Supplementary Materials

Methods

Ancient hominins and primatesgenomes reconstruction

Aligned bam-files (IceMan, Hunters, VindijaNea, AltaiNea, MezmaiskayaNea) or vcf-files (Denisovan/pinky, Chimpanzee) were obtained from thecorresponding open sources (ST2) and re-processed by using one of thefollowing pipelines – depending on type of the source file available. In the case if bam-files were available – a "consensus DNA sequence" was generated by using "SAMtools"and "seqtk" [1] software:

$samtoolsmpileup -ufgenome.fainput.bam | bcftools view -cg - | vcfutils.pl vcf2fq >consencus.fq
$ seqtkseq -l 50 -A consencus.fq>consencus.fa

In the case when only vcf-files were available, - the consensus DNA sequence was generated by using Genome Analysis Toolkit (GATK)software [2]:

$ java -Xmx2g –jar GenomeAnalysisTK.jar -lw 50 -R genome.fa -T FastaAlternateReferenceMaker
-L TARGET.intervals -o consencus.fa --variant input.vcf.gz

Having consensus DNA sequences and coordinates of regions of interest (ROI) around each SNPs position it wasstraightforward to extract ROI's nucleotide sequences and/or exons/introns nucleotide sequences corresponding to aparticular ancient DNA sample by using "fastahack" software [3] and custom Python scripts (available upon request).

Results

rs2218404. Neither WM nor PO and PS domain scores (N = 19 patients for the T allele carriers for WM and PS) were significantly attenuated comparing to G/G carriers (N = 34 patients for WM and PS; $p = 0.12$ (F = 2.50) and $p = 0.11$ (F = 2.59) for the WM and PIQ, respectively, and $p = 0.14$ (F = 2.19) for the PS) (Figure 1B-1). None of the healthy human subjects IQ scores were affected by the presence of T allele (N = 62 for T and N = 81 for G/G) and were significantly higher than patients' scores (see ST1 for the diagnosis effect statistics and interaction).

rs96501. No substantial difference in the IQ scores has been observed among the patients either C (N = 20) or T/T (N = 41) allele carriers despite a decline across the all scores comparing to healthy subjects (ST1).

rs1105684. Within SCZ patients cohort a cognitive endophenotype was not detectable (N = 33 and N = 67 for the A and T/T alleles carriers, respectively) with very close scores for both genotyping groups but substantially lower than for the healthy subjects (ST1).

rs2149171. WM contributing scores of letter number sequencing (LNS), digit span (DS) and arithmetic (AM) did not differ significantly among the genotypes (Figure 1B-2, lower panel) indicative of rather a distributive nature of the mutation allele effect on the final WM score. None of the healthy human subjects IQ scores were affected by the C allele presence and were repeatedly higher than for the SCZ patients' scores at the diagnosis level (ST1).

Supplementary table 1 (ST1):

HP118														
	WAIS-III	Schizophrenia (N = 61)				Control (N = 143)				ANCOVA (sex, education) *p* values (*F* values)				
	Variables	G/G		T carrier		G/G		T carrier		Diagnosis	Genotype	Interaction		
	NETRIN G1; NTNG1	(N = 36)		(N = 25)		(N = 81)		(N = 62)		effect	effect			
	rs2218404	Mean	SD	Mean	SD	Mean	SD	Mean	SD				SZ	HC
	Verbal IQ	96.1	14.3	83.6	16.3	108.9	13.2	108.6	13.8	**5.18×10^{-13} (59.81)**	**0.0033 (8.87)**	**0.0058 (7.78)**	**0.0059 (8.19)**	0.81 (0.06)
	Performance IQ	84.5	16.5	77.0	13.5	109.0	10.3	109.5	12.5	**1.05×10^{-29} (180.76)**	0.11 (2.59)	0.061 (3.54)	0.089 (3.00)	0.77 (0.08)
	Full-scale IQ	90.0	15.9	78.9	14.9	109.9	11.6	109.5	12.5	**1.05×10^{-23} (131.74)**	**0.0057 (7.80)**	**0.012 (6.48)**	**0.017 (6.01)**	0.76 (0.09)
										-				
	Verbal Comprehension	97.5	13.2	84.4	16.6	106.6	13.2	107.1	12.9	**1.30×10^{-9} (40.58)**	**0.0050 (8.08)**	**0.0018 (9.99)**	**0.0037 (9.14)**	0.79 (0.07)
	Perceptual Organization	86.6	16.8	81.5	14.2	108.0	10.0	107.7	14.6	**2.57×10^{-21} (114.05)**	0.19 (1.74)	0.28 (1.19)	0.22 (1.52)	0.82 (0.05)
	Working Memory[a]	91.1	13.8	84.2	13.5	106.3	15.9	105.4	16.1	**7.14×10^{-10} (42.17)**	0.12 (2.50)	0.25 (1.33)	0.16 (2.06)	0.64 (0.21)
	Processing Speed[a]	82.9	14.9	75.1	12.9	110.0	12.9	108.8	13.3	**1.02×10^{-26} (157.61)**	0.14 (2.19)	0.23 (1.44)	0.24 (1.41)	0.71 (0.14)
Verbal Comprehension	Vocabulary	9.7	2.8	7.1	3.2	11.4	3.0	11.7	3.1	**1.69×10^{-7} (29.42)**	**0.020 (5.49)**	**0.0041 (8.44)**	**0.0053 (8.39)**	0.65 (0.20)
Verbal Comprehension	Similarities	9.6	2.9	7.6	3.3	11.2	2.3	11.2	2.4	**4.75×10^{-7} (27.13)**	**0.041 (4.23)**	**0.022 (5.35)**	**0.044 (4.26)**	0.79 (0.07)
Verbal Comprehension	Information	9.3	2.6	7.1	3.0	10.9	3.0	10.9	2.4	**1.57×10^{-7} (29.58)**	**0.0067 (7.50)**	**0.014 (6.09)**	**0.0047 (8.39)**	0.78 (0.08)
		[a]N=34		[a]N=19										

rs2218404

rs96501

HP122												
	WAIS-III	Schizophrenia (N = 61)				Control (N = 143)				ANCOVA (sex, education) *p* values (*F* values)		
	Variables	C carrier		T/T		C carrier		T/T		Diagnosis	Genotype	Interaction
	NETRIN G1; NTNG1	(N = 20)		(N = 41)		(N = 45)		(N = 98)		effect	effect	
	rs96501	Mean	SD	Mean	SD	Mean	SD	Mean	SD			
	Verbal IQ	90.3	12.6	91.3	17.9	109.7	12.7	108.3	13.8	**3.97×10^{-11} (48.95)**	0.58 (0.31)	0.33 (0.95)
	Performance IQ	79.7	13.5	82.2	16.7	108.5	10.1	109.6	11.8	**3.36×10^{-27} (159.45)**	0.53 (0.40)	0.57 (0.33)
	Full-scale IQ	84.1	13.4	86.1	17.7	110.1	10.7	109.5	12.5	**4.87×10^{-21} (112.06)**	0.94 (0.01)	0.33 (0.96)
										-		.
	Verbal Comprehension	89.0	12.5	93.6	17.3	107.6	12.4	106.5	13.3	**1.15×10^{-8} (35.50)**	0.72 (0.13)	0.07 (3.31)
	Perceptual Organization	83.2	14.3	85.1	16.7	108.5	12.1	107.6	12.3	**3.95×10^{-20} (105.62)**	>0.99 (<0.01)	0.39 (0.73)
	Working Memory[a]	91.4	11.3	87.0	15.2	107.9	15.3	105.0	16.2	**7.39×10^{-9} (36.65)**	0.084 (3.01)	0.87 (0.03)
	Processing Speed[a]	77.9	10.5	81.3	16.5	104.9	11.4	111.6	13.2	**3.65×10^{-25} (144.85)**	**0.028 (4.89)**	0.52 (0.42)
Processing Speed	Digit Symbol	6.2	2.4	6.3	3.0	11.0	2.6	12.3	2.6	**2.16×10^{-24} (137.00)**	0.12 (2.40)	0.23 (1.48)
Processing Speed	Symbol Search[a]	5.9	2.0	6.7	3.5	10.9	2.7	12.1	2.8	**2.98×10^{-19} (100.48)**	**0.053 (3.79)**	0.71 (0.14)
		[a]N=19		[a]N=34								

PPI	Schizophrenia (N = 54)							Control (N = 105)							ANCOVA (age, sex) *p* values (*F* values)		
Variables	C/C		C/T		T/T			C/C		C/T		T/T			Diagnosis	Genotype	Interaction
	(N = 4)		(N = 17)		(N = 33)			(N = 3)		(N = 30)		(N = 72)			effect	effect	
	Mean	SD	Mean	SD	Mean	SD		Mean	SD	Mean	SD	Mean	SD				
SR	52.2	24.4	160.5	148.3	128.4	87.4		102.8	51.1	154.1	119.3	177.2	124.1		0.59 (0.29)	0.50 (0.71)	0.42 (0.88)
HAB	32.8	21.1	42.7	20.3	43.4	20.2		44.6	24.1	53.6	19.7	53.0	19.2		0.07 (3.37)	0.58 (0.55)	0.98 (0.02)
PPI90	-5.7	18.4	37.7	25.2	41.3	32.2		36.9	31.8	45.2	24.1	43.0	44.0		0.08 (3.04)	0.15 (1.90)	0.35 (1.05)
PPI86	-14.0	19.3	28.6	49.5	37.9	27.7		16.7	25.6	37.9	29.0	33.8	43.3		0.23 (1.42)	0.054 (2.97)	0.35 (1.05)
PPI82	-30.1	41.0	31.3	27.6	37.8	23.5		12.9	24.1	32.7	27.8	29.4	35.5		0.11 (2.56)	**0.0010 (7.23)**	0.065 (2.78)

HP124	NETRIN G2; NTNG2													
rs1105684	rs1105684													
	WAIS	Schizophrenia (N = 100)				Control (N = 145)				ANCOVA (sex, education) *p* values (*F* values)				
	Variables	A carriers		T/T		A carriers		T/T		Diagnosis	Genotype	Interaction		
		(N = 33)		(N = 67)		(N = 49)		(N = 96)		effect	effect			
		Mean	SD	Mean	SD	Mean	SD	Mean	SD				SZ	HC
	Verbal IQ	91.5	18.9	92.4	14.9	106.3	13.2	110.5	13.5	**2.59×10^{-11} (48.98)**	0.10 (2.72)	0.32 (1.01)	0.82 (0.05)	**0.029 (4.90)**
	Performance IQ	84.4	18.7	82.6	15.4	107.2	12.1	110.6	10.8	**3.03×10^{-27} (151.16)**	0.58 (0.31)	0.11 (2.51)	0.52 (0.41)	**0.048 (3.99)**
	Full-scale IQ	87.2	19.8	86.7	15.5	107.3	12.5	111.4	11.6	**6.24×10^{-21} (106.78)**	0.23 (1.45)	0.17 (1.92)	0.81 (0.06)	**0.018 (5.70)**

rs1105684

rs2149171

HP127	NETRIN G2; NTNG2																
rs2149171																	
	WAIS-III	Schizophrenia (N = 58)						Control (N = 142)						ANCOVA (sex, education) *p* values (*F* values)				
	Variables	C/C		C/T		T/T		C/C		C/T		T/T		Diagnosis	Genotype	Interaction		
		(N = 14)		(N = 29)		(N = 15)		(N = 39)		(N = 73)		(N = 30)		effect	effect			
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD				SZ	HC
	Verbal IQ	88.3	18.9	88.2	14.2	99.4	15.7	109.8	11.4	107.8	15.1	109.5	11.9	**1.88×10^{-9} (39.82)**	**0.029 (3.60)**	0.083 (2.52)		
	Performance IQ	77.9	17.8	79.1	13.9	88.9	16.8	109.1	11.1	108.9	11.1	110.5	12.2	**1.57×10^{-25} (147.32)**	**0.035 (3.42)**	0.15 (1.89)		-
	Full-scale IQ	82.3	18.7	82.6	14.3	94.1	17.2	110.4	10.7	108.9	12.8	110.8	11.8	**5.72×10^{-19} (98.28)**	**0.014 (4.35)**	0.066 (2.76)		-
														-				
	Verbal Comprehension	90.1	18.0	90.7	15.2	97.4	15.0	107.0	10.6	106.5	14.6	107.2	12.2	**1.62×10^{-6} (24.50)**	0.27 (1.31)	0.39 (0.95)		
	Perceptual Organization	81.4	17.5	82.0	14.9	92.2	16.1	107.6	12.1	107.9	11.9	108.5	13.4	**4.71×10^{-18} (92.03)**	0.050 (3.04)	0.11 (2.20)		-
	Working Memory[a]	83.9	15.0	86.5	11.0	97.0	16.0	107.3	14.9	104.5	16.8	107.7	15.2	**1.21×10^{-8} (35.60)**	**0.040 (3.27)**	0.16 (1.85)		-

	Processing Speed[a]	76.3	18.2	79.3	14.7	84.3	11.4	106.6	12.1	110.6	12.5	110.7	15.0	**2.67×10^{-24} (139.08)**	0.19 (1.69)	0.71 (0.34)		
	Comprehension	6.6	3.9	7.7	3.3	11.0	3.5	12.5	2.9	12.1	3.2	11.7	3.1	**1.15×10^{-8} (35.59)**	**0.012 (4.54)**	**3.56×10^{-4} (8.28)**	**0.0030 (6.48)**	0.61 (0.49)
Working Memory	Arithmetic	8.6	3.7	7.4	2.7	9.1	3.4	11.8	2.8	11.3	3.5	11.8	3.1	**2.83×10^{-7} (28.34)**	0.092 (2.42)	0.58 (0.54)		
Working Memory	Digit Span	8.9	2.2	9.0	2.9	10.7	3.2	11.1	3.6	10.6	3.0	11.5	3.6	**0.0043 (8.34)**	0.066 (2.75)	0.61 (0.50)		-
Working Memory	Letter Number Sequencing[a]	5.9	2.5	7.1	2.8	8.5	3.6	11.1	3.6	10.7	3.4	10.8	2.7	**1.34×10^{-8} (35.37)**	0.27 (1.32)	0.18 (1.73)		-
		N=12		N=25		N=14												

rs2274855

HP128	NETRIN G2; NTNG2													
rs2274855														
	WAIS-III	Schizophrenia (N = 59)				Control (N = 143)				ANCOVA (sex, education) *p* values (*F* values)				
	Variables	A carrier		G/G		A carrier		G/G		Diagnosis	Genotype	Interaction		
		(N = 33)		(N = 26)		(N = 91)		(N = 52)		effect	effect			
		Mean	SD	Mean	SD	Mean	SD	Mean	SD				SZ	HC
	Verbal IQ	86.1	16.2	96.5	15.0	108.8	14.2	108.6	12.2	**9.68×10^{-11} (46.83)**	**0.017 (5.83)**	**0.018 (5.70)**	**0.020 (5.70)**	0.98 (<0.01)
	Performance IQ	78.3	15.8	86.0	15.1	108.9	11.1	109.8	11.6	**3.77×10^{-27} (159.57)**	**0.036 (4.46)**	0.11 (2.58)		-
	Full-scale IQ	81.1	16.2	90.9	15.6	109.6	12.2	109.9	11.6	**1.03×10^{-20} (110.02)**	**0.012 (6.44)**	**0.025 (5.07)**	**0.033 (4.80)**	0.77 (0.09)
										-				
	Verbal Comprehension	88.8	17.0	95.5	14.1	106.8	13.4	106.7	12.5	**1.37×10^{-7} (29.91)**	0.14 (2.23)	0.16 (1.98)		
	Perceptual Organization	80.9	15.9	90.0	14.7	107.5	11.8	108.6	12.8	**1.49×10^{-19} (101.86)**	**0.012 (6.36)**	0.056 (3.68)		-
	Working Memory[a]	84.0	11.8	95.0	14.3	106.0	16.7	105.8	14.6	**1.30×10^{-8} (35.36)**	**0.023 (5.28)**	**0.023 (5.23)**	**0.0031 (9.73)**	0.99 (<0.01)
	Processing Speed[a]	77.0	16.9	84.0	10.6	109.2	11.8	109.9	15.0	**1.01×10^{-25} (149.62)**	0.12 (2.45)	0.20 (1.64)		
	Comprehension	13.4	5.2	17.4	5.0	21.0	4.4	19.9	4.1	**1.99×10^{-9} (39.61)**	**0.035 (4.49)**	**2.27×10^{-4} (14.12)**	**0.0063 (8.06)**	0.12 (2.44)
Working Memory	Arithmetic	11.9	3.7	13.7	4.2	17.2	4.6	17.5	4.3	**2.06×10^{-8} (34.19)**	0.067 (3.39)	0.27 (1.24)		
Working Memory	Digit Span	16.1	3.2	18.2	4.2	19.0	4.3	19.3	4.5	**0.0040 (8.47)**	**0.026 (5.04)**	0.10 (2.70)		-
Working Memory	Letter Number Sequencing[a]	10.0	3.1	11.4	2.4	13.9	2.8	13.3	2.3	**4.77×10^{-8} (32.39)**	0.27 (1.22)	**0.031 (4.73)**	0.053 (3.95)	0.32 (0.98)
										-				-
Perceptual Organization	Picture Completion	7.2	3.0	7.7	3.2	11.1	2.4	11.3	2.9	**3.98×10^{-15} (72.74)**	0.46 (0.56)	0.77 (0.09)		-
Perceptual Organization	Block Design	7.0	3.9	9.1	3.4	11.1	2.7	11.3	2.9	**2.07×10^{-8} (34.18)**	**0.010 (6.82)**	**0.041 (4.23)**	**0.028 (5.08)**	0.57 (0.32)
Perceptual Organization	Matrix Reasoning	6.8	3.0	8.7	3.2	11.6	2.7	11.6	2.4	**2.11×10^{-14} (68.26)**	**0.024 (5.19)**	**0.038 (4.34)**	**0.038 (4.54)**	0.84 (0.04)
		N=29		N=23										

Dataset Name	Link to the dataset	File name(s) used
Saqqaq-Eskimos [13]	http://www.binf.ku.dk/saqqaq/	chr1.diff.annotation chr1.same.annotation.gz chr9.diff.annotation chr9.same.annotation.gz chr1.diff.highconfidence.PP chr1.same.highconfidence.PP chr9.diff.highconfidence.PP chr9.same.highconfidence.PP
IceMan[14]*	ftp://ftp.sra.ebi.ac.uk/vol1/ERA081/ERA081021/bam/	ellen1.bam
	ftp://ftp.sra.ebi.ac.uk/vol1/ERA081/ERA081149/bam/	ellen2.bam
	ftp://ftp.sra.ebi.ac.uk/vol1/ERA081/ERA081149/bam/	teresa.bam
Hunters [15]**	http://www.ebi.ac.uk/ena/data/view/PRJEB6272	Loschbour.hg19_1000g.bam Motala1.bam Motala3.bam Motala12.bam
VindijaNea[16]*	ftp://hgdownload.cse.ucsc.edu/gbdb/hg18/neandertal/seqAlis/	all-hg18.bam
Denisovan/pinky[17]**	http://cdna.eva.mpg.de/neandertal/altai/Denisovan/	DenisovaPinky.hg19_1000g.1.mod.vcf.gz DenisovaPinky.hg19_1000g.9.mod.vcf.gz
AltaiNea[17]**	http://cdna.eva.mpg.de/neandertal/altai/AltaiNeandertal/bam/	AltaiNea.hg19_1000g.1.dq.bam AltaiNea.hg19_1000g.9.dq.bam
MezmaiskayaNea[17]**	http://cdna.eva.mpg.de/neandertal/Mezmaiskaya/bam/	E733.bam
Chimpanzee [18]*	http://www.biologiaevolutiva.org/greatape/data.html https://eichlerlab.gs.washington.edu/greatape/data/VCFs/SNPs/	Pan_troglodytes.vcf.gz

Supplementary Table 2 (ST2)

*aligned against hg18: http://hgdownload.cse.ucsc.edu/goldenPath/hg18/bigZips/

**aligned against hg19 (iGenome by Illumina): http://support.illumina.com/sequencing/sequencing_software/igenome.html

rs2274855. Similarly to the rs2149171 C-allele-produced phenotype, the attenuated PIQ was due to lower PO scores ($p = 0.012$ ($F = 6.36$)) while the PS stays unaffected by the A-allele presence, contrary to the rs2218404 endophenotype. None of the healthy subject's scores were affected by the A-allele presence but meanwhile diagnostically being higher than that of SCZ patients (ST1).

None of the other 3 ROIs of *NTNG1* (rs7851893, rs3824574 and rs2149171) display any significantly different evolutionary rate changes (Figure 3C) possibly due to the highly conservative nature of the embedding them DNA (Figure 3B). The ROI of rs2218404 is 100% conserved across all hominins (no other mutations are detectable in the ±50 nu regions as well, SM) but contains two point mutations in chimpansee, 6 in marmoset (due to «GCC» duplication) and 9 in mouse (Figure 2D, SM: rs2218404). Evolution of the other 3 ROIs non-affecting IQ SNPs (rs1373336, rs1444042 and rs4915045) shows much in common such as major allele fixation in primates (C, A and A, respectively), stronger conservation of the primates' ROIs and the following 100% conservation of ROIs across the hominin species. There is 1 mutation for chimpanzee and 3 for marmoset in rs1373336 ROI plus an extra marmoset-related mutation in rs1444042 ROI. Other primates ROIs are 100% identical to the human ones. Next on the gene is a non-affecting IQ ROI of rs628117 also characterised by 100% conservation in hominins and has only 1 mutation in chimpanzee comparing to human but 9 mutations in the marmoset ROI with 7 of them being asymmetrically nested in the preceding the allele 5'-sequence and resulting in the characteristic inflection point of brake on the percent identity plot on Figure 2B. It is also interesting to note the C/T alleles intermingled transition across the all analysed species. rs96501 ROI is 100% conserved starting from chimpanzee but dramatically different in the following the allele's positioning 3'-sequence shoulder of marmoset (contains 6 mutations). The T-allele of rs1105684 is strongly conserved across all primates and hominins however its ROI undergoes dramatic changes with 8 and 15 out of 20 nu mutated in marmoset and mice, respectively. Mice rs7851893 ROI is poorly conserved (5 mutations) and displays the third highest value for the evolutionary advancements rate (Figure 3C) from mice to marmoset after rs1105684 and rs2274855 (Figure 3C). Mice rs3824574 ROI ccontains 1 mutation. The T-allele of rs3824574 is not detectable in any of the analysed species but its analogue in rs7851893 is gradually substituted on G-allele starting from MezmayskayaNea and Mesolithic hominins. -50 nu DNA stretch preceeding the rs3824574 allele position contains a MezmayskayaNea-specific substitution however 30 nu away (Figure 3B, right panel). A Denisovan-specific mutation is found 46 nu downstream of the rs2149171 mutation allele position within the ROI. The ROI of rs2274855 is strongly conserved in primates with 1 mutation per chimpanzee and marmoset each but containing significant evolutionary changes in mice ROI (8 mutations). This is the another ROI (alongside with the rs2218404 ROI affecting the VC, Figure 2C) undergoing an AE from chimapanzee to human.

There are 5 point mutations in Eskimos Netrin-G1 relative to modern human reference genome (N363S, V385G, D405G, E461G, P506R) located in the middle and end of the protein with no other aa substitutions found across all other analysedhominins (due to poor genome coverage the protein information for both genes is not always complete, see SM). One mutation found for chimpanzee (A81S) is absent in all other analysed species; marmoset contains 4 mutations (L32V, D553G, P564R, A565T), and mouse – 20 mutations relative to the human ortholog (P24L, L25F, L32V, T35S, Q36L, M48T, T78S, D184H, T215S, V305S, V307L, N363T, H501Y, H523Q, L529A, A532D, D553E, H561R, T571M, S577G). Two mutations are shared among the mice and marmoset proteins (L32V and D553E/G) and one mice mutation is also found in the reconstructed Eskimos protein (N363T/S). As for Netrin-G2, there are 5 mutations in the Eskimos protein (T346A, K352R, A375G, D409G, E465G, P510R). Loschbour, Motala12, and AltaiNea each contain only one mutation - T346A. Motala3 – 2 mutations (S200L, R201C), VindijaNea – 3 mutations (H10Y, T346A, L573I), Denisovan – 2 mutations (R181Q, T346A),

Mezmayskaya Nea – at least 1 mutation (T311I), chimpanzee – 2 mutations (T346A, S371A), marmoset – 12 mutations (A104V, M298V, A345S, T346A, S364G, A370V, S371V, A381T, Y389S, Q521L, R568H, R587H), and mouse – 49 mutations (H3R, A15V, K98R, E100D, I110V, T130S, M139V, S187P, L237M, L249F, S297T, M298V, K318R, S327A, T346A, K352S, W353Q, R355K, S357P, A359M, R365F, S367P, A370S, R372S, G377A, T378I, P379S, A380V, A382V, A384S, P385Q, G388D, Y389S, K390T, Q393E, K397R, M402I, A491V, L527Q, R536P, V541I, P550L, G555A, L557P, A565I, A566V, A570D, G586A, R587C). The mice protein is also 2 aa longer than human Netrin-G2 due to extra 3 aa present in {LE1(ex5)-Ukd} area: X350A; X364G; X365D (Figure 3F, as per mice coordinates) and one aa preceding the GPI-link (D554X) is absent (see SM: Netrin-G2). Similarly to Netrin-G1, Eskimos and mouse have the same aa residue affected – K352R/S (see SM). Marmoset and mouse share four mutations not found in other species (M298V, A370V/S, Y389S, R587H/C), and the aforementioned T346A shared with other hosts. Interestingly that despite truly belonging anatomically to primates (as a "New World Monkey") marmoset shares higher number of coding (Figure 3F) and non-coding (Figure 3D) mutations with mouse rather than with the "Old World Monkey" – chimpanzee, though corroborating the reported conservation of the cortical area genetic markers with mice [4]. It is also noted that if the mice-specific Netrin-G1 point mutations are predominantly located at the N- and C-termini of the protein (domains LE1 and GPI-link, respectively), 26 out of total 52 Netrin-G2 point mutations are middle part nested (exon5 + Ukd domain), another 12 are at the C-termini and the rest is spread across the other areas (SM).

Discussion

The T-allele of rs1105684 is strongly conserved across all primates and hominins but not the ROI area which in turn undergoes dramatic changes with 8 and 15 out of 20 nu being mutated in marmoset and mice, respectively, demonstrating some potential evolutionary siginifcant events taking place and possibly linked to the T-allele emergence. The mutation alleles of rs11056893 (affecting VIQ and PIQ) and rs7851893 (non-affecting IQ) occupy strategically important locations at the beginning of the gene (intron (1-2) preceeds the coding exon 2, Figure 3A) but are separated from each other by 428 nu. Displaying distinct cognitive endophenotypes functionality (while both being associated with SCZ) they represent an example of a strong positive selection of an intron loci as opposite to a neutral drift [5] of the non-coding area. The fact that rs1105684 A-allele substitution of a strongly conserved T allele (>40 mlnyrs as per marmoset) affects healthy individuals may be indicative of its devastating impact on *NTNG2* function which does not require an interaction with the diagnosis of SCZfor an effect to take a place (a variation of a TATA-box embedding the allele is envisaged). In general, two distinct patterns emerge when the ROIs embedding all 11 analysed *NTNG* SNPs are compared. 1 pattern: a ROI either undergoes a sequence identity relaxation when moving away down from human across the evolutionary distant species (rs2218404, rs1373336, rs628117, rs96501, rs1105684), or opposite, 2 pattern: a ROI represents a high level of conservation (>95%) starting from marmoset (rs1444042, rs4915045, rs7851893, rs3824574, rs2149171, and rs2274855) and sometimes even including the mice ROI (rs3824574, rs2149171). Two ROIs belonging to group 1 pattern, rs2218404 and rs1105684, share substantial visual geometrical similarity represented as a funel-like plots where the low percent of sequence identity preceding and following the mutation allele (and embedded predominantly into the defined ROIs) is gradually removed upon moving both sides away from the allele position. Two other ROIs from this group, rs628117 and rs96501, display unique among others asymmetrically dramatic changes in the identity affecting either of the area sides (Figure 3B, low left and middle). Notably that rs628117 and rs96501 are both located in the centre of the identity plot inclination points possibly demarcating the boundaries of the following and preceding,

respectively, non-coding conserved islands of the intron (9-10) (Figure 3A). Significance of such intron-specific location manifests itself in the PS score affected by the rs96501 (Figure 1B-1).Group 2 pattern can be further subdivided as: a) with the percent identity becoming stronger when moving away from the ROI (rs1444042); b) weaker (rs7851893 and rs2149171); and c) mixed (rs4915045, rs3824574, rs2274855). Interestingly, that 3 out of the group 1 pattern ROIs (rs2218404, rs96501, rs1105684) affect IQ scores in human with two of them being *NTNG1* borne (rs2218404, rs96501) and one - *NTNG2* (rs1105684). In the second group there are two out of six ROIs (rs2149171 and rs2274855) affecting IQ and both are *NTNG2*-nested (Figure 3A).

With all cares being taken to represent the extinct *NTNGs* composition as accurately as possible some potential technical faults cannot be excluded. C to T and G to A are known drastically elevated transitions in the ancient DNA strands breaks [6]. And C deamination at 5'-overhangs leads to greater C/T misincorporation rates and converts into a complementary increase in G/A (3), e.g.5′-TCC-3′→5′-TTC-3′ (potentially affecting rs96501), which is a common somatic mutation in melanoma skin cancers [7].

Eskimos is described as having had brown eyes, A+ blood type, not a light skin but having had dark and thick hair [8]. Based on the prevailing number of SNPs (coinciding with our findings for *NTNG2* protein, Figure 3F) its genome does not show affinity to the modern-day Europeans but rather to the modern populations of northeast Siberia [9]. As for Neanderthals, some of them were probably red-haired and pale skinned, were able to taste bitter, have had a blood type *O* [10] and even were able to "enunciate as we do" [11] interpreting [12].

Supplementary References

1. Heng Li. "seqtk" - tool for processing sequences in the FASTA or FASTQ format. [https://github.com/lh3/seqtk]
2. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR,et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet43:491-498. [Crossref]
3. Erik Garrison. "fastahack" - FASTA file indexing and sequence extraction tool. [https://github.com/ekg/fastahack]
4. Mashiko H, Yoshida AC, Kikuchi SS, Niimi K, Takahashi E, et al. (2012) Comparative anatomy of marmoset and mouse cortex from genomic expression. *J Neurosci* 32: 5039-5053. [Crossref]
5. Noirel J, Simonson T (2008) Neutral evolution of proteins: The superfunnel in sequence space and its relation to mutational robustness. *J Chem Phys* 129: 185104. [Crossref]
6. Briggs AW, Stenzel U, Johnson PL, Green RE, Kelso J, et al. (2007) Patterns of damage in genomic DNA sequences from a Neandertal. *Proc Natl Acad Sci U S A*104: 14616-14621. [Crossref]
7. Harris K (2015) Evidence for recent, population-specific evolution of the human mutation rate. *Proc Natl Acad Sci U S A*112: 3439-3444. [Crossref]
8. Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, et al. (2010) Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature*463: 757-762. [Crossref]
9. Der Sarkissian C, Allentoft ME, Ávila-Arcos MC, Barnett R, Campos PF, et al. (2015) Ancient genomics. *Philos Trans R Soc Lond B Biol Sci* 370: 20130387. [Crossref]
10. Sanchez-Quinto F, Lalueza-Fox C (2015) Almost 20 years of Neanderthal palaeogenetics: adaptation, admixture, diversity, demography and extinction. *Philos Trans R Soc Lond B BiolSci* 370: 20130374.
11. Pollard KS (2009) What makes us human? *Sci Am* 300: 44-49. [Crossref]

12. Krause J, Lalueza-Fox C, Orlando L, Enard W, Green RE, et al. (2007) The derived FOXP2 variant of modern humans was shared with Neandertals. *CurrBiol*17: 1908-1912. [Crossref]
13. Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, et al. (2010) Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature*463: 757-762. [Crossref]
14. Keller A, Graefen A, Ball M, Matzas M, Boisguerin V, et al. (2012) New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nat Commun*3: 698. [Crossref]
15. Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, et al. (2014) Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*513: 409-413. [Crossref]
16. Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, et al. (2010) A draft sequence of the Neandertal genome. *Science*328: 710-722. [Crossref]
17. Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, et al.(2013) The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*505:43-49. [Crossref]
18. Prado-Martinez J, Sudmant PH, Kidd JM, Li H, Kelley JL, et al. (2013) Great ape genetic diversity and population history. *Nature*499: 471-475. [Crossref]

Results

rs2218404. Neither WM nor PO and PS domain scores (N = 19 patients for the T allele carriers for WM and PS) were significantly attenuated comparing to G/G carriers (N = 34 patients for WM and PS; $p = 0.12$ ($F = 2.50$) and $p = 0.11$ ($F = 2.59$) for the WM and PIQ, respectively, and $p = 0.14$ ($F = 2.19$) for the PS) (Figure 1B-1). None of the healthy human subjects IQ scores were affected by the presence of T allele (N = 62 for T and N = 81 for G/G) and were significantly higher than patients' scores (see ST1 for the diagnosis effect statistics and interaction).

rs96501. No substantial difference in the IQ scores has been observed among the patients either C (N = 20) or T/T (N = 41) allele carriers despite a decline across the all scores comparing to healthy subjects (ST1).

rs1105684. Within SCZ patients cohort a cognitive endophenotype was not detectable (N = 33 and N = 67 for the A and T/T alleles carriers, respectively) with very close scores for both genotyping groups but substantially lower than for the healthy subjects (ST1).

rs2149171. WM contributing scores of letter number sequencing (LNS), digit span (DS) and arithmetic (AM) did not differ significantly among the genotypes (Figure 1B-2, lower panel) indicative of rather a distributive nature of the mutation allele effect on the final WM score. None of the healthy human subjects IQ scores were affected by the C allele presence and were repeatedly higher than for the SCZ patients' scores at the diagnosis level (ST1).

rs2274855. Similarly to the rs2149171 C-allele-produced phenotype, the attenuated PIQ was due to lower PO scores ($p = 0.012$ ($F = 6.36$)) while the PS stays unaffected by the A-allele presence, contrary to the rs2218404 endophenotype. None of the healthy subject's scores were affected by the A-allele presence but meanwhile diagnostically being higher than that of SCZ patients (ST1).

None of the other 3 ROIs of *NTNG1* (rs7851893, rs3824574 and rs2149171) display any significantly different evolutionary rate changes (Figure 3C) possibly due to the highly conservative nature of the embedding them DNA (Figure 3B). The ROI of rs2218404 is 100% conserved across all hominins (no other mutations are detectable in the ±50 nu regions as well, SM) but

contains two point mutations in chimpansee, 6 in marmoset (due to «GCC» duplication) and 9 in mouse (Figure 2D, SM: rs2218404). Evolution of the other 3 ROIs non-affecting IQ SNPs (rs1373336, rs1444042 and rs4915045) shows much in common such as major allele fixation in primates (C, A and A, respectively), stronger conservation of the primates' ROIs and the following 100% conservation of ROIs across the hominin species. There is 1 mutation for chimpanzee and 3 for marmoset in rs1373336 ROI plus an extra marmoset-related mutation in rs1444042 ROI. Other primates ROIs are 100% identical to the human ones. Next on the gene is a non-affecting IQ ROI of rs628117 also characterised by 100% conservation in hominins and has only 1 mutation in chimpanzee comparing to human but 9 mutations in the marmoset ROI with 7 of them being asymmetrically nested in the preceding the allele 5'-sequence and resulting in the characteristic inflection point of brake on the percent identity plot on Figure 2B. It is also interesting to note the C/T alleles intermingled transition across the all analysed species. rs96501 ROI is 100% conserved starting from chimpanzee but dramatically different in the following the allele's positioning 3'-sequence shoulder of marmoset (contains 6 mutations). The T-allele of rs1105684 is strongly conserved across all primates and hominins however its ROI undergoes dramatic changes with 8 and 15 out of 20 nu mutated in marmoset and mice, respectively. Mice rs7851893 ROI is poorly conserved (5 mutations) and displays the third highest value for the evolutionary advancements rate (Figure 3C) from mice to marmoset after rs1105684 and rs2274855 (Figure 3C). Mice rs3824574 ROI ccontains 1 mutation. The T-allele of rs3824574 is not detectable in any of the analysed species but its analogue in rs7851893 is gradually substituted on G-allele starting from MezmayskayaNea and Mesolithic hominins. -50 nu DNA stretch preceeding the rs3824574 allele position contains a MezmayskayaNea-specific substitution however 30 nu away (Figure 3B, right panel). A Denisovan-specific mutation is found 46 nu downstream of the rs2149171 mutation allele position within the ROI. The ROI of rs2274855 is strongly conserved in primates with 1 mutation per chimpanzee and marmoset each but containing significant evolutionary changes in mice ROI (8 mutations). This is the another ROI (alongside with the rs2218404 ROI affecting the VC, Figure 2C) undergoing an AE from chimapanzee to human.

There are 5 point mutations in Eskimos Netrin-G1 relative to modern human reference genome (N363S, V385G, D405G, E461G, P506R) located in the middle and end of the protein with no other aa substitutions found across all other analysedhominins (due to poor genome coverage the protein information for both genes is not always complete, see SM). One mutation found for chimpanzee (A81S) is absent in all other analysed species; marmoset contains 4 mutations (L32V, D553G, P564R, A565T), and mouse – 20 mutations relative to the human ortholog (P24L, L25F, L32V, T35S, Q36L, M48T, T78S, D184H, T215S, V305S, V307L, N363T, H501Y, H523Q, L529A, A532D, D553E, H561R, T571M, S577G). Two mutations are shared among the mice and marmoset proteins (L32V and D553E/G) and one mice mutation is also found in the reconstructed Eskimos protein (N363T/S). As for Netrin-G2, there are 5 mutations in the Eskimos protein (T346A, K352R, A375G, D409G, E465G, P510R). Loschbour, Motala12, and AltaiNea each contain only one mutation - T346A. Motala3 – 2 mutations (S200L, R201C), VindijaNea – 3 mutations (H10Y, T346A, L573I), Denisovan – 2 mutations (R181Q, T346A), MezmayskayaNea – at least 1 mutation (T311I), chimpanzee – 2 mutations (T346A, S371A), marmoset – 12 mutations (A104V, M298V, A345S, T346A, S364G, A370V, S371V, A381T, Y389S, Q521L, R568H, R587H), and mouse – 49 mutations (H3R, A15V, K98R, E100D, I110V, T130S, M139V, S187P, L237M, L249F, S297T, M298V, K318R, S327A, T346A, K352S, W353Q, R355K, S357P, A359M, R365F, S367P, A370S, R372S, G377A, T378I, P379S, A380V, A382V, A384S, P385Q, G388D, Y389S, K390T, Q393E, K397R, M402I, A491V, L527Q, R536P, V541I, P550L, G555A, L557P, A565I, A566V, A570D, G586A, R587C). The mice protein is also 2 aa longer than human Netrin-G2 due to extra 3 aa present in {LE1(ex5)-Ukd} area: X350A; X364G; X365D (Figure 3F, as per mice coordinates) and one aa preceding the GPI-link

(D554X) is absent (see SM: Netrin-G2). Similarly to Netrin-G1, Eskimos and mouse have the same aa residue affected – K352R/S (see SM). Marmoset and mouse share four mutations not found in other species (M298V, A370V/S, Y389S, R587H/C), and the aforementioned T346A shared with other hosts. Interestingly that despite truly belonging anatomically to primates (as a "New World Monkey") marmoset shares higher number of coding (Figure 3F) and non-coding (Figure 3D) mutations with mouse rather than with the "Old World Monkey" – chimpanzee, though corroborating the reported conservation of the cortical area genetic markers with mice [4]. It is also noted that if the mice-specific Netrin-G1 point mutations are predominantly located at the N- and C-termini of the protein (domains LE1 and GPI-link, respectively), 26 out of total 52 Netrin-G2 point mutations are middle part nested (exon5 + Ukd domain), another 12 are at the C-termini and the rest is spread across the other areas (SM).

Discussion

The T-allele of rs1105684 is strongly conserved across all primates and hominins but not the ROI area which in turn undergoes dramatic changes with 8 and 15 out of 20 nu being mutated in marmoset and mice, respectively, demonstrating some potential evolutionary siginifcant events taking place and possibly linked to the T-allele emergence. The mutation alleles of rs11056893 (affecting VIQ and PIQ) and rs7851893 (non-affecting IQ) occupy strategically important locations at the beginning of the gene (intron (1-2) preceeds the coding exon 2, Figure 3A) but are separated from each other by 428 nu. Displaying distinct cognitive endophenotypes functionality (while both being associated with SCZ) they represent an example of a strong positive selection of an intron loci as opposite to a neutral drift [5] of the non-coding area. The fact that rs1105684 A-allele substitution of a strongly conserved T allele (>40 mlnyrs as per marmoset) affects healthy individuals may be indicative of its devastating impact on *NTNG2* function which does not require an interaction with the diagnosis of SCZfor an effect to take a place (a variation of a TATA-box embedding the allele is envisaged). In general, two distinct patterns emerge when the ROIs embedding all 11 analysed *NTNG* SNPs are compared. 1 pattern: a ROI either undergoes a sequence identity relaxation when moving away down from human across the evolutionary distant species (rs2218404, rs1373336, rs628117, rs96501, rs1105684), or opposite, 2 pattern: a ROI represents a high level of conservation (>95%) starting from marmoset (rs1444042, rs4915045, rs7851893, rs3824574, rs2149171, and rs2274855) and sometimes even including the mice ROI (rs3824574, rs2149171). Two ROIs belonging to group 1 pattern, rs2218404 and rs1105684, share substantial visual geometrical similarity represented as a funel-like plots where the low percent of sequence identity preceding and following the mutation allele (and embedded predominantly into the defined ROIs) is gradually removed upon moving both sides away from the allele position. Two other ROIs from this group, rs628117 and rs96501, display unique among others asymmetrically dramatic changes in the identity affecting either of the area sides (Figure 3B, low left and middle). Notably that rs628117 and rs96501 are both located in the centre of the identity plot inclination points possibly demarcating the boundaries of the following and preceding, respectively, non-coding conserved islands of the intron (9-10) (Figure 3A). Significance of such intron-specific location manifests itself in the PS score affected by the rs96501 (Figure 1B-1).Group 2 pattern can be further subdivided as: a) with the percent identity becoming stronger when moving away from the ROI (rs1444042); b) weaker (rs7851893 and rs2149171); and c) mixed (rs4915045, rs3824574, rs2274855). Interestingly, that 3 out of the group 1 pattern ROIs (rs2218404, rs96501, rs1105684) affect IQ scores in human with two of them being *NTNG1* borne (rs2218404, rs96501) and one - *NTNG2* (rs1105684). In the second group there are two out of six ROIs (rs2149171 and rs2274855) affecting IQ and both are *NTNG2*-nested (Figure 3A).

With all cares being taken to represent the extinct *NTNGs* composition as accurately as possible some potential technical faults cannot be excluded. C to T and G to A are known drastically elevated transitions in the ancient DNA strands breaks [6]. And C deamination at 5'-overhangs leads to greater C/T misincorporation rates and converts into a complementary increase in G/A (3), e.g.5′-TCC-3′→5′-TTC-3′ (potentially affecting rs96501), which is a common somatic mutation in melanoma skin cancers [7].

Eskimos is described as having had brown eyes, A+ blood type, not a light skin but having had dark and thick hair[8]. Based on the prevailing number of SNPs (coinciding with our findings for *NTNG2* protein, Figure 3F) its genome does not show affinity to the modern-day Europeans but rather to the modern populations of northeast Siberia [9]. As for Neanderthals, some of them were probably red-haired and pale skinned, were able to taste bitter, have had a blood type *O*[10] and even were able to "enunciate as we do" [11]interpreting [12].

References

1. Mashiko H, Yoshida AC, Kikuchi SS, Niimi K, Takahashi E, et al. (2012) Comparative anatomy of marmoset and mouse cortex from genomic expression. *J Neurosci* 32: 5039-5053. [Crossref]
2. Noirel J, Simonson T (2008) Neutral evolution of proteins: The superfunnel in sequence space and its relation to mutational robustness. *J ChemPhys*129: 185104. [Crossref]
3. Briggs AW, Stenzel U, Johnson PL, Green RE, Kelso J, et al. (2007) Patterns of damage in genomic DNA sequences from a Neandertal. *ProcNatl Acad Sci U S A* 104: 14616-14621. [Crossref]
4. Harris K (2015) Evidence for recent, population-specific evolution of the human mutation rate. *Proc Natl Acad Sci U S A* 112: 3439-3444. [Crossref]
5. Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, et al. (2010) Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* 463: 757-762. [Crossref]
6. Der Sarkissian C, Allentoft ME, Ávila-Arcos MC, Barnett R, Campos PF, et al. (2015) Ancient genomics. *Philos Trans R SocLond B BiolSci* 370: 20130387. [Crossref]
7. Sanchez-Quinto F, Lalueza-Fox C (2015) Almost 20 years of Neanderthal palaeogenetics: adaptation, admixture, diversity, demography and extinction. *Philos Trans R SocLond B Biol Sci* 370:20130374.
8. Pollard KS (2009) What makes us human? *Sci Am*300: 44-49.[Crossref]
9. Krause J, Lalueza-Fox C, Orlando L, Enard W, Green RE, et al. (2007) The derived FOXP2 variant of modern humans was shared with Neandertals. *Curr Biol* 17: 1908-1912. [Crossref]