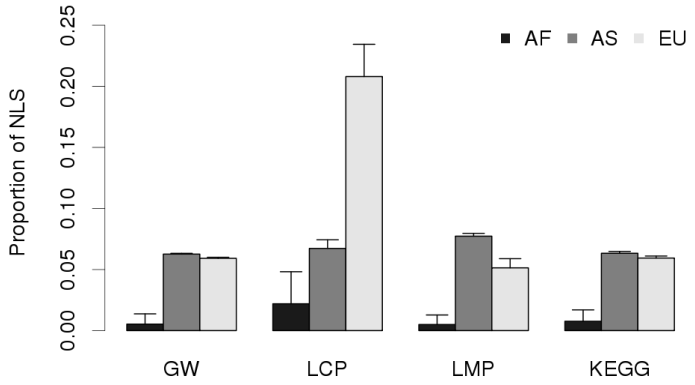
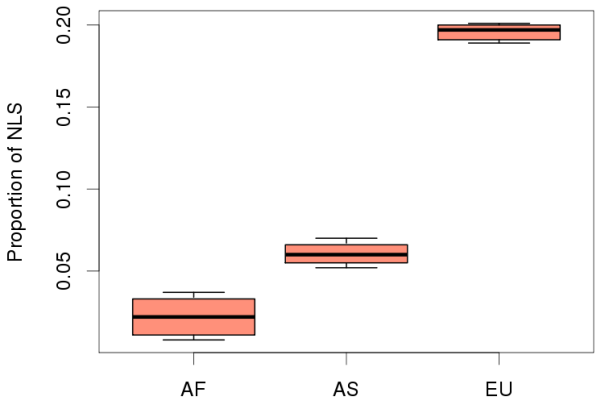


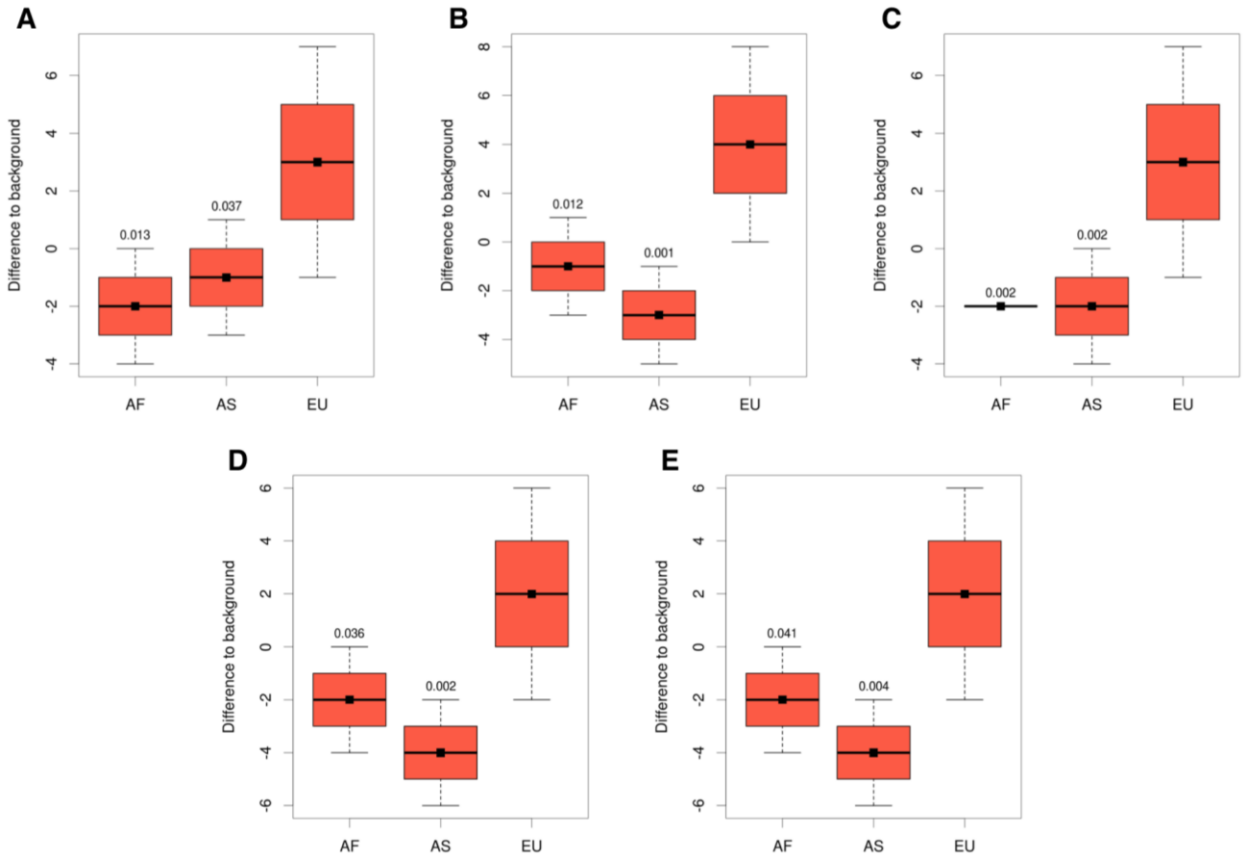
Supplementary Figure 1. Average proportions of NLS in contemporary African (AF, black), Asian (AS, dark gray) and European (EU, light gray) populations calculated based on sequence data from the 1000 genomes project[13].

The proportions were calculated for the whole genome (GW, *n*=1,158,559 sites), lipid catabolic process (LCP, *n*=498 sites) genes, lipid metabolic process (LMP, *n*=5,851 sites) genes, and all genes annotated in KEGG pathways (KEGG, *n*=140,915 sites). The error bars show the standard deviation of the NLS proportion estimates.



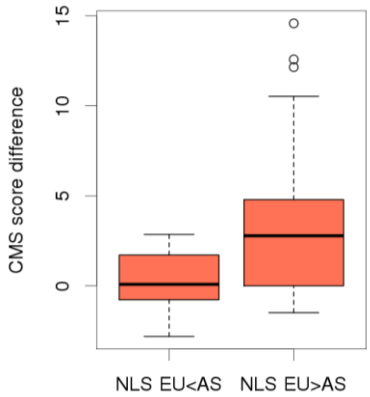
Supplementary Figure 2. Robustness estimates for the average proportions of NLS located within genes annotated in the LCP term in contemporary African, European and Asian populations.

The boxes show quartiles and the median of the data, the whiskers extend to the minimum and maximum data values located within 0.5 interquartile range from the box. The variance shown by the boxplots was calculated by bootstrapping genomic sites (*n*=498) 1,000 times.



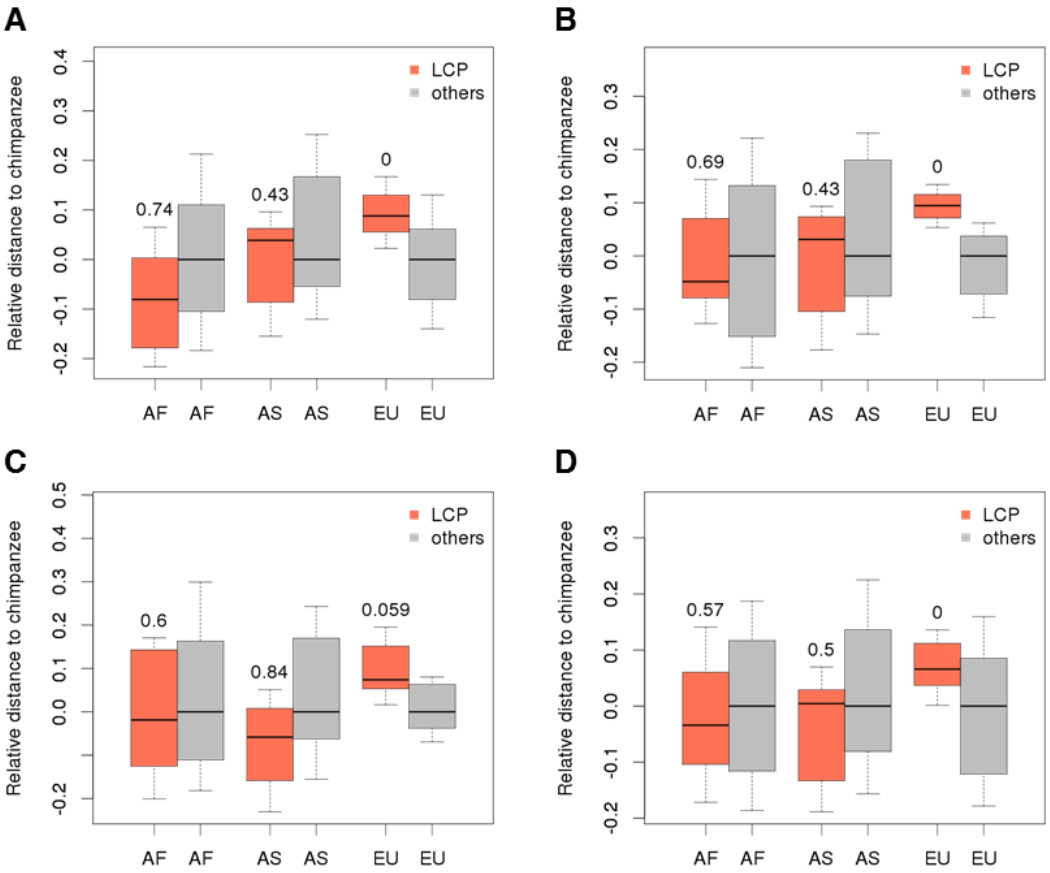
Supplementary Figure 3. Positive selection signal in LCP genes estimated using CMS scores[25].

Black squares represent the number of sites with the positive selection signal, measured as a CMS score higher or equal to a CMS score threshold of 2 (**A**), 1 (**B**) or 0.5 (**C**-**E**) in the LCP gene regions in human populations of European (EU), Asian (AS) and African (AF) descent. For each population, the numbers of sites in the LCP term were normalized to the numbers of sites passing the same CMS score threshold in the same number of gene regions sampled within the same population among all annotated genes (**A**-**C**), among all protein-coding genes (**D**), and among protein-coding genes directly linked with non-LCP metabolic pathways according to KEGG annotation (**E**). The boxplots represent variation of normalized site number estimates obtained by bootstrapping LCP gene regions 1,000 times. The boxes show quartiles and the median of the data, the whiskers extend to the minimum and maximum data values located within 0.5 interquartile range from the box. The numbers above the boxplots show the proportion of the normalized site numbers in Europeans (EU), observed in 1,000 bootstraps over LCP gene regions (n=38), that were less or equal to the normalized site numbers in Africans (AF) or Asians (AS).



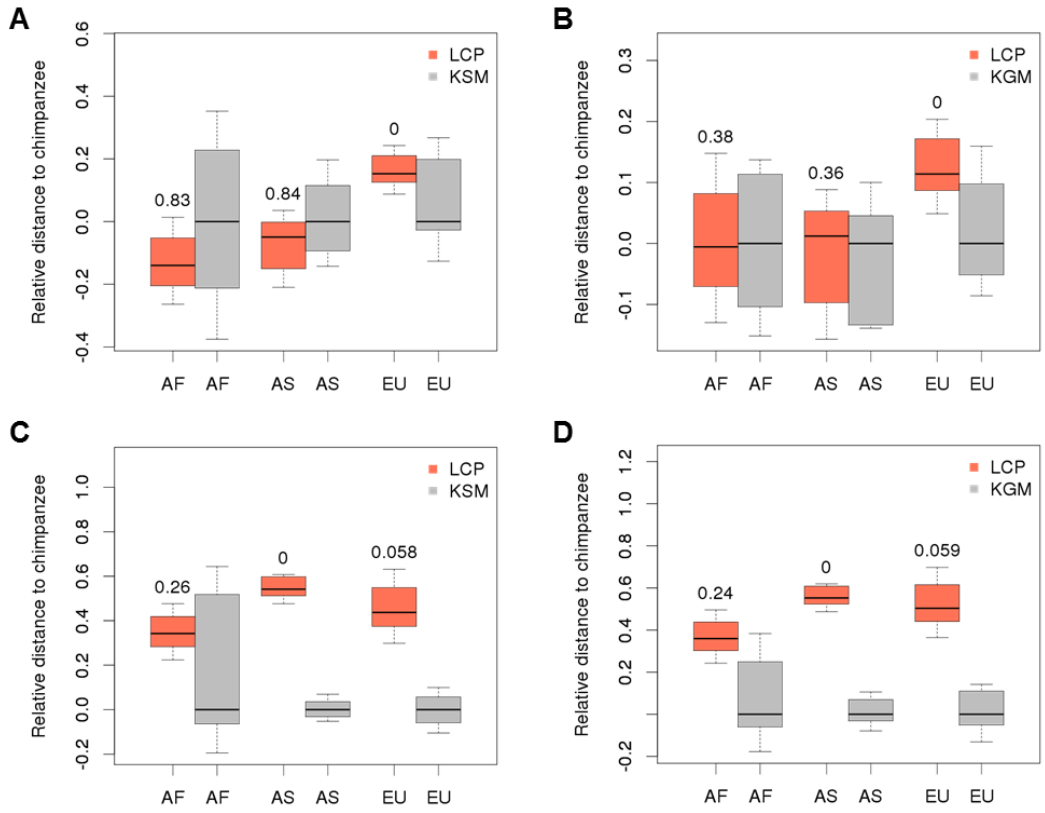
Supplementary Figure 4. Positive selection signal measured using CMS scores for genes within the LCP term in contemporary humans of European decent.

On the right, difference of CMS scores in Europeans to CMS scores in Asians for LCP genes (*n*=45) containing excess of NLS in Europeans compared to Asians (EU>AS). On the left, difference of CMS scores in Europeans to CMS scores in Asians for LCP genes (*n*=20) containing fewer NLS in Europeans compared to Asians (EU<AS). The boxes show quartiles and the median of the data, the whiskers extend to the minimum and maximum data values located within 0.5 interquartile range from the box.



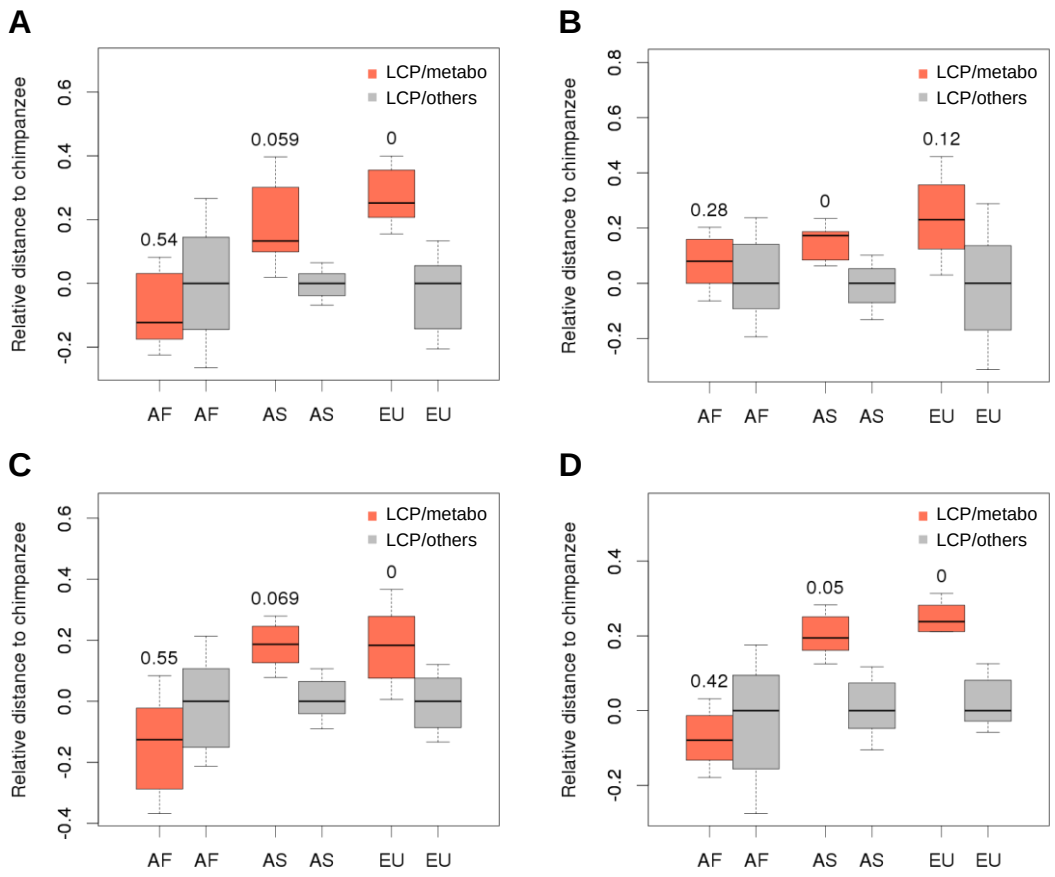
Supplementary Figure 5. Metabolite concentration divergence between chimpanzees and three human populations in the LCP term and other metabolic pathways, and controlled for differences in sex, age, postmortem delay and tissue preservation based on RNA integrity estimates (RIN values) among species and populations.

The boxplots show the distribution of divergence estimates measured between chimpanzees and humans of African (AF), Asian (AS) and European (EU) descent for seven metabolic categories directly linked to LCP genes (red) and metabolites in other metabolic pathways (gray). The boxes show quartiles and the median of the data, the whiskers extend to the minimum and maximum data values located within 0.5 interquartile range from the box. The variance shown by the boxplots was calculated by bootstrapping mass-spectrometry peaks (n=1,090 for LCP and n=163 for other metabolite categories) 1,000 times. To minimize the influence of environmental differences among populations, metabolic divergence in the LCP term was normalized to the divergence based on all other metabolic pathways within the same population. The numbers above the red boxplots show the proportion of values from the LCP divergence distribution that were smaller than, or equal to, the divergence values calculated based on other metabolic pathways shown by the gray boxplots. The distributions shown are based on residuals of the linear regression models used to exclude possible linear effects of differences in (**A**) sex, (**B**) age, (**C**) postmortem delay and (**D**) tissue preservation based on RNA integrity estimates (RIN values) among species and populations.



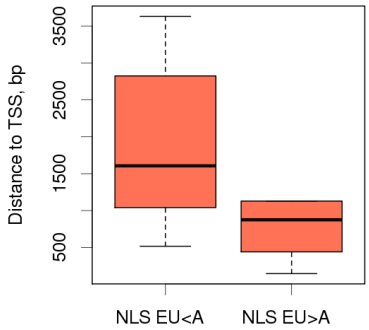
Supplementary Figure 6. Lipid concentration and gene expression divergence in LCP and two other metabolic pathways: KEGG sphingolipid metabolism (KSM) and KEGG glycerolipid metabolism (KGM).

(**A**) and (**B**) Lipid concentration divergence in LCP relative to KSM (**A**) or KGM (**B**) pathways in three human populations. The boxplots show the distribution of divergence estimates measured between chimpanzees and humans of African (AF), Asian (AS) and European (EU) descent for seven metabolic categories directly linked to LCP genes (red), as well as for five metabolic categories in KSM pathway or for five metabolic categories in KGM pathway (gray). (**C**) and (**D**) Gene expression divergence in LCP relative to KSM (**C**) or KGM (**D**) pathways in three human populations. The boxplots show the distribution of divergence estimates measured between chimpanzees and humans of AF, AS and EU descent for LCP genes directly linked to the seven metabolic categories (red), as well as for KSM or KGM genes directly linked to corresponding metabolites (gray). The boxes show quartiles and the median of the data, the whiskers extend to the minimum and maximum data values located within 0.5 interquartile range from the box. The variance shown by the boxplots was calculated by bootstrapping mass-spectrometry peaks (n=1,090 for LCP, n=154 for KSM, and n=474 for KGM) and RPKM values (n=6 for LCP, n=24 for KSM, and n=12 for KGM) 1,000 times. To minimize the influence of environmental differences among populations, metabolic/gene expression divergence in the LCP term was normalized to the divergence of KSM or KGM pathways within the same population. The numbers above the red boxplots show the proportion of values from the LCP divergence distribution that were smaller than, or equal to, the divergence values calculated based on KSM or KGM pathways shown by the gray boxplots.



Supplementary Figure 7. Gene expression divergence between chimpanzees and three human populations in the LCP term and other genes expressed in the prefrontal cortex, controlled for differences in sex, age, postmortem delay and tissue preservation based on RNA integrity estimates (RIN values) among species and populations.

The boxplots show the distribution of divergence estimates measured between chimpanzees and humans of African (AF), Asian (AS) and European (EU) descent for LCP genes directly linked to seven LCP metabolic categories (red) and other LCP genes (gray). The boxes show quartiles and the median of the data, the whiskers extend to the minimum and maximum data values located within 0.5 interquartile range from the box. The variance shown by the boxplots was calculated by bootstrapping RPKM values (n=6 for LCP/metabo and n=26 for LCP/others) 1,000 times. To minimize the influence of environmental differences among populations, gene expression divergence in the LCP term was normalized to the divergence based on all other metabolic pathways within the same population. The numbers above the red boxplots show the proportion of values from the LCP divergence distribution that were smaller than, or equal to, the divergence values calculated based on other metabolic pathways shown by the gray boxplots. The distributions shown are based on residuals of the linear regression models used to exclude possible linear effects of differences in (**A**) sex, (**B**) age, (**C**) postmortem delay and (**D**) tissue preservation based on RNA integrity estimates (RIN values) among species and populations.



Supplementary Figure 8. Distances between NLS located within six LCP genes connected to European-specific metabolic changes and transcription start sites (TSS) of these genes.

On the right, distances to TSS for the polymorphic sites with higher D-score in Europeans compared to Asians and Africans (EU>A, *n*=14), representing European NLS. On the left, distances to TSS for the polymorphic sites with lower D-score in Europeans compared to Asians and Africans (EU<A, *n*=27), representing non-European (Asian) NLS. The boxes show quartiles and the median of the data, the whiskers extend to the minimum and maximum data values located within 0.5 interquartile range from the box.

Supplementary Table 1. Average genome-wide proportion of NLS in eleven contemporary human populations calculated as (#ABBA - #BABA) / (#ABBA + #BABA).

ASW	0.000	0.014	0.015	-0.051	-0.050	-0.048	-0.047	-0.048	-0.046	-0.046	-0.044
LWK	-0.014	0.000	0.001	-0.063	-0.062	-0.061	-0.059	-0.060	-0.058	-0.058	-0.057
YRI	-0.015	-0.001	0.000	-0.063	-0.063	-0.062	-0.060	-0.06	-0.059	-0.059	-0.057
CHB	0.051	0.063	0.063	0.000	0.001	0.003	0.005	0.004	0.006	0.006	0.008
CHS	0.050	0.062	0.063	-0.001	0.000	0.002	0.004	0.003	0.005	0.005	0.007
JPT	0.048	0.061	0.062	-0.003	-0.002	0.000	0.002	0.002	0.003	0.004	0.005
CEU	0.047	0.059	0.060	-0.005	-0.004	-0.002	0.000	-0.001	0.002	0.001	0.003
FIN	0.048	0.060	0.060	-0.004	-0.003	-0.002	0.001	0.000	0.002	0.002	0.004
GBR	0.046	0.058	0.059	-0.006	-0.005	-0.003	-0.002	-0.002	0.000	0.000	0.002
IBS	0.046	0.058	0.059	-0.006	-0.005	-0.004	-0.001	-0.002	0.000	0.000	0.001
TSI	0.045	0.058	0.058	-0.007	-0.006	-0.004	-0.003	-0.003	-0.002	-0.001	0.000

All sites with a sequence difference between chimpanzee and Neanderthal genomes were used (*n*=1,158,559).

Supplementary Table 2. Average genome-wide proportion of NLS in eleven contemporary human populations calculated as (#ABBA - #BABA) / (#ABBA + #BABA).

	ASW	LWK	YRI	CHB	CHS	JPT	CEU	FIN	GBR	IBS	TSI
ASW	0.000	0.014	0.015	-0.050	-0.049	-0.047	-0.047	-0.047	-0.046	-0.045	-0.044
LWK	-0.014	0.000	0.000	-0.062	-0.061	-0.060	-0.059	-0.059	-0.058	-0.058	-0.056
YRI	-0.015	0.000	0.000	-0.062	-0.062	-0.061	-0.059	-0.059	-0.058	-0.058	-0.057
CHB	0.050	0.062	0.062	0.000	0.002	0.003	0.005	0.004	0.006	0.006	0.007
CHS	0.049	0.061	0.062	-0.002	0.000	0.002	0.004	0.003	0.005	0.005	0.007
JPT	0.047	0.060	0.061	-0.003	-0.002	0.000	0.002	0.001	0.003	0.004	0.005
CEU	0.047	0.059	0.059	-0.005	-0.004	-0.002	0.000	-0.001	0.001	0.001	0.003
FIN	0.047	0.059	0.059	-0.004	-0.003	-0.001	0.001	0.000	0.002	0.003	0.004
GBR	0.046	0.058	0.058	-0.006	-0.005	-0.003	-0.001	-0.002	0.000	0.000	0.002
IBS	0.045	0.058	0.058	-0.006	-0.005	-0.004	-0.001	-0.003	0.000	0.000	0.001
TSI	0.045	0.057	0.057	-0.007	-0.006	-0.004	-0.002	-0.003	-0.001	-0.001	0.001

Only sites with a sequence difference between chimpanzee and Neanderthal genomes and showing substitution types potentially representing the results of DNA damage characteristic of ancient DNA samples (C/T and A/G substitutions) were used (*n*=803,089).

Supplementary Table 3. Average genome-wide proportion of NLS in eleven contemporary human populations calculated as (#ABBA - #BABA) / (#ABBA + #BABA).

	ASW	LWK	YRI	CHB	CHS	JPT	CEU	FIN	GBR	IBS	TSI
ASW	0.000	0.014	0.015	-0.052	-0.053	-0.050	-0.049	-0.049	-0.048	-0.048	-0.046
LWK	-0.014	0.000	0.001	-0.065	-0.064	-0.062	-0.061	-0.061	-0.060	-0.060	-0.059
YRI	-0.015	-0.001	0.000	-0.066	-0.065	-0.064	-0.062	-0.062	-0.062	-0.062	-0.059
CHB	0.052	0.065	0.066	0.000	0.001	0.003	0.006	0.004	0.006	0.006	0.008
CHS	0.053	0.064	0.065	-0.001	0.000	0.002	0.004	0.004	0.005	0.005	0.008
JPT	0.050	0.062	0.064	-0.003	-0.002	0.000	0.003	0.002	0.003	0.003	0.006
CEU	0.049	0.061	0.062	-0.006	-0.004	-0.003	0.000	-0.001	0.002	0.001	0.003
FIN	0.049	0.061	0.062	-0.004	-0.004	-0.002	0.001	0.000	0.001	0.000	0.003
GBR	0.048	0.060	0.062	-0.006	-0.005	-0.003	-0.002	-0.001	0.000	0.000	0.002
IBS	0.048	0.060	0.062	-0.006	-0.005	-0.003	-0.001	0.000	0.000	0.000	0.001
TSI	0.047	0.059	0.060	-0.007	-0.007	-0.005	-0.003	-0.003	-0.002	-0.001	0.000

Only sites with a sequence difference between chimpanzee and Neanderthal genomes, excluding sites showing substitution types potentially representing the results of DNA damage characteristic of ancient DNA samples (C/T and A/G substitutions), were used (*n*=355,470).

Supplementary Table 4. Average proportion of NLS in ten contemporary human populations compared to YRI population and calculated as (#ABBA - #BABA) / (#ABBA + #BABA).

	LCP genes			Whole genome		
	C/T + A/G	**Others**	**Total**	**C/T + A/G**	**Others**	**Total**
ASW_YRI	0.046	0.064	0.051	0.015	0.015	0.015
CEU_YRI	0.201	0.229	0.208	0.059	0.062	0.060
CHB_YRI	0.017	0.183	0.068	0.062	0.066	0.063
CHS_YRI	0.009	0.179	0.060	0.062	0.065	0.063
FIN_YRI	0.169	0.210	0.182	0.059	0.062	0.06
GBR_YRI	0.191	0.234	0.202	0.058	0.062	0.059
IBS_YRI	0.236	0.292	0.252	0.058	0.062	0.059
JPT_YRI	0.020	0.198	0.074	0.061	0.064	0.062
LWK_YRI	0.007	0.035	0.015	0.000	0.001	0.001
TSI_YRI	0.182	0.231	0.196	0.057	0.060	0.058

The left panel represents the proportions of sites in genes annotated in the LCP term. The right panel represents the genome average. The proportions are shown for the genomic sites with substitution types potentially caused by DNA damage characteristic of ancient DNA samples (C/T + A/G), for genomic sites containing all other substitution types (Others), and for the total of all genomic sites with a sequence difference between the chimpanzee and the Neanderthal genomes (Total).

Supplementary Table 5. Genomic coordinates of genes in the LCP term (hg19 version), which demonstrate the greatest excess of NLS in populations with European ancestry.

Ensembl Gene ID	Gene Name	Chr	Start (bp)	End (bp)	Description
ENSG00000117054	ACADM	1	76190036	76253260	acyl-CoA dehydrogenase, C-4 to C-12 straight chain
ENSG00000122971	ACADS	12	121163538	121177811	acyl-CoA dehydrogenase, C-2 to C-3 short chain
ENSG00000072778	ACADVL	17	7120444	7128592	acyl-CoA dehydrogenase, very long chain
ENSG00000122787	AKR1D1	7	137687070	137802732	aldo-keto reductase family 1, member D1
ENSG00000112294	ALDH5A1	6	24495080	24537435	aldehyde dehydrogenase 5 family, member A1
ENSG00000132855	ANGPTL3	1	63063158	63071830	angiopoietin-like 3
ENSG00000110244	APOA4	11	116691419	116694022	apolipoprotein A-IV
ENSG00000110243	APOA5	11	116660083	116663136	apolipoprotein A-V
ENSG00000164039	BDH2	4	104000592	104021040	3-hydroxybutyrate dehydrogenase, type 2
ENSG00000170835	CEL	9	135937365	135947248	carboxyl ester lipase
ENSG00000176194	CIDEA	18	12254318	12277594	cell death-inducing DFFA-like effector a
ENSG00000110090	CPT1A	11	68522088	68611878	carnitine palmitoyltransferase 1A (liver)
ENSG00000205560	CPT1B	22	51007290	51017899	carnitine palmitoyltransferase 1B (muscle)
ENSG00000036530	CYP46A1	14	100150641	100193638	cytochrome P450, family 46, subfamily A, polypeptide 1
ENSG00000104823	ECH1	19	39306062	39322645	enoyl CoA hydratase 1, peroxisomal
ENSG00000127884	ECHS1	10	135175984	135187193	enoyl CoA hydratase, short chain, 1, mitochondrial
ENSG00000102393	GLA	X	100652791	100662913	galactosidase, alpha
ENSG00000131373	HACL1	3	15602211	15643338	2-hydroxyacyl-CoA lyase 1
ENSG00000138029	HADHB	2	26466038	26513336	hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase (trifunctional protein), beta subunit
ENSG00000101323	HAO1	20	7863628	7921121	hydroxyacid oxidase (glycolate oxidase) 1 [Source:HGNC Symbol;Acc:4809]
ENSG00000116882	HAO2	1	119911402	119936753	hydroxyacid oxidase 2 (long chain)
ENSG00000198189	HSD17B11	4	88257762	88312538	hydroxysteroid (17-beta) dehydrogenase 11
ENSG00000087076	HSD17B14	19	49316274	49339935	hydroxysteroid (17-beta) dehydrogenase 14
ENSG00000025423	HSD17B6	12	57145945	57181574	hydroxysteroid (17-beta) dehydrogenase 6
ENSG00000162139	NEU3	11	74699179	74729938	sialidase 3 (membrane sialidase)
ENSG00000103066	PLA2G15	16	68279207	68294961	phospholipase A2, group XV
ENSG00000243708	PLA2G4B	15	42120283	42140345	phospholipase A2, group IVB (cytosolic)
ENSG00000105499	PLA2G4C	19	48551100	48614074	phospholipase A2, group IVC (cytosolic, calcium-independent)
ENSG00000100344	PNPLA3	22	44319619	44360368	patatin-like phospholipase domain containing 3
ENSG00000186951	PPARA	22	46546424	46639653	peroxisome proliferator-activated receptor alpha
ENSG00000112033	PPARD	6	35310335	35395968	peroxisome proliferator-activated receptor delta
ENSG00000131238	PPT1	1	40538379	40563375	palmitoyl-protein thioesterase 1
ENSG00000117592	PRDX6	1	173446405	173457946	peroxiredoxin 6
ENSG00000103056	SMPD3	16	68392231	68482591	sphingomyelin phosphodiesterase 3, neutral membrane (neutral sphingomyelinase II)
ENSG00000136699	SMPD4	2	130908981	130940323	sphingomyelin phosphodiesterase 4, neutral

					membrane (neutral sphingomyelinase-3)
ENSG00000101846	STS	X	7137497	7272851	steroid sulfatase (microsomal), isozyme S
ENSG00000156096	UGT2B4	4	70345883	70391732	UDP glucuronosyltransferase 2 family, polypeptide B4
ENSG00000128245	YWHAH	22	32340447	32353590	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide

The genes are located in 38 independent genomic regions and are uniformly distributed on human chromosomes.

Supplementary Table 6. Numbers of NLS and Neanderthal-chimpanzee divergent sites in LCP genes.

Gene	NLS	Divergent sites	Gene length (nt)
ENSG00000072778	0	0	8148
ENSG00000101846	0	0	135354
ENSG00000102393	0	0	10122
ENSG00000103066	0	0	15754
ENSG00000110244	0	0	2603
ENSG00000128245	0	0	13143
ENSG00000136699	0	0	31342
ENSG00000243708	0	0	10372
ENSG00000025423	0	1	35629
ENSG00000110243	0	1	3053
ENSG00000132855	0	1	8672
ENSG00000116882	0	2	25351
ENSG00000170835	0	2	9883
ENSG00000122971	0	4	14273
ENSG00000176194	0	4	23276
ENSG00000104823	1	3	16429
ENSG00000112294	1	3	42355
ENSG00000162139	1	6	30759
ENSG00000164039	1	13	20448
ENSG00000138029	2	21	47298
ENSG00000087076	3	5	23655
ENSG00000205560	3	6	10609
ENSG00000156096	3	21	45849
ENSG00000036530	3	23	42997
ENSG00000110090	3	26	89790
ENSG00000131238	4	12	24996
ENSG00000117592	5	6	11541
ENSG00000112033	6	38	85633
ENSG00000103056	7	21	90360
ENSG00000127884	9	12	11209
ENSG00000198189	9	25	54776
ENSG00000100344	9	35	40749
ENSG00000122787	9	40	115662
ENSG00000131373	10	14	41127
ENSG00000101323	14	29	57493
ENSG00000117054	14	29	63224
ENSG00000186951	16	42	93229
ENSG00000105499	35	53	63009

Supplementary Table 7. Average proportion of NLS for the samples with high genome coverage (12x whole genome coverage or higher) calculated as (#ABBA - #BABA) / (#ABBA + #BABA).

	LCP genes			Whole genome		
	C/T + A/G	**Others**	**Total**	**C/T + A/G**	**Others**	**Total**
ASW_YRI	0.097	0.117	0.101	0.015	0.016	0.015
CEU_YRI	0.334	0.110	0.263	0.060	0.065	0.061
CHB_YRI	0.112	0.205	0.141	0.063	0.066	0.064
CHS_YRI	NA	NA	NA	NA	NA	NA
FIN_YRI	0.174	0.165	0.172	0.059	0.061	0.059
GBR_YRI	0.248	0.237	0.242	0.059	0.062	0.060
IBS_YRI	0.296	0.258	0.282	0.057	0.061	0.058
JPT_YRI	0.090	0.219	0.132	0.060	0.065	0.062
LWK_YRI	0.059	-0.022	0.037	0.002	0.003	0.002
TSI_YRI	NA	NA	NA	NA	NA	NA

The left panel represents the proportions of sites in genes annotated in the LCP term. The right panel represents the genome average. The proportions are shown for the genomic sites with substitution types potentially caused by DNA damage characteristic of ancient DNA samples (C/T + A/G), for genomic sites containing all other substitution types (Others), and for the total of all genomic sites with a sequence difference between the chimpanzee and the Neanderthal genomes (Total).

Supplementary Table 8. Average proportion of NLS in individuals of African, East Asian and European ancestry with high genome coverage available in the Complete Genomics diversity set.

	LCP genes			Whole genome		
	C/T + A/G	**Others**	**Total**	**C/T + A/G**	**Others**	**Total**
ASW_YRI	0.090	0.011	**0.065**	0.013	0.013	**0.013**
LWK_YRI	0.043	0.179	**0.084**	0.000	0.001	**0.001**
MKK_YRI	0.138	0.052	**0.112**	0.010	0.012	**0.011**
CHB_YRI	0.067	0.241	**0.124**	0.059	0.062	**0.060**
JPT_YRI	0.029	0.106	**0.054**	0.063	0.063	**0.063**
CEU_YRI	0.277	0.178	**0.245**	0.056	0.054	**0.055**
TSI_YRI	0.201	0.187	**0.197**	0.054	0.058	**0.055**

ASW - HapMap African ancestry individuals from SW US, CEU - CEPH individuals, CHB - Han Chinese in Beijing, JPT - Japanese individuals, LWK - Luhya individuals, MKK - Maasai in Kenya, TSI - Toscan individuals, and YRI - Yoruba individuals. The proportion was calculated as (#ABBA - #BABA) / (#ABBA + #BABA). The left panel represents the proportions of sites in genes annotated with the LCP term. The right panel represents the genome average. The proportions are shown for the genomic sites with substitution types potentially caused by DNA damage characteristic of ancient DNA samples (C/T + A/G), for genomic sites containing all other substitution types (Others), and for the total of all genomic sites with a sequence difference between the chimpanzee and the Neanderthal genomes (Total).

Supplementary Table 9. Average proportion of NLS in ten contemporary human populations compared to YRI population.

	LCP genes			Whole genome		
	Altai	**Vindjia**	**Combined**	**Altai**	**Vindjia**	**Combined**
ASW_YRI	0.039	0.052	**0.051**	0.015	0.015	**0.015**
CEU_YRI	0.153	0.240	**0.208**	0.060	0.059	**0.060**
CHB_YRI	0.103	0.118	**0.068**	0.063	0.063	**0.063**
CHS_YRI	0.111	0.127	**0.060**	0.063	0.063	**0.063**
FIN_YRI	0.129	0.242	**0.182**	0.060	0.059	**0.059**
GBR_YRI	0.133	0.233	**0.202**	0.059	0.059	**0.059**
IBS_YRI	0.180	0.261	**0.252**	0.059	0.058	**0.059**
JPT_YRI	0.107	0.129	**0.074**	0.062	0.062	**0.062**
LWK_YRI	0.033	0.017	**0.015**	0.001	0.001	**0.001**
TSI_YRI	0.155	0.240	**0.196**	0.058	0.057	**0.058**

Proportions were calculated using the Altai and Vindjia Neanderthal genomes separately, or using the consensus genome built using both Neanderthal genomes (Combined). The proportions are shown for LCP genes and for the entire genome: the total of all genomic sites with a sequence difference between the chimpanzee and the Neanderthal genomes.

Supplementary Table 10. Genes within the LCP term detected in the RNA-seq data that are directly linked to the seven metabolic categories within the LCP pathway.

Gene name	Linked metabolites *	Gene function **
CEL	C00422, C02530	GO:0006629 lipid metabolic process; GO:0006641 triglyceride metabolic process; GO:0006707 cholesterol catabolic process; GO:0009062 fatty acid catabolic process; GO:0018350 protein esterification; GO:0030157 pancreatic juice secretion; GO:0030299 intestinal cholesterol absorption; GO:0044241 lipid digestion; GO:0044258 intestinal lipid catabolic process; GO:0044281 small molecule metabolic process; GO:0046514 ceramide catabolic process
PLA2G15	C04230	GO:0006672 ceramide metabolic process; GO:0009062 fatty acid catabolic process; GO:0046470 phosphatidylcholine metabolic process; GO:0006629 lipid metabolic process
PLA2G4B	C00350, C04230, C00157	GO:0006644 phospholipid metabolic process; GO:0006654 phosphatidic acid biosynthetic process; GO:0006954 inflammatory response; GO:0006987 activation of signaling protein activity involved in unfolded protein response; GO:0007567 parturition; GO:0019369 arachidonic acid metabolic process; GO:0019722 calcium-mediated signaling; GO:0030968 endoplasmic reticulum unfolded protein response; GO:0036148 phosphatidylglycerol acyl-chain remodeling; GO:0036150 phosphatidylserine acyl-chain remodeling; GO:0036151 phosphatidylcholine acyl-chain remodeling; GO:0036152 phosphatidylethanolamine acyl-chain remodeling; GO:0044267 cellular protein metabolic process; GO:0044281 small molecule metabolic process; GO:0046474 glycerophospholipid biosynthetic process; GO:0046475 glycerophospholipid catabolic process
PLA2G4C	C00350, C04230, C00157	GO:0006644 phospholipid metabolic process; GO:0006954: inflammatory response; GO:0007567 parturition; GO:0019369 arachidonic acid metabolic process; GO:0035556 intracellular signal transduction; GO:0036149 phosphatidylinositol acyl-chain remodeling; GO:0036151 phosphatidylcholine acyl-chain remodeling; GO:0036152 phosphatidylethanolamine acyl-chain remodeling; GO:0044281 small molecule metabolic process; GO:0046474 glycerophospholipid biosynthetic process; GO:0046475 glycerophospholipid catabolic process
PNPLA3	C00422	GO:0036155 acylglycerol acyl-chain remodeling; GO:0046474 glycerophospholipid biosynthetic process; GO:0006644 phospholipid metabolic process; GO:0044281 small molecule metabolic process; GO:0019432 triglyceride biosynthetic process; GO:0019433 triglyceride catabolic process
SMPD3	C00195, C00550	GO:0007049 cell cycle; GO:0006687 glycosphingolipid metabolic process; GO:0007275 multicellular organismal development; GO:0030072 peptide hormone secretion; GO:0044281 small molecule metabolic process; GO:0006685 sphingomyelin catabolic process

* KEGG accession numbers of the linked metabolic categories.

** Gene function according to Gene Ontology annotation[20]

Supplementary Table 11. Sample information for the prefrontal cortex samples used in metabolite and RNA-seq measurements.

Species	Sample_#	RNA-seq	Gender	Age	RNA RIN	Postmortem	Source
Human_EU	1	Y	M	21,9	9,7	13	Maryland
Human_AF	2	Y	M	22,9	7,3	4	Maryland
Human_EU	3	Y	M	27,1	8,8	15	Maryland
Human_EU	4	Y	M	30	7,4	19	Maryland
Human_AS	5	Y	M	37	6,9	5	Wuhan
Human_AF	6	Y	M	38,6	7,3	5	Maryland
Human_AS	7	Y	F	47	7,4	5	Wuhan
Human_AF	8	Y	M	50,4	6,9	8	Maryland
Human_AS	9	Y	F	52	6,8	5	Wuhan
Human_AS	10	Y	F	52	7,1	4	Wuhan
Human_EU	11	Y	M	53,6	6,7	19	Maryland
Human_AF	12	Y	M	58,1	8,4	9	Maryland
Human_AS	13	Y	F	59	6,5	5	Wuhan
Human_EU	14	Y	M	61,5	8,8	6	Maryland
Chimpanzee	1	Y	F	27,9	8,2	5	Irchel
Chimpanzee	2	Y	F	41,8	7,1	6	Irchel
Chimpanzee	3	Y	M	37,8	8,5	5	Irchel
Chimpanzee	4	Y	F	21,4	7,6	5	Yerkes
Chimpanzee	5	Y	M	26,4	7,9	5	Yerkes
Chimpanzee	6	Y	M	22,7	7	4	Yerkes
Chimpanzee	7	N	F	43,4	8,7	4	BPRC
Chimpanzee	8	N	F	39,9	8,3	5	BPRC
Chimpanzee	9	N	F	31,7	7,8	5	BPRC
Chimpanzee	10	N	M	28,3	7,5	10	BPRC
Chimpanzee	11	N	F	20,2	7,7	5	BPRC
Chimpanzee	12	N	F	20,5	6,9	5	BPRC
Chimpanzee	13	N	M	24,3	7,8	6	BPRC
Chimpanzee	14	N	M	10,4	7,3	5	BPRC

In the "Species" column, EU stands for European ancestry, AS stands for Asian ancestry, and AF stands for African ancestry. Six of the 14 chimpanzee samples that were used for RNA-seq measurements are marked with "Y" in the "RNA-seq" column. For all individuals, listed are: sex (M/F), age (in years), RNA RIN (RNA Integrity Number – values over 7 indicate excellent RNA preservation), postmortem delay (in hours), and sample source ('Maryland' stands for the NICHD Brain and Tissue Bank for Developmental Disorders at the University of Maryland, USA; 'Wuhan' stands for the Chinese Wuhan Xiehe Hospital, China; 'Irchel' stands for the Anthropological Institute & Museum of the University of Zürich-Irchel, Switzerland; 'Yerkes' stands for the Yerkes Primate Center, USA; and 'BPRC' stands for the Biomedical Primate

Research Centre, the Netherlands).

Supplementary Table 12. RNA-seq read mapping statistics for the 14 human samples and 6 chimpanzee prefrontal cortex samples used for RNA-seq measurements.

Species	Sample_#	Total_reads	Unique_mapped	Multiple_mapped	Unique_%	Multiple_%
Human_EU	1	17052134	14653755	1085887	85.94	6.37
Human_AF	2	17670579	15215644	1118161	86.11	6.33
Human_EU	3	15644246	13421192	1020370	85.79	6.52
Human_EU	4	14955507	11691156	2082619	78.17	13.93
Human_AS	5	14269905	12248034	837581	85.83	5.87
Human_AF	6	14842000	12837513	805565	86.49	5.43
Human_AS	7	14171211	12079821	942723	85.24	6.65
Human_AF	8	14982206	12679276	1024410	84.63	6.84
Human_AS	9	15540198	13496533	723758	86.85	4.66
Human_AS	10	15027750	13153203	679557	87.53	4.52
Human_EU	11	14190548	10436837	2584520	73.55	18.21
Human_AF	12	15134633	13188007	709308	87.14	4.69
Human_AS	13	15124547	13065560	815535	86.39	5.39
Human_EU	14	16085282	13994069	829696	87.00	5.16
Chimpanzee	1	3977099	3086707	185344	77.61	4.66
Chimpanzee	2	3887122	2979788	194046	76.66	4.99
Chimpanzee	3	5965770	4179508	304563	70.06	5.11
Chimpanzee	4	4473074	3417107	240112	76.39	5.37
Chimpanzee	5	5029692	3769145	259822	74.94	5.17
Chimpanzee	6	4312534	3357557	207772	77.86	4.82

In the “Species” column, EU stands for European ancestry, AS stands for Asian ancestry, and AF stands for African ancestry. Listed are: total sequence reads number (Total_reads), number of sequence reads that uniquely mapped to the corresponding genome (Unique_mapped) and the corresponding percentage of total reads (Unique_%), number of sequence reads that mapped to the corresponding genome at multiple genomic locations (Multiple_mapped) and the corresponding percentage of total reads (Multiple_%).

Supplementary Table 13. Human disease terms containing enriched numbers of LCP genes, according to the Genetic Association Database[34].

Disease term	Count	PValue	Genes	Fold Enrichment	Benjamini	FDR
obesity	8	3,3E-06	ENSG00000176194, ENSG00000186951, ENSG00000205560, ENSG00000112033, ENSG00000110243, ENSG00000110244, ENSG00000100344, ENSG00000110090	10,3	5,03E-04	0,0040
Hypertriglyceridemia	3	0,0011	ENSG00000186951, ENSG00000110243, ENSG00000110244	53,1	0,0836	1,3896
cholesterol, HDL cholesterol, LDL lipoprotein triglycerides	3	0,0033	ENSG00000186951, ENSG00000110243, ENSG00000110244	31,8	0,1534	3,9215
coronary heart disease	4	0,0034	ENSG00000186951, ENSG00000112033, ENSG00000110243, ENSG00000110244	12,0	0,1202	4,0180
triglycerides	3	0,0043	ENSG00000186951, ENSG00000110243, ENSG00000110244	28,1	0,1206	5,0175
cholesterol	3	0,0065	ENSG00000036530, ENSG00000110243, ENSG00000110244	22,7	0,1507	7,5465
cholesterol; triglycerides	2	0,0118	ENSG00000186951, ENSG00000110243	159,3	0,2242	13,260
cholesterol, HDL; triglycerides; diabetes, type 2; cholesterol, LDL	2	0,0118	ENSG00000186951, ENSG00000112033	159,3	0,2242	13,260
medium-chain acyl-CoA dehydrogenase deficiency	2	0,0118	ENSG00000122971, ENSG00000117054	159,3	0,2242	13,260
left ventricular hypertrophy	3	0,0131	ENSG00000186951, ENSG00000205560, ENSG00000110090	15,9	0,2186	14,614
metabolic syndrome	3	0,0140	ENSG00000186951, ENSG00000112033, ENSG00000110243	15,4	0,2084	15,496
triglycerides; atherosclerosis, coronary; hypertriglyceridemia	2	0,0177	ENSG00000110243, ENSG00000110244	106,2	0,2340	19,218
triglycerides; atherosclerosis, coronary; lipoprotein	2	0,0177	ENSG00000186951, ENSG00000112033	106,2	0,2340	19,218
body mass; lipids	2	0,0235	ENSG00000110243, ENSG00000110244	79,6	0,2761	24,769
atherosclerosis, coronary; lipoprotein	2	0,0235	ENSG00000110243, ENSG00000110244	79,6	0,2761	24,769
cholesterol cholesterol, HDL triglycerides	2	0,0235	ENSG00000110243, ENSG00000110244	79,6	0,2761	24,769
heart disease, ischemic	3	0,0260	ENSG00000186951, ENSG00000110243, ENSG00000110244	11,1	0,2798	27,052
atherosclerosis, coronary hyperlipidemia	2	0,0351	ENSG00000186951, ENSG00000112033	53,1	0,3365	34,757
cholesterol, HDL triglycerides	2	0,0466	ENSG00000110243, ENSG00000110244	39,8	0,3984	43,424

Alzheimer's Disease	5	0,0480	ENSG00000186951, ENSG00000112033, ENSG00000036530, ENSG00000110243, ENSG00000110244	3,3	0,3865	44,391

Fisher's exact test *p*-values, fold enrichments, Benjamini-corrected Fisher's exact test *p*-values and FDR were estimated using DAVID functional annotation tool[33] with default parameters. Only diseases with *p*-values lower than 0.05 are listed.

Supplementary Table 14. Denisova sequence at sites showing sequence divergence between the Neanderthal and the chimpanzee genomes located within LCP genes.

Population	Denisova sequence	Divergent sites	NLS (D>0.3)	%	**NLS (D>0.5)**	**%**	NLS (D>0.7)	%
EU	C	80	48	60.0	**40**	**50.00**	33	41.25
EU	H	17	3	17.6	**3**	**17.65**	3	17.65
EU	N	259	137	52.9	**116**	**44.79**	107	41.31
AS	C	72	29	40.3	**23**	**31.94**	22	30.56
AS	H	19	6	31.6	**6**	**31.58**	6	31.58
AS	N	256	136	53.1	**130**	**50.78**	121	47.27
AF	C	75	12	16.0	**8**	**10.67**	8	10.67
AF	H	17	0	0.0	**0**	**0.00**	0	0.00
AF	N	258	37	14.3	**9**	**3.49**	4	1.55

'C' stands for the chimpanzee state, 'H' stands for the heterozygous sites, 'N' stands for the Neanderthal state. The divergent sites are sites showing sequence divergence between the Neanderthal and the chimpanzee genomes located within LCP genes. We estimated whether these sites have a high frequency of Neanderthal genotype in contemporary human populations using the D statistic, *i.e.* whether these sites could be classified as NLS in specific human populations. We note, however, that the D statistic estimates at individual sites are far less reliable than the estimates based on a large numbers of sites. Due to the relatively small numbers of divergent sites, we used relaxed D statistic cutoffs to determine putative NLS resulting in higher frequencies of NLS sites than in the genome wide analysis.