Supplementary Table S4. Polymorphism and divergence in hominids at homologous positions in the 1,115 bp sequence including the 870 bp *NAT2* coding exon.

The screened sequence spans from 18'257'489 to 18'258'603 on chromosome 8 in the human reference sequence GRCh37/hg19. Positions downstream the *NAT2* coding exon are reported in italic font. Non-synonymous mutations are shown in bold type. The 38 substitutions corresponding to inter-species divergence (between two species at least) are highlighted by stars to the left part of the table. Polymorphisms shared between species or sub-species are highlighted in blue font. (N) indicates that undefined positions were also observed in some individuals. Variants also recorded in sequenced ancient genomes of hominins (Denisova, Neanderthal or 45'000 years old *Homo sapiens* from Ust'-Ishim) are boxed (see Supplementary Table S8).

	Position				Common chimpanzee[2]					Gorilla[4]		Orangutan[5]	
	human reference sequence	human cds	SNP ID in Ensembl	Human[1]	Western	Nigeria-Cameroon	Eastern	Central	Bonobo[3]	*Gorilla gorilla*	*Gorilla beringei*	*Pongo abelii*	*Pongo pygmaeus*
	18257520	**7**	**rs200893121**	**A/G**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**
	18257521	**8**	**rs186884477**	**T/A**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**
	18257542	**29**	**rs72466456**	**T/C**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**
	18257549	36		T	T	T	T/C	T	T	T	T	T	T
	18257576	**63**		**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G/T**
	18257583	**70**	**rs45477599**	**T/A**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**
	18257585	**72**		**A**	**A**	**A**	**A**	**A**	**A/C**	**A**	**A**	**A**	**A**
	18257612	99		G	G	G	G	G	G	G/A(N)	A	G	G
	18257617	**104**	**rs199780526**	**T/C**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**
	18257621	108		C	C	C	C	C	C	C	C	C/T	C/T
	18257624	111	rs72554615	T/C	T	T	T	T	T	T	T	T	T
	18257630	117		C	C	C	C	C	C	C	C	C/T(N)	C(N)
	18257634	**121**	**rs149283608**	**A/T**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**
*	18257654	141		C	C	C	C	C	C	C	C	T	T
*	**18257658**	**145**		**G**	**G**	**G**	**G**	**G**	**A**	**G**	**G**	**G**	**G**
	18257665	**152**	**rs72466457**	**G/T**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**
	18257670	**157**		**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G/A(N)**	**G/A(N)**
	18257703	**190**	**rs1805158**	**C/T**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**
*	**18257704**	**191**	**rs1801279**	**G/A**	**G**	**G**	**G**	**G**	**G/A**	**T**	**T**	**G**	**G**
	18257712	**199**	**rs201345576**	**T/C**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**

	21	**18257716**	**203**	**rs72466458**	**G/A**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**
	22	18257732	219	rs750089838	A	A	A	A	A	A	A/G(N)	A	A	A
	23	18257738	225	rs200775849	G/C	G	G	G	G	G	G	G	G	G
	24	18257741	228	rs72466459	C/T	C	C	C	C	C	C	C	C	C
	25	18257795	282	rs1041983	C/T	C	C	C	C	C	C	C	C	C
*	**26**	**18257806**	**293**		**T**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**T**	**T**
	27	**18257821**	**308**	**rs749903527**	**C/T**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**
	28	**18257854**	**341**	**rs1801280**	**T/C**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**
*	**29**[6]	**18257858**	**345**	**rs45532639**	**C/T**	**A**	**A**	**A**	**A**	**A**	**C/T**	**C/T**	**T**	**T**
	30	**18257859**	**346**	**rs183409091**	**G/A**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**
	31	18257867	354	rs146405047	T/C	T	T	T	T	T	T	T	T	T
*	32	18257876	363	rs372383623	C	C	C	C	C	C	C	C	T	T
	33	**18257877**	**364**	**rs4986996**	**G/A**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**
	34	**18257916**	**403**	**rs12720065**	**C/G**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**
*	35	18257921	408		A	G	G	G	G	G	G	G	G	G
	36	**18257924**	**411**	**rs4986997**	**A/T**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**
	37	**18257947**	**434**	**rs72554616**	**A/C**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**
	38	18257951	438		G	G	G	G	G	G	G/A	A	G	G
*	**39**	**18257956**	**443**		**G**	**C**	**C**	**C**	**C**	**C**	**G**	**G**	**G**	**G**
*	**40**	**18257964**	**451**	**rs747336755**	**T**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**T/C**	**C**
	41	**18257965**	**452**		**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G/A**	**G**
	42	**18257971**	**458**	**rs72466460**	**C/T**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**
*	**43**	**18257978**	**465**	**rs368704280**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**C**	**C**
*	44	18257984	471		A	A	A	A	A	A	A	A	C	C
	45	**18257985**	**472**	**rs139351995**	**A/C**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**
	46	18257994	481	rs1799929	C/T	C	C	C	C	C	C	C	C	C
	47	**18258004**	**491**	**rs139512288**	**T/C**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**
	48	**18258012**	**499**	**rs72554617**	**G/A**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**
	49	**18258019**	**506**	**rs200585149**	**A/G**	**A**	**A**	**A**	**A**	**A**	**A/G**	**A**	**A**	**A**
*	**50**	**18258024**	**511**		**A**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**
	51	**18258027**	**514**		**A**	**A**	**A/G**	**A/G**	**A/G**	**A**	**A**	**A**	**A**	**A**
	52	18258029	516		C	C	C	C	C	C	C	C	C	C/T

	53	18258030	517		A	A	A	A	A	A	A	A	A/G	A/G
	54	18258031	518	rs369500066	A/G	A	A	A	A	A	A	A	A	A
	55	18258033	520	rs753856548	G	G	G	G	G	G	G/T	G	G	G
*	56	18258044	531	rs757090914	T	T	T	T	T	T	T	T	C	C
*	57	18258047	534		T	T	T	T	T	T	T	T	A	A
	58	18258049	536	rs572750517	A	A	A	A	A	A	A	A	A/G	A
*	59	18258054	541		C	C	C	C	C	C	C	C	G	G
*	60	18258085	572		T	T	T	T	T	T	T	T	G	G
*	61	18258086	573		A	C	C	C	C	C	C	C	C	C
	62	18258091	578	rs79050330	C/T	C/T	C	C	C	C	C	C	C	C
	63	18258092	579	rs144176822	G/T	G	G	G	G	G	G	G	G	G
*	64	18258096	583	rs759059031	G	G	G	G	G	G	G	G	C	C
*	65	18258101	588		T	T	T	T	T	T	T	T	C	C
	66	18258102	589	rs375746304	C/T	C	C	C	C	C	C	C	C	C
	67	18258103	590	rs1799930	G/A	G	G	G	G	G	G	G	G	G
*	68	18258107	594	rs563563855	A	G	G	G	G	G	A	A	A	A
*	69	18258108	595		A	G	G	G	G	G	A	A	A	A
	70	18258113	600	rs72466461	A/G	A	A	A	A	A	A	A	A	A
	71	18258122	609	rs45618543	G/T	G	G	G	G	G	G	G	G	G
	72	18258127	614		T	T	T	T	T	T	T/C	T	T	T
	73	18258135	622	rs56387565	T/C	T	T	T	T	T	T	T	T	T
	74	18258146	633		G/A	G	G	G	G	G	G	G	G	G
	75	18258151	638	rs138707146	C/T	C	C	C	C	C	C	C	C	C
	76	18258154	641		C/T	C	C	C	C	C	C	C	C	C
	77	18258161	648		A	A	A	A	A	A	A/G	G	A	A
*	78	18258177	664		T	C	C	C	C	C	C	C	T	T/C
	79	18258178	665		T/G	T	T	T	T	T	T	T	T	T
	80	18258196	683	rs45518335	C/T	C	C	C	C	C	C	C	C	C
*	81	18258207	694		T	T	T	T	T	T	T	T	C	C
*	82	18258241	728		A	A	A	A	A	A	A	A	T	T
	83	18258272	759	rs56011192	C/T	C	C	C	C	C	C	C	C	C
	84	18258279	766	rs55700793	A/G	A	A	A	A	A	A	A	A	A

*	85	18258287	774		C	G	G	G	G	G	G	G	G	G
	86	**18258301**	**788**		**T**	**T**	**T**	**T**	**T**	**T**	**T/C**	**T**	**T**	**T**
	87	18258302	789		T	G/T	T	T	T	T	T	T	T	T
	88[7]	**18258316**	**803**	**rs1208**	**A/G**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**
	89	**18258322**	**809**		**T/C**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**
*	**90**	**18258347**	**834**		**T**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**
*	**91[8]**	**18258351**	**838**	**rs56393504**	**G/A**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**C**	**C**
	92	**18258358**	**845**	**rs56054745**	**A/C**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**
	93	**18258370**	**857**	**rs1799931**	**G/A**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**G**
	94	**18258372**	**859**	**rs72554618**	**T/del**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**
*	*95*	*18258390*	*877*		*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*T*	*T*
	96[9]	*18258407*	*894*		*C*	*C*	*C*	*C*	*C*	*C*	*C/CC*	*C*	*C*	*C*
	97	*18258415*	*902*		*T*	*T*	*T*	*T*	*T*	*T*	*T*	*T*	*T*	*T/A*
*	*98*	*18258441*	*928*		*A*	*G*	*G*	*G*	*G*	*G*	*A*	*A*	*A*	*A*
*	*99*	*18258444*	*931*		*T*	*T*	*T*	*T*	*T*	*T*	*T*	*T*	*G*	*G*
	100	*18258447*	*934*		*G*	*G/A*	*G/A*	*G/A*	*G/A*	*G*	*G*	*G*	*G*	*G*
*	*101*	*18258456*	*943*		*G*	*G*	*G*	*G*	*G*	*G*	*G(N)*	*G(N)*	*C*	*C*
	102	*18258462*	*949*		*C*	*C/G*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*
	103	*18258469*	*956*		*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C/T*	*C*
	104	*18258475*	*962*	*rs754854317*	*C*	*C*	*C*	*C*	*C*	*C*	*C/T(N)*	*C*	*C*	*C*
*	*105*	*18258476*	*963*	*rs539028538*	*G*	*G*	*G*	*G*	*G*	*G*	*G*	*G*	*A*	*A*
	106	*18258480*	*967*		*T*	*T*	*T*	*T*	*T*	*T*	*T/C*	*T*	*T*	*T*
*	*107*	*18258491*	*978*		*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*C*	*C*
*	*108*	*18258497*	*984*		*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*C*	*C*
	109	*18258502*	*989*		*C/G*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*
*	*110*	*18258508*	*995*		*C*	*C*	*C*	*C*	*C*	*C*	*C*	*C*	*G*	*G*
*	*111*	*18258518*	*1005*		*T*	*C*	*C*	*C*	*C*	*C*	*T*	*T*	*T*	*T*
	112	*18258524*	*1011*		*T*	*T*	*T*	*T*	*T*	*T*	*T/C*	*C*	*T*	*T*
	113	*18258534*	*1021*		*T/C*	*T*	*T*	*T*	*T*	*T*	*T*	*T*	*T*	*T*
*	*114*	*18258538*	*1025*		*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*G*	*G*
	115	*18258539*	*1026*	*rs185257970*	*A/G*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*

*	*116*	*18258551*	*1038*	*T*	*T*	*T*	*T*	*T*	*T*	*T*	*T*	*C*	*C*
	117	*18258598*	*1085*	*G/A*	*G*	*G*	*G*	*G*	*G*	*G*	*G*	*G*	*G*

[1] Human polymorphisms are those recorded by the consensus gene nomenclature of human *NAT* alleles (http://nat.mbg.duth.gr/), complemented with haplotype data from 1000Genomes Phase 1 (The Genomes Project Consortium 2012), (Sabbagh et al. 2008), (Patin et al. 2006a), (Mortensen et al. 2011) and (Podgorna et al. 2015). Human positions not recorded as polymorphic but associated with a SNP identifier are reported with a highest population MAF < 0.01 in Ensembl (Yates et al. 2016) (http://www.ensembl.org/Homo_sapiens/Info/Index).

[2] Western, Niger-Cameroon, Eastern and Central chimpanzees are *Pan troglodytes verus*, *P. t. ellioti*, *P. t. schweinfurthii* and *P. t. troglodytes*, respectively. Polymorphism recording is based on the individuals of the present study (Supplementary Figure S1) and the chimpanzees of Prado-Martinez et al. (2013) cross-checked with the *P. t. verus* assembly reference sequence (panTro4, February 2011). For Eastern chimpanzees (*P. t. schweinfurthii*), no successful *NAT2* Sanger sequencing of the single DNA sample available in this study (CHarriet) was achieved, so that *NAT2* polymorphism recording for this individual is based solely on the NGS sequence from Prado-Martinez et al. (2013).

[3] Based on the *Pan paniscus* individual of this study (Bonobo), the bonobos of Prado-Martinez et al. (2013) and the *Pan paniscus* draft assembly reference sequence (PanPan1, May 2012).

[4] Based on the individuals of this study, the gorillas of Prado-Martinez et al. (2013) and the *Gorilla gorilla gorilla* draft assembly reference sequence (gorGor4, December 2014).

[5] Based on the individuals of this study, the orangutans of Prado-Martinez et al. (2013) and the *Pongo pygmaeus abelii* draft assembly reference sequence (ponAbe2, July 2007).

[6] At this position (18'257'858), a fixed non-synonymous mutation (A) differentiates the *Pan* genus from the other hominids. Among the latter, a synonymous C/T polymorphism is observed in humans and gorillas, whereas a T is observed in the orangutans.

[7] At this position (18'258'316), a synonymous A/G polymorphism is observed in humans, whereas all other hominids are apparently fixed on A; an A is also observed in the Neanderthal sequences; in Denisova there seems to be an inconsistency at this position between the high-coverage sequence reads of the Denisova genome in the UCSC Genome Browser (reporting A) and the ancient genome browser from the Max Planck Institute for Evolutionary Anthropology, which apparently reports G (see Supplementary Table S8).

[8] At this position (18'258'351), a non-synonymous G/A polymorphism is observed in humans, and a non-synonymous mutation (C) differentiates the orangutans from all other hominids (G).

[9] At this position (18'258'407), immediately downstream the coding exon, we hypothesize that a C insertion (C/CC) polymorphism could be present in gorillas. The four gorillas Sanger sequenced in this study have the C insertion and so does the gorGor3 reference sequence (gorGor3, May 2011). However, the newer gorGor4 reference sequence does not have this insertion, and it is not reported, either, in the gorilla NGS sequences of the GAGP from Prado-Martinez et al. (2013).

References

Mortensen HM, Froment A, Lema G, Bodo JM, Ibrahim M, Nyambo TB, Omar SA, and Tishkoff SA. 2011. Characterization of genetic variation and natural selection at the arylamine N-acetyltransferase genes in global human populations. Pharmacogenomics 12(11):1545-1558.

Patin E, Barreiro LB, Sabeti PC, Austerlitz F, Luca F, Sajantila A, Behar DM, Semino O, Sakuntabhai A, Guiso N et al. . 2006. Deciphering the ancient and complex evolutionary history of human arylamine N-acetyltransferase genes. American journal of human genetics 78(3):423-436.

Podgorna E, Diallo I, Vangenot C, Sanchez-Mazas A, Sabbagh A, Cerny V, and Poloni ES. 2015. Variation in NAT2 acetylation phenotypes is associated with differences in food-producing subsistence modes and ecoregions in Africa. BMC Evol Biol 15:263.

Prado-Martinez J, Sudmant PH, Kidd JM, Li H, Kelley JL, Lorente-Galdos B, Veeramah KR, Woerner AE, O'Connor TD, Santpere G et al. . 2013. Great ape genetic diversity and population history. Nature 499(7459):471-475.

Sabbagh A, Langaney A, Darlu P, Gerard N, Krishnamoorthy R, and Poloni ES. 2008. Worldwide distribution of NAT2 diversity: implications for NAT2 evolutionary history. BMC Genet 9:21.

The Genomes Project C. 2012. An integrated map of genetic variation from 1,092 human genomes. Nature 491:56.

Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, Cummins C, Clapham P, Fitzgerald S, Gil L et al. . 2016. Ensembl 2016. Nucleic Acids Res 44(D1):D710-D716.