir hosts for viruses are bats. Numerous viruses that can cause serious

illness in humans have been identified in them, such as SARS-Cov2.<sup>44</sup> It is not surprising that the cave-residing Neanderthals have co-lived with bats as bat fossils have been discovered in caves where tools typically used by Neanderthals were found.<sup>45</sup> Interestingly, since the COVID-19 pandemic, SARS-Cov2 infection has become a known cause of secondary HLH.<sup>46</sup> Therefore, bats may have been one of the sources of new pathogen exposure which could have triggered serious damaging immunological events in Neanderthals with a genetic predisposition. In Neanderthal DNA, fragments of various infectious agents were detected in dental pulp from Neanderthals, including bacteria from the Actinobacter phylum.<sup>47</sup> One of the most studied bacteria from the Actinobacteria phylum is *Mtb*.

The evolutionary timing and spread of the *Mtb* was reported by Wirth et al, using mycobacterial tandem repeat sequences as genetic markers.<sup>48</sup> Using Bayesian statistics and experimental data on the variability of the mycobacterial markers in infected patients, they estimated the age of the *Mtb* at 40,000 years. This coincides with the expansion of early modern human populations out of Africa. An interesting hypothesis could be that migration of early modern humans out of Africa carried along *Mtb*, spreading the pathogen across Eurasia, thus exposing archaic humans such as Neanderthals to this “new” pathogen. As indicated previously, induction of HLH needs a trigger. This was also shown in a mouse model of HLH. The mice that had a donor splice site in *Unc13d*, the mouse orthologue of human Munc13-4, did not spontaneously develop an HLH phenotype, but only upon a specific infectious trigger.<sup>49</sup> It could thus be that “new” pathogens like *Mtb* were the specific trigger that caused an HLH phenotype in the archaic humans.

Glycosylation is one of the most fundamental protein modification processes in all living organisms as the majority of secreted and membrane-bound proteins are glycoproteins. N-glycosylation refers to the attachment of glycans to the nitrogen atom of an asparagine residue in a protein, whereas O-glycosylation is the attachment of glycans to the oxygen atom of serine or threonine residues. N-glycosylation is the most common co-translational/post-translational protein modification in human cells.<sup>31</sup> MOGS is expressed ubiquitously in the endoplasmic reticulum, where it functions as the first enzyme involved in the trimming of N-glycans in the N-glycosylation pathway. This initial step is crucial to the further processing of N-glycans. When the enzyme is damaged, various important biological processes are impaired, including protein folding, stability and transition. Genetic loss-of-function variants in the *MOGS* gene causes the extremely rare congenital disorder of glycosylation type IIb (CDG-IIb), also known as MOGS-CDG. MOGS-CDG presents as a paradoxical immunologic phenotype associated with severe hypogammaglobulinemia, in absence of overt clinical indications of infection along with various neurologic complications.<sup>50</sup> The mechanism behind the hypogammaglobulinemia was suggested as an unexplained shorter immunoglobulin half-life. Resistance to viral infections such as HIV and influenza viruses could be explained by impaired viral replication and cellular entry

of these enveloped viruses which are heavily glycosylated.<sup>50</sup> Glycoproteins are vital for microbes itself. The SARS-CoV2 spike protein, for instance, is an extensively glycosylated protein which protrudes from the viral surface and mediates host-cell entry by binding to angiotensin-converting enzyme 2 (ACE2).<sup>51</sup> Yet on the other hand, glycosylation is also a factor in the immune defense against microorganisms and pathogens. Numerous effector systems of the immune system, including cytokines, cytokine receptors, antibodies and HLA require fine-tuned glycosylation for their optimal function.

In 2014, Pääbo et al. reported the same *MOGS* variant ((NM\_006302):c.1484G>A,p.R495Q (rs34075781; CADD: 23.7) in relation to bone morphology, abnormalities in palates in the genomes of three Neanderthals (El Sidrón, Vindija, and Altai genomes).<sup>7</sup> The variant is located in the C-terminal catalytic domain which catalyzes the cleavage of N-glycans. This variant is ultra-rare (with allele frequency of up to 0.001%) in control cohorts, and it was previously described once in a COVID patient, as a heterozygous VUS. *MOGS* is one of the key genes regulating HLA-I expression and plays a significant role in immune defense by regulating immunoglobulin half-life, viral replication, and cell entry.

The absence of the homozygous p.R495Q variant in the modern human population, suggests that this variant may be damaging upon exposure to specific pathogens.

The relationship between archaic alleles and disease in modern humans remains unclear. Since 2019, the SARS-Cov2 pandemic has provided evidence linking the disease severity to genes inherited from Neanderthals.<sup>21,52</sup> In South Asians, especially Bangladeshis, a Neanderthal haplotype of a gene cluster on Chromosome 3 is adaptively retained and is linked to a more severe response to SARS-CoV-2 infection.<sup>21</sup> Increased levels of a Neanderthal haplotype of 2'5' oligoadenylate synthase (*OAS1*) in Europeans, on the other hand, confers protection against Sars-CoV-2 infection.<sup>52</sup>

In conclusion, our study is the first to specifically describe the presence of IEI-related gene-variants in the genomes of Neanderthals and Denisovan. We discovered three variants on two genes that are linked to rare IEI disorders in modern humans. All three of these variants were present in all the explored genomes of Neanderthals and Denisovan, but absent from modern human genomes (or were found heterozygous and in extremely low allele frequencies) and other primate genomes. This observation strongly suggests a deleterious effect of these variants on the immune system of archaic



**Figure 4. Illustration of the hypothesis on transmission of tropical pathogens by early modern humans to Eurasia expedited the extinction of Neanderthals carrying possible damaging variants in IEI-related genes.**

humans. Neanderthals and Denisovans survived a long period of time until, shortly after (by evolutionary standards) the Paleolithic migrations of modern humans (*Homo sapiens*). Studies by Houldcroft et al, indicated that infectious diseases are tens of thousands of years older than previously believed and thus co-evolved with humans for long time.<sup>53</sup> They hypothesized that spreading of pathogens out of Africa could have been catastrophic to Neanderthal populations in Eurasia. Our findings add a novel layer to this with the identification of archaic-human-specific variants in genes associated with IEI in modern human. We therefore postulate that: 1); these archaic-human-specific IEI gene-variants could be detrimental and cause severe and fatal disease upon encounter of pathogens brought along by modern human, 2); such unfortunate host-pathogen interactions may have expedited the extinction of Neanderthals (**Figure 4**).

### Acknowledgements

The research for this manuscript was performed within the framework of the Erasmus Postgraduate School of Molecular Medicine. This work was supported by the China Scholarship Council for funding PhD fellowships. (No.201908440363 Z.Z). The authors have no conflict of interest.

### Author contributions

- PMvH designed the study.
- ZZ, SS, HIJ, PMvH acquired data, analyzed data, and drafted the manuscript.
- ZZ, SS, ML, NS, PS, VD, WD, HIJ, PMvH interpreted the data and contributed substantially to manuscript revisions.
- All the authors approved the final version submitted for publication.

### References

1. Duveau J, Berillon G, Verna C, Laisne G, Cliquet D. The composition of a Neanderthal social group revealed by the hominin footprints at Le Rozel (Normandy, France). *Proc Natl Acad Sci U S A*. 2019;116:19409-14.
2. Reich D, Patterson N, Kircher M, Delfin F, Nandineni MR, Pugach I, et al. Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania. *Am J Hum Genet*. 2011;89:516-28.
3. Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, et al. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*. 2014;505:43-9.
4. Meyer M, Arsuaga JL, de Filippo C, Nagel S, Aximu-Petri A, Nickel B, et al. Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. *Nature*. 2016;531:504-7.
5. Gomez-Robles A. Dental evolutionary rates and its implications for the Neanderthal-modern human divergence. *Sci Adv*. 2019;5:eaaw1268.
6. Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S, et al. A high-coverage genome sequence from an archaic Denisovan individual. *Science*. 2012;338:222-6.
7. Castellano S, Parra G, Sanchez-Quinto FA, Racimo F, Kuhlwil M, Kircher M, et al. Patterns of coding variation in the complete exomes of three Neanderthals. *Proc Natl Acad Sci U S A*. 2014;111:6666-71.
8. Higham T, Douka K, Wood R, Ramsey CB, Brock F, Basell L, et al. The timing and spatiotemporal patterning of Neanderthal disappearance. *Nature*. 2014;512:306-9.
9. Wall JD, Yang MA, Jay F, Kim SK, Durand EY, Stevison LS, et al. Higher levels of neanderthal ancestry in East Asians than in Europeans. *Genetics*. 2013;194:199-209.
10. Sankararaman S, Mallick S, Dannemann M, Prüfer K, Kelso J, Paabo S, et al. The genomic landscape of Neanderthal ancestry in present-day humans. *Nature*. 2014;507:354-7.
11. Simonti CN, Vernot B, Bastarache L, Bottinger E, Carrell DS, Chisholm RL, et al. The phenotypic legacy of admixture between modern humans and Neanderthals. *Science*. 2016;351:737-41.
12. Dannemann M, Kelso J. The Contribution of Neanderthals to Phenotypic Variation in Modern Humans. *Am J Hum Genet*. 2017;101:578-89.
13. Abi-Rached L, Jobin MJ, Kulkarni S, McWhinnie A, Dalva K, Gragert L, et al. The shaping of modern human immune systems by multiregional admixture with archaic humans. *Science*. 2011;334:89-94.
14. Deschamps M, Laval G, Fagny M, Itan Y, Abel L, Casanova JL, et al. Genomic Signatures of Selective Pressures and Introgression from Archaic Hominins at Human Innate Immunity Genes. *Am J Hum Genet*. 2016;98:5-21.
15. Kerner G, Patin E, Quintana-Murci L. New insights into human immunity from ancient genomics. *Curr Opin Immunol*. 2021;72:116-25.
16. Enard D, Petrov DA. Evidence that RNA Viruses Drove Adaptive Introgression between Neanderthals and Modern Humans. *Cell*. 2018;175:360-71 e13.
17. Migliorini F, Torsiello E, Spiezia F, Oliva F, Tingart M, Maffulli N. Association between HLA genotypes and COVID-19 susceptibility, severity and progression: a comprehensive review of the literature. *Eur J Med Res*. 2021;26:84.
18. Parham P, Moffett A. Variable NK cell receptors and their MHC class I ligands in immunity, reproduction and human evolution. *Nat Rev Immunol*. 2013;13:133-44.
19. Smith R, Moots RJ, Murad M, Wallace GR. A Darwinian view of Behçet's disease. *Rheumatology and Immunology Research*. 2021;2:91-9.
20. Sankararaman S, Mallick S, Dannemann M, Prüfer K, Kelso J, Pääbo S, et al. The genomic landscape of Neanderthal ancestry in present-day humans. *Nature*. 2014;507:354-7.
21. Zeberg H, Paabo S. The major genetic risk factor for severe COVID-19 is inherited from Neanderthals. *Nature*. 2020;587:610-2.
22. Tangye SG, Al-Herz W, Bousfiha A, Cunningham-Rundles C, Franco JL, Holland SM, et al. Human Inborn Errors of Immunity: 2022 Update on the Classification from the International Union of Immunological Societies Expert Committee. *J Clin Immunol*. 2022;
23. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res*. 2018;46:W537-W44.
24. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38:e164.
25. Rentsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*. 2019;47:D886-D94.
26. Prüfer K, de Filippo C, Grote S, Mafessoni F, Korlevic P, Hajdinjak M, et al. A high-coverage Neanderthal genome from Vindija Cave in Croatia. *Science*. 2017;358:655-8.
27. Mafessoni F, Grote S, de Filippo C, Slon V, Kolobova KA, Viola B, et al. A high-coverage Neanderthal genome from Chagyrskaya Cave. *Proc Natl Acad Sci U S A*. 2020;117:15132-6.
28. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24-6.
29. Janka GE. Familial and acquired hemophagocytic lymphohistiocytosis. *Annu Rev Med*. 2012;63:233-46.
30. Zhang K, Chandrakasan S, Chapman H, Valencia CA, Husami A, Kissell D, et al. Synergistic defects of different molecules in the cytotoxic pathway lead to clinical familial hemophagocytic lymphohistiocytosis. *Blood*. 2014;124:1331-4.
31. Peiwei Z, Peng X, Luo S, Huang Y, Tan L, Shao J, et al. Identification and characterization of novel mutations in MOGS in a Chinese patient with infantile spasms. *Neurogenetics*. 2020;21:97-104.
32. Jongsma MLM, de Waard AA, Raaben M, Zhang T, Cabukusta B, Platzer R, et al. The SPPL3-Defined Glycosphingolipid Repertoire Orchestrates HLA Class I-Mediated Immune Responses. *Immunity*. 2021;54:132-50 e9.
33. Bisgin A, Sonmezler O, Boga I, Yilmaz M. The impact of rare and low-frequency genetic variants in common variable immunodeficiency (CVID). *Sci Rep*. 2021;11:8308.
34. Jackson TJ, Chow G, Suri M, Byrd P, Taylor MR, Whitehouse WP. Longitudinal analysis of the neurological features of ataxia-telangiectasia. *Dev Med Child Neurol*. 2016;58:690-7.

35. Jamwal M, Aggarwal A, Palodhi A, Sharma P, Bansal D, Trehan A, et al. Next-Generation Sequencing–Based Diagnosis of Unexplained Inherited Hemolytic Anemias Reveals Wide Genetic and Phenotypic Heterogeneity. *J Mol Diagn.* 2020;22:579-90.
36. Namkoong H, Edahiro R, Takano T, Nishihara H, Shirai Y, Sonehara K, et al. DOCK2 is involved in the host genetics and biology of severe COVID-19. *Nature.* 2022;609:754-60.
37. Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, Chen J, et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science.* 2020;370:eabd4570.
38. Zhang SY, Jouanguy E, Ugolini S, Smahi A, Elain G, Romero P, et al. TLR3 deficiency in patients with herpes simplex encephalitis. *Science.* 2007;317:1522-7.
39. Bogershausen N, Gatinois V, Riehmer V, Kayserili H, Becker J, Thoernes M, et al. Mutation Update for Kabuki Syndrome Genes KMT2D and KDM6A and Further Delineation of X-Linked Kabuki Syndrome Subtype 2. *Hum Mutat.* 2016;37:847-64.
40. Feldmann J, Callebaut I, Raposo G, Certain S, Bacq D, Dumont C, et al. Munc13-4 is essential for cytolytic granules fusion and is mutated in a form of familial hemophagocytic lymphohistiocytosis (FHL3). *Cell.* 2003;115:461-73.
41. Bichon A, Bourenne J, Allardet-Servent J, Papazian L, Hraiech S, Guervilly C, et al. High Mortality of HLH in ICU Regardless Etiology or Treatment. *Front Med (Lausanne).* 2021;8:735796.
42. Reiner AP, Spivak JL. Hematophagic histiocytosis. A report of 23 new patients and a review of the literature. *Medicine (Baltimore).* 1988;67:369-88.
43. Janka G, Imashuku S, Elinder G, Schneider M, Henter JL. Infection- and malignancy-associated hemophagocytic syndromes. Secondary hemophagocytic lymphohistiocytosis. *Hematol Oncol Clin North Am.* 1998;12:435-44.
44. Wang LF, Anderson DE. Viruses in bats and potential spillover to animals and humans. *Curr Opin Virol.* 2019;34:79-89.
45. Galán J, Cuenca-Bescós G, López-García JM, Sauqué V, Núñez-Lahuerta C. Fossil bats from the Late Pleistocene site of the Aguilón P7 Cave (Zaragoza, Spain). *Comptes Rendus Palevol.* 2016;15:501-14.
46. Bandaru SS, Capace A, Busa V, Williams A. Secondary Hemophagocytic Lymphohistiocytosis in a Post-COVID-19 Patient. *Cureus.* 2022;14:e22620.
47. Weyrich LS, Duchene S, Soubrier J, Arriola L, Llamas B, Breen J, et al. Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus. *Nature.* 2017;544:357-61.
48. Wirth T, Hildebrand F, Allix-Beguec C, Wolbeling F, Kubica T, Kremer K, et al. Origin, spread and demography of the Mycobacterium tuberculosis complex. *PLoS Pathog.* 2008;4:e1000160.
49. Crozat K, Hoebe K, Ugolini S, Hong NA, Janssen E, Rutschmann S, et al. Jinx, an MCMV susceptibility phenotype caused by disruption of Unc13d: a mouse model of type 3 familial hemophagocytic lymphohistiocytosis. *J Exp Med.* 2007;204:853-63.
50. Sadat MA, Moir S, Chun T-W, Lusso P, Kaplan G, Wolfe L, et al. Glycosylation, Hypogammaglobulinemia, and Resistance to Viral Infections. *N Engl J Med.* 2014;370:1615-25.
51. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell.* 2020;181:271-80 e8.
52. Zeberg H, Paabo S. A genomic region associated with protection against severe COVID-19 is inherited from Neandertals. *Proc Natl Acad Sci U S A.* 2021;118:
53. Houldcroft CJ, Underdown SJ. Neanderthal genomics suggests a pleistocene time frame for the first epidemiologic transition. *Am J Phys Anthropol.* 2016;160:379-88.