

Supplementary Materials for

Divergent Ah receptor ligand selectivity during hominin evolution

Troy D. Hubbard, Iain A. Murray, William H. Bisson, Alexis P. Sullivan, Aswathy Sebastian, George H. Perry, Nina G. Jablonski and Gary H. Perdew*

*Corresponding author. E-mail: ghp2@psu.edu

This PDF file includes:

Supplementary tables 1 and 2

Supplementary figure 1, 2, 3, 4 and 5.

Supplementary Table 1. Exome SNPs fixed for a derived allele in humans and the ancestral allele in Neandertals and Denisovans

Synonymous SNPs				Nonsynonymous SNPs							
Gene	Chr:Pos	RS ID	Ref/Alt Alleles	Gene	Chr:Pos	RS ID	Ref/Alt Alleles	dbSNP	MAF	dbSN P %	
FHL3	chr1:38464668	.	A/G	TTC22	chr1:55248120	rs61733130	T/C	C	0.02	14	2.84
RLF	chr1:40702570	rs16827078	C/T	BCAR3	chr1:94048153	rs12062278	T/C	C	0.05	26	5.31
SLC2A1	chr1:43394612	rs2228490	T/C	SPAG17	chr1:118558632	.	T/C
MUTYH	chr1:45796899	rs74318065	C/G	SPAG17	chr1:118634297	.	A/C
TTC4	chr1:55188382	rs11550481	C/T	RPTN	chr1:152128838	rs75497821	T/A	A	0.01	79	1.58
TLL7	chr1:84387062	.	T/C	SLC27A3	chr1:153751869	.	G/T
BCAR3	chr1:94032969	rs34294233	C/G	CR1L	chr1:207868022	rs72468038	G/T	T	0.03	19	3.95
VCAM1	chr1:101196855	.	C/T	USH2A	chr1:215808022	rs41308435	T/C	C	0.02	12	2.44
SPAG17	chr1:118558666	.	G/A	USH2A	chr1:215821939	rs41315587	C/T	T	0.03	18	3.67
ARHGEF2	chr1:155934849	.	A/G	ABHD1	chr2:27351355	rs34127901	C/A	A	0.01	67	1.34
FCRL3	chr1:157665381	.	A/G	MOGS	chr2:74689671	rs34838944	G/C	C	0.00	48	0.96
CD1D	chr1:158153772	.	G/A	TTC31	chr2:74719816	rs111415294	G/A	A	0.00	46	0.92
AIM2	chr1:159043041	rs34479821	C/T	COL6A3	chr2:238283472	rs11896521	T/G	G	0.07	37	7.49
KCTD3	chr1:215792535	rs75651965	C/G	ZNF502	chr3:44762392	rs6798400	T/C	C	0.03	15	3.12
USH2A	chr1:215821971	rs41304083	G/A	KIF15	chr3:44843440	rs72875737	A/G	G	0.01	78	1.56
HLX	chr1:221054648	rs146391223	A/C	CDCP1	chr3:45130609	rs35428731	G/A	A	0.02	14	2.86
RAD51A	chr2:17699275	rs114644105	G/T	DNAH1	chr3:52381865	rs61734653	A/G	G	0.00	39	0.78
PAIP2B	chr2:71429687	rs13001756	C/G	DNAH1	chr3:52417991	rs61749019	G/A	A	0.00	39	0.78
MRPS9	chr2:105705507	rs11538198	C/T	CD86	chr3:121825197	rs2681417	G/A	G	0.13	66	13.2
TTN	chr2:179500790	rs55847232	A/G	TMCC1	chr3:129546729	rs784689	T/C	T	0.03	17	3.47
UNC80	chr2:210705279	rs116304707	G/A	SLC34A2	chr4:25673231	rs112461275	T/G	G	0.01	58	1.16
UNC80	chr2:210818971	rs16843977	G/A	GPRIN3	chr4:90169025	rs17015286	C/T	T	0.05	29	5.97
FARSB	chr2:223488419	rs143154293	G/T	FGB	chr4:155491759	rs4220	G/A	A	0.15	76	15.34
IQCA1	chr2:237396748	rs61754936	A/C	DMXL1	chr5:118485238	rs112912947	G/C	C	0.02	10	2.18
PPARG	chr3:12475497	rs41516544	A/G	PPIC	chr5:122364538	rs34341374	T/C	C	0.03	15	3.06
Synonymous SNPs				Nonsynonymous SNPs							
Gene	Chr:Pos	RS ID	Ref/Alt Alleles	Gene	Chr:Pos	RS ID	Ref/Alt Alleles	dbSNP	MAF	dbSN P %	
ZNF35	chr3:44700845	rs115467258	G/A	ZNF292	chr6:87971068	rs114321262	T/C	C	0.00	28	0.56
QARS	chr3:49136666	rs113144521	T/C	ZBTB24	chr6:109802500	.	C/G

PRKCD	chr3:53223 927	rs10758 65	C/G	PLEK	chr6:15116 1086	rs17080 410	T/C	C	0.04	20	4.05
IMPG2	chr3:10096 3267	rs35648 234	A/G	HG1	chr6:15939 8700	rs37569 87	C/T	=	0.21	10	21.25
TNIK	chr3:17084 3782	rs13074 171	C/A	RSPH3	chr7:74930 97	rs17167 927	C/G	T	0.06	32	6.47
POLN	chr4:21299 05	.	C/T	COL28A1	.	.	T/C	G	0.06	32	6.47
AFAP1	chr4:78736 95	rs62636 597	T/C	AHR	chr7:17375 392	.	T/C	=	0.04	20	4.05
BOD1L	chr4:13603 037	.	T/C	DNAH11	chr7:21892 230	.	T/C	T	0.06	34	6.89
LAP3	chr4:17609 125	rs62641 702	A/G	SCRN1	chr7:29980 471	rs35960 711	C/T	=	0.07	39	7.83
UTP3	chr4:71555 705	.	A/G	PCLO	chr7:82583 388	rs10487 647	A/G	G	0.07	39	7.83
MTTP	chr4:10051 8283	rs17533 489	T/C	CALD1	chr7:13464 2991	rs11263 3425	A/G	G	0.00	21	0.42
INTS12	chr4:10660 4364	rs34072 732	A/G	EPHB6	chr7:14256 3253	rs81771 46	T/G	G	0.06	34	6.83
DCHS2	chr4:15518 0853	rs13957 8314	A/T	SH2D4A	chr8:19221 726	rs11160 0003	G/A	A	0.00	9	0.18
FGB	chr4:15549 0832	rs4681	C/T	GSR	chr8:30557 599	rs81909 76	C/T	T	0.00	10	0.2
CTNND2	chr5:10981 921	rs13180 095	A/T	ADAM18	chr8:39564 352	.	G/A	=	0.00	2	
NUP155	chr5:37341 232	.	C/T	RB1C1	chr8:53568 742	.	C/T	A	0.15	75	15.06
C5orf30	chr5:10261 2166	rs17155 192	T/C	SPAG1	chr8:10120 6391	rs17335 870	G/A	=	0.12	62	12.48
HSD17B4	chr5:11887 7677	rs12714	C/T	ANKS6	chr9:10153 3220	rs64158 47	C/T	C	0.12	62	12.48
FBXW11	chr5:17130 3477	rs10052 973	A/G	NEK6	chr9:12711 3155	.	C/G	=	0.00	22	0.44
DSP	chr6:75857 63	.	A/C	NR6A1	chr9:12730 0473	.	T/G	T	0.00	22	0.44
TFAP2D	chr6:50697 021	.	T/C	TTF1	chr9:13527 7523	rs14734 1865	C/T	=	0.00	44	
PHIP	chr6:79707 184	rs19038 2214	C/A	C9orf86	chr9:13973 4151	.	C/G	C	0.00	5	0.1
GRIK2	chr6:10213 4102	rs14011 4474	C/T	FBXW5	chr9:13983 6554	rs14996 7967	T/C	=	0.00	1	
SYNE1	chr6:15258 9205	rs11600 7471	G/A	ARRD1	chr9:14050 7366	.	T/C	A	0.01	71	1.42
RSPH3	chr6:15939 8633	rs37569 88	T/C	ASCC1	chr10:7388 7894	rs79032 469	T/A	=	0.01	42	
MACC1	chr7:20197 911	rs11708 4218	C/T	FAM178A	chr10:1026 76434	.	G/A	C	0.01	74	1.48

Synonymous SNPs

Nonsynonymous SNPs

Gene	Chr:Pos	RS ID	Ref/Alt Alleles	Gene	Chr:Pos	RS ID	Ref/Alt Alleles	dbSNP MAF	dbSNP P %		
GLI3	chr7:42004 600	rs34089 404	G/A	DCHS1	chr11:6654 769	.	C/T				
HGF	chr7:81388 042	rs57456 35	T/C	OR4C16	chr11:5534 0170	rs61916 568	A/C	C	0.02	12	2.4
PCLO	chr7:82595 324	rs99693 58	C/T	OR10	chr11:5573 5518	rs61896 165	G/A	A	0.03	19	3.81
KPNA7	chr7:98790 669	rs77836 06	G/C	AG1	chr11:5590 5025	rs19479 24	T/G	G	0.06	30	6.03
TSC22D4	chr7:10007 2031	rs78098 01	T/C	OR8J3	chr11:5590 5025	rs19479 24	T/G	G	0.06	30	6.03
SMO	chr7:12885 0875	rs20166 07	T/C	TECTA	chr11:1209 73430	rs35507 522	A/G	G	0.04	20	4.13
SLC37A3	chr7:14006 9426	rs11311 0951	G/T	CCDC15	chr11:1248 57495	.	A/C	=	0.04	13	7
SLC25A37	chr8:23429 329	rs10473 84	T/C	PRDM10	chr11:1297 72293	.	T/G	T	0.00	22	0.44
ADAMD1	chr8:24256 512	rs70052 58	T/C	EFCA	chr12:3747 472	rs11749 2949	C/T	G	0.15	78	15.63
EC1				B4B	chr12:6678 6091	rs79703 87	G/C	=	0.15	63	3

ADAMD EC1	chr8:24261 518	rs61731 545	G/A	PTPR B	chr12:7101 6340	rs72981 47	C/T	T =	0.01 98	99	1.98
INTS9	chr8:28625 777	.	C/T	SLITR K1	chr13:8445 4655	rs14562 8951	A/C	C =	0.00 24	12	0.24
EFCAB1	chr8:49647 669	.	G/A	NOVA 1	chr14:2691 8100	.	C/T				
RB1CC 1	chr8:53569 194	rs16918 075	C/T	TTLL5	chr14:7624 9759	rs14867 7697	A/G	G =	0.00 62	31	0.62
PUF60	chr8:14489 8801	.	A/G	GPR1 32	chr14:1055 17492	.	C/G				
NFKBIL 2	chr8:14566 5516	rs23063 83	T/C	C15orf 55	chr15:3464 6769	rs61737 334	G/A	A =	0.02 3	11 5	2.3
KIAA15 39	chr9:35106 315	rs18213 4491	G/A	KIAA1 199	chr15:8117 3308	.	G/A				
AMBP	chr9:11682 3774	rs11210 9765	G/A	CDH1 6	chr16:6694 7064	.	T/C				
PAPPA	chr9:11913 0030	rs38276 68	A/G	GGT6	chr17:4463 716	rs62066 363	G/A	A =	0.01 3	65	1.3
TLN1	chr9:35706 519	.	G/T	DNAH 2	chr17:7721 315	rs61749 477	C/A	A =	0.00 7	35	0.7
NET1	chr10:5498 223	rs77152 706	T/C	SPAG 5	chr17:2692 5578	.	A/G				
FRMD4 A	chr10:1373 5947	rs22965 96	A/G	SSH2	chr17:2795 9034	.	T/C				
FAM171 A1	chr10:1525 5193	.	A/G	BCAS 3	chr17:5878 6681	rs34712 615	A/G	G =	0.08 31	41 6	8.31
MPP7	chr10:2834 5418	rs75802 203	T/C	AZI1	chr17:7917 6142	rs11168 1839	A/G	G =	0.02 1	10 5	2.1
ERCC6	chr10:5073 2681	.	A/G	RFNG	chr17:8000 6980	.	G/A				
P4HA1	chr10:7476 7998	rs34162 681	C/G	KIAA1 772	chr18:1908 5379	.	A/G				
LDB3	chr10:8844 6985	rs76615 432	T/C	DSG4	chr18:2898 3565	rs72292 52	T/C	C =	0.02 28	11 4	2.28

Synonymous SNPs

Nonsynonymous SNPs

Gene	Chr:Pos	RS ID	Ref/Alt Alleles	Gene	Chr:Pos	RS ID	Ref/Alt Alleles	dbSNP	MAF	dbSN P %	
CRTAC 1	chr10:9966 4456	rs70685 03	T/C	C19orf 28	chr19:3547 315	.	C/G				
PYROX D2	chr10:1001 52825	.	G/A	PPAN	chr19:1022 1642	rs11559 188	A/G	G =	0.06 75	33 8	6.75
ADRA2 A	chr10:1128 39026	.	A/G	PRKC SH	chr19:1155 9907	rs34351 170	A/G	G =	0.00 78	39	0.78
TDRD1	chr10:1159 50727	rs75742 995	A/G	AKAP8	chr19:1548 0981	rs34063 092	G/A	A =	0.02 28	11 4	2.28
PNLIP	chr10:1183 14919	rs11246 4025	C/T	NPHS 1	chr19:3633 0277	rs34736 717	C/G	G =	0.04 77	23 9	4.77
FXC1	chr11:6502 780	rs11555 935	C/A	C5AR1	chr19:4782 3038	rs44671 85	G/A	G =	0.03 06	15 3	3.06
ST5	chr11:8751 577	.	A/G	C5AR1	chr19:4782 3871	rs11880 097	G/T	G =	0.03 06	15 3	3.06
OR5AS 1	chr11:5579 8533	rs17528 803	A/C	TRPM 4	chr19:4967 1207	rs11398 4787	G/A	A =	0.01 6	80	1.6
OR8J3	chr11:5590 4751	rs19479 23	T/C	RSPH 1	chr21:4389 7491	.	G/T				
NAALA DL1	chr11:6481 3918	.	C/T	OSBP 2	chr22:3128 9891	rs34240 867	A/G	G =	0.00 6	30	0.6
NAALA DL1	chr11:6481 4965	.	A/G	ADSL	chr22:4076 0978	.	T/C				
FOXMI	chr12:2983 636	rs28990 692	A/G	PPP1 R3F	chrX:49137 916	rs17148 347	T/C	C =	0.09 01	34 0	9.01
PTPN6	chr12:7061 221	.	T/C	RPGR	chrX:38182 130	rs11163 1988	T/C	C =	0.02 54	96	2.54
ITGA5	chr12:5479 5410	rs73322 228	A/G								
BAZ2A	chr12:5700 3964	rs22305 80	T/A								
MYL2	chr12:1113 53556	rs23016 10	A/G								

SFRS8	chr12:1322	rs34754	T/C
	10063	096	
NARG1	chr13:4190	rs38128	T/C
L	2939	96	
DIAPH3	chr13:6059	.	T/C
	0255		
TEP1	chr14:2084	rs22280	G/A
	9169	39	
DHRS2	chr14:2410	rs11618	A/G
	9065	9534	
DHRS2	chr14:2411	rs11204	C/T
	4054	4307	
LRRC16	chr14:2453	rs11260	C/A
B	4474	6230	
TINF2	chr14:2470	rs10141	C/T
	9351	326	
SSTR1	chr14:3867	.	G/C
	9302		
FERMT	chr14:5341	rs61741	G/C
2	7158	989	

Synonymous SNPs

Nonsynonymous SNPs

Gene	Chr:Pos	RS ID	Ref/Alt Alleles	Gene	Chr:Pos	RS ID	Ref/Alt Alleles	dbSNP MAF	dbSNP %
GALNTL1	chr14:6979	.	C/T						
	2114								
KIAA1199	chr15:8117	rs35541	G/A						
	2064	581							
TM6SF1	chr15:8378	.	C/T						
	8350								
ACAN	chr15:8938	.	A/C						
	2027								
ANPEP	chr15:9034	rs11354	A/T						
	8682	7384							
UNC45A	chr15:9148	rs80345	C/T						
	6305	51							
CDH16	chr16:6694	.	C/T						
	6646								
ZNF469	chr16:8849	rs74384	G/C						
	6452	633							
MYO1C	chr17:1386	rs61756	C/G						
	915	682							
MYBBP1A	chr17:4455	rs14500	T/C						
	763	6324							
DNAH2	chr17:7720	rs78005	C/T						
	920	021							
MSL1	chr17:3828	rs28526	C/T						
	2624	298							
CCR7	chr17:3871	rs22290	T/C						
	1222	95							
SPOP	chr17:4767	rs11380	G/A						
	9280	8176							
KRT14	chr17:3974	rs75795	T/C						
	0629	684							
DAZAP1	chr19:1430	rs11377	T/C						
	303	8676							
FEM1A	chr19:4792	rs73919	A/G						
	697	322							
PRKCSH	chr19:1155	rs13914	T/C						
	9801	4555							
MAN2B1	chr19:1277	.	A/G						
	4194								
KIAA0892	chr19:1945	.	G/A						
	6169								
ATP1A3	chr19:4247	rs61733	A/G						
	4639	017							
FGF21	chr19:4926	.	C/A						
	0260								
MUC16	chr19:8982	rs78717	A/G						
	212	907							

HOMER 3	chr19:1904 2387	rs57792 670	A/G
JAG1	chr20:1063 7057	rs10485 741	T/C
NCOA6	chr20:3332 9689	.	C/T

Synonymous SNPs

Gene	Chr:Pos	RS ID	Ref/Alt Alleles
C20orf1 18	chr20:3551 7690	.	C/T
SLC12A 5	chr20:4466 5946	rs37465 19	T/C
AGPAT 3	chr21:4538 7870	.	A/G

Nonsynonymous SNPs

Gene	Chr:Pos	RS ID	Ref/Alt Alleles	dbSNP MAF	dbSNP %
------	---------	-------	-----------------	-----------	---------

Supplementary Table 2. Exome SNPs fixed for a derived allele in humans (and not variable in dbSNP) and the ancestral allele in Neandertals and Denisovans

Synonymous SNPs			Nonsynonymous SNPs		
Gene	Chr:Pos (1)	Ref/Alt Bases (2)	Gene	Chr:Pos	Ref/Alt Bases
FHL3	chr1:38464668	A/G	SPAG17	chr1:118558632	T/C
TLL7	chr1:84387062	T/C	SPAG17	chr1:118634297	A/C
VCAM1	chr1:101196855	C/T	SLC27A3	chr1:153751869	G/T
SPAG17	chr1:118558666	G/A	ZBTB24	chr6:109802500	C/G
ARHGEF2	chr1:155934849	A/G	AHR	chr7:17375392	T/C
FCRL3	chr1:157665381	A/G	DNAH11	chr7:21892230	T/C
CD1D	chr1:158153772	G/A	ADAM18	chr8:39564352	G/A
POLN	chr4:2129905	C/T	RB1CC1	chr8:53568742	C/T
BOD1L	chr4:13603037	T/C	NEK6	chr9:127113155	C/G
UTP3	chr4:71555705	A/G	NR6A1	chr9:127300473	T/G
NUP155	chr5:37341232	C/T	C9orf86	chr9:139734151	C/G
DSP	chr6:7585763	A/C	ARRDC1	chr9:140507366	T/C
TFAP2D	chr6:50697021	T/C	FAM178A	chr10:102676434	G/A
INTS9	chr8:28625777	C/T	DCHS1	chr11:6654769	C/T
EFCAB1	chr8:49647669	G/A	CCDC15	chr11:124857495	A/C
PUF60	chr8:144898801	A/G	PRDM10	chr11:129772293	T/G
TLN1	chr9:35706519	G/T	NOVA1	chr14:26918100	C/T
FAM171A1	chr10:15255193	A/G	GPR132	chr14:105517492	C/G
ERCC6	chr10:50732681	A/G	KIAA1199	chr15:81173308	G/A
PYROXD2	chr10:100152825	G/A	CDH16	chr16:66947064	T/C
ADRA2A	chr10:112839026	A/G	SPAG5	chr17:26925578	A/G
ST5	chr11:8751577	A/G	SSH2	chr17:27959034	T/C
NAALADL1	chr11:64813918	C/T	RFNG	chr17:80006980	G/A
NAALADL1	chr11:64814965	A/G	KIAA1772	chr18:19085379	A/G
Synonymous SNPs			Nonsynonymous SNPs		
Gene	Chr:Pos (1)	Ref/Alt Bases (2)	Gene	Chr:Pos	Ref/Alt Bases
PTPN6	chr12:7061221	T/C	C19orf28	chr19:3547315	C/G
DIAPH3	chr13:60590255	T/C	RSPH1	chr21:43897491	G/T
SSTR1	chr14:38679302	G/C	ADSL	chr22:40760978	T/C
GALNTL1	chr14:69792114	C/T			
TM6SF1	chr15:83788350	C/T			
ACAN	chr15:89382027	A/C			
CDH16	chr16:66946646	C/T			
MAN2B1	chr19:12774194	A/G			
KIAA0892	chr19:19456169	G/A			
FGF21	chr19:49260260	C/A			
NCOA6	chr20:33329689	C/T			
C20orf118	chr20:35517690	C/T			

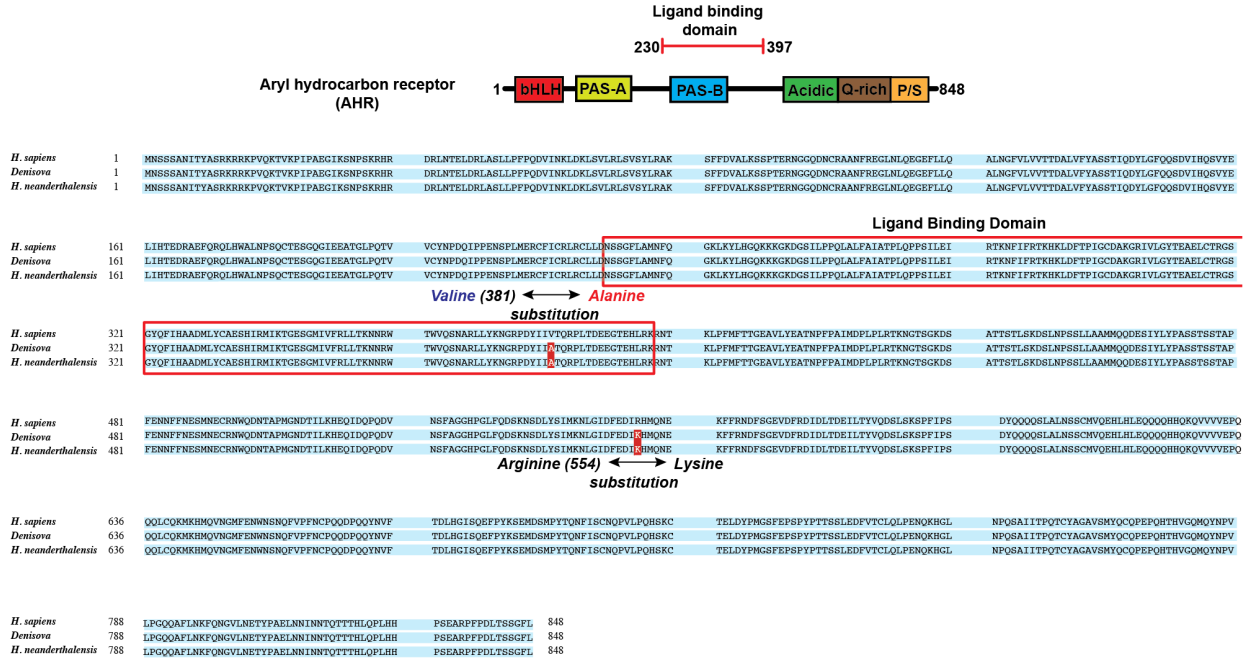
AGPAT3

chr21:45387870

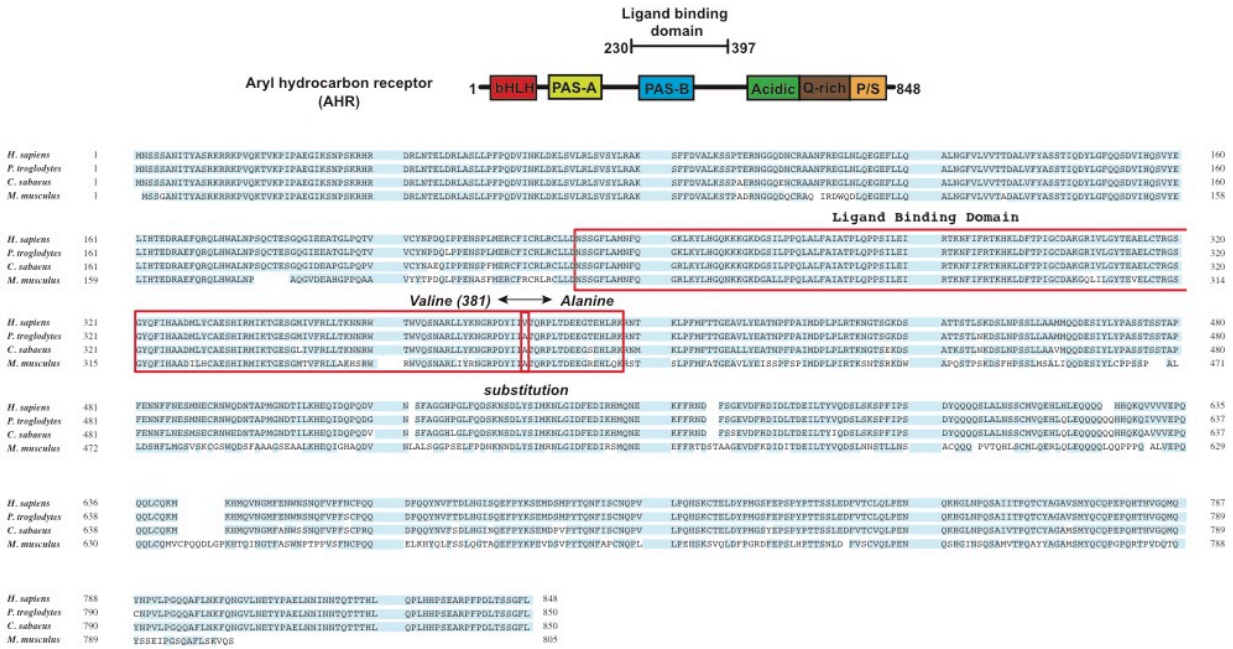
A/G

1: Chromosome Number:Reference Base Position in Chromosome

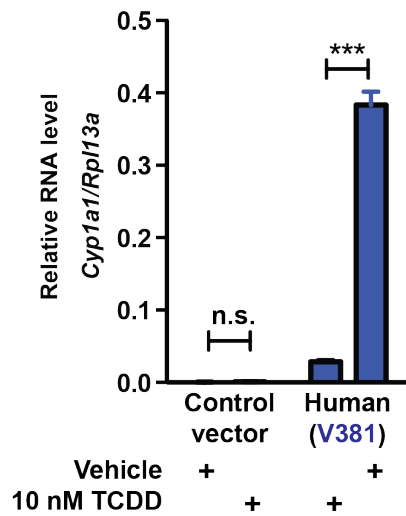
2: Reference/Alternate Bases at specified chromosome position



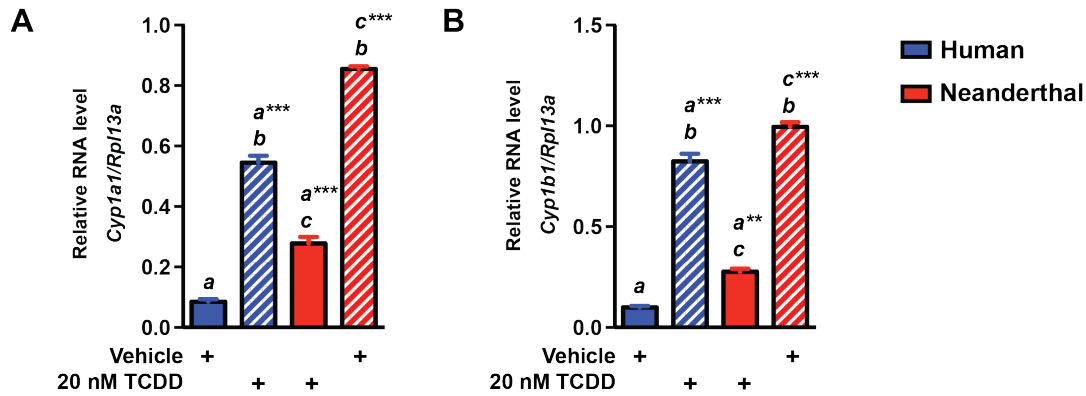
Supplementary fig. 1. Schematic representation of *H. sapiens* AHR domain organization highlighting the position of the ligand binding domain (residues 230-397). Amino acid sequences for hominin (*H. sapiens*, *H. neanderthalensis* and Denisova) AHR were aligned using the NCBI protein-protein blast (blastp) algorithm. AHR ligand binding domain is indicated by red border. Non-synonymous amino acid variants A ↔ V381 and R ↔ K554 are highlighted in red.



Supplementary fig. 2. Schematic representation of *H. sapiens* AHR domain organization highlighting the position of the ligand binding domain (residues 230-397). Amino acid sequences for hominin (*H. sapiens*, *P. troglodytes*, *C. sabaeus* and *M. musculus*) AHR were aligned using the NCBI protein-protein blast (blastp) algorithm. AHR ligand binding domain is indicated by red border. Non-synonymous amino acid variants A↔V381 is highlighted in red.



Supplementary fig. 3. TCDD fails to induce *Cyp1a1* in BP8 cells in the absence of ectopic AHR expression. BP8 cells were transfected with either a control or human AHR expressing plasmid, after 24 h cells were treated with TCDD for 4 h. *Cyp1a1* mRNA levels were measured by qRT-PCR. Data were analyzed using one-way ANOVA followed by Tukey's multiple-comparisons test. ***< 0.001.



Supplementary fig. 4. At a saturating dose of TCDD the ectopic expression of the Neanderthal and human AHR in BP8 cells results in similar levels of *Cyp1a1* and *Cyp1b1*. BP8 cells were transfected with either a Neanderthal or human AHR expressing plasmid, after 24 h cells were treated with TCDD for 4 h. *Cyp1a1* mRNA levels were measured by qRT-PCR. Data were analyzed using one-way ANOVA followed by Tukey's multiple-comparisons test. *** $P < 0.001$, ** $P < 0.01$. Alphabetical characters indicate statistical comparisons between groups.

Codon optimized green monkey AHR cDNA sequence

ATGAACTCCAGCTCCGCCAACATTACTTACGCCAGCAGGAAAAGAAGAAAGCCCGTCCAGAAGACAGTGAAGCCA
ATCCCAGCCGAAGGCATCAAATCCAATCCTTCTAAGCGGCACAGAGACAGGCTGAACACAGAGCTGGATCGACTG
GCATCCCTGCTGCCATTCCCTCAGGACGTGATTAACAAACTGGATAAGCTGTTCAGTGTGAGACTGTCCGTCTCT
TACCTGAGGGCAAAAAGCTTCTTTGACGTGGCCCTGAAGAGCTCCCCAGCAGAGCGAAAACGGAGGACAGGAAAAT
TGCAGGGCCGCTAACTTTTCGCGAAGGACTGAATCTGCAGGAGGGCGAATTCCTGCTGCAGGCTCTGAACGGGTTT
GTGCTGGTGGTCACCACAGACGCTCTGGTCTTCTACGCATCTAGTACCATCCAGGACTATCTGGGATTTTCAGCAG
TCAGATGTGATCCACCAGAGCGTCTATGAGCTGATTTCATACCCGAGGATCGAGCCGAATTCAGCGGCAGCTGCAT
TGGGCTCTGAATCCCTCTCAGTGCACAGAGAGTGGGCAGGGAATCGATGAAGCCCTGGACTGCCTCAGCCAGTG
GTCTGTTACAATGCTGAGCAGATCCCCCTGAGAACAGCCCTTTTCATGGAAAGGTGCTTTATTTGTGCGCTGCGA
TGCCTGCTGGATAAATCAAGCGGCTTCTGGCCATGAACTTTTCAGGGGCGACTGAAGTATCTGCACGGACAGAAG
AAAAAGGGAAAGGACGGCTCTATCCTGCCACCCCAGCTGGCTCTGTTTCGCCATTGCTACACCTCTGCAGCCTCCA
AGCATCCTGGAGATTTCGCACAAAGAACTTCATCTTTTCGAACTAAGCATAAGCTGGACTTTACCCCAATCGGATGC
GATGCAAAGGGCAGAATTGTGCTGGGGTACACCGAGGCCGAATGTGCACAAGGGGCTCTGGCTACCAGTTCATC
CACGCAGCCGATATGCTGTATTGTGCCGAGAGCCATATCAGAATGATCAAGACCCGGGAATCCGGACTGATCGTG
TTCCGGCTGCTGACTAAGAACAATAGATGGACCTGGGTCCAGTCCAATGCTAGGCTGCTGTACAAGAACGGCAGA
CCCGACTATATCATTGCAACACAGAGGCCTCTGACTGATGAGGAAGGGAGCGAGCACCTGCGGAAAAGAAAACATG
AAGCTGCCCTTCATGTTTACTACCGGCGAGGCCCTGCTGTATGAAGCTACCAATCCATTCCCCGCAATCATGGAC
CCTCTGCCACTGCGCACCAAAAACGGCACAAGCGAAAAGGACTCCGCCACAAAGTCTACCCTGAATAAGGATAGT
CTGAACCCATCCTCTCTGCTGGCTGCAGTGCAGCAGGATGAGTCCATCTACCTGTATCCAGCTAGTTCAACT
AGTCCACCGCACCTTCGAAAACAATTTCTGAATGAGAGCATGAGTGAATGCCGAATTTGGGAAGACAACACA
GCTCCCATGGGCAATGATACTATCCTGAAGCACGAGCAGATTGACCAGCCCAGGATGTGAACAGCTTCGCAGGA
GGCCATCTGGGGCTGTTTTCAGGACAGTAAAAATTCAGATCTGTACTCCATCATGAAGAACCTGGGAATCGACTTC
GAGGATATCAAGCACATGCAGAATGAAAAGTTCTTTTCGAAACGACTTCAGCAGTGAGGTGGATTTTTCGGGACATC
GATCTGACAGACGAAATTTCTGACTTACATTCAGGATAGTCTGTCAAAAAGCCCTTCATCCCCTCCGACTATCAG
CAGCAGCAGTCTCTGGCCCTGAATTTCTAGTTGTATGGTGCAGGAGCACCTGCAGCTGGAACAGCAGCAGCAGCAG
CAGCACCATCAGAAGCAGGCAGTGGTTCGAGCCCAGCAGCAGCTGTGCCAGAAAATGAAGCATATGCAAGTGAAT
GGCATGTTTCGCTAACTGGTCCAGTAACCAGTTCGTTCCATTTTCATGTCCAGACAGGACCCTCAGCAGTACAAC
GTGTTCTCAGATCTGCACGGGATCAATCAGGAGTTTTCTTACAAAAGTGAATGGACCCAGTGCCATATACACAG
AACTTCATCTCTTGCAACCAGCCTGTGCTGCCACAGCATAGTAAGTGTACTGAGCTGGACTACCCAATGGGGAGC
TATGAACCCCTCCCCTTATCCAACCACATCAAGCCTGGAGGATTTTGTGACTTGCCTGCAGGTCCCCGAAAATCAG
AAACAGGGCTGAACCCACAGTCCGCCATCGTGACACCCCAGACTTGCTACGCAGGAGCCATGTCTATGTATCAG
TGTCAGCCCGGCTCAGCACACCCATGTGGGACAGATGCAGTATAAACCCTGCTGCCTGGCAGCAGGCCTTTC
CTGAATAAGTTTTCAGAACGGCGTGCTGAATGAGACCTACCCTGCTGAACCTGAACAATATCAACAATACAGGCC
ACCACACATCTGCAGCCCCTGCACCATCCAAGCGAAGCAAGACCATTTCCCGACCTGACATCATCTGGGTTCTCTG
TAA

Codon optimized chimpanzee AHR cDNA sequence

ATGAACTCCAGCTCCGCCAACATTACTTACGCCAGCAGGAAAAGAAGAAAGCCCGTCCAGAAGACAGTGAAGCCA
ATCCCAGCCGAAGGCATCAAATCCAATCCTTCTAAGCGGCACAGAGACAGGCTGAACACAGAGCTGGATCGACTG
GCATCCCTGCTGCCATTCCCTCAGGACGTGATTAACAAACTGGATAAGCTGTTCAGTGTGAGACTGTCCGTCTCT
TACCTGAGGGCAAAAAGCTTCTTTGACGTGGCCCTGAAGAGCTCCCCAACCAGAGCGAAAACGGAGGACAGGATAAT
TGCAGGGCCGCTAACTTTTCGCGAAGGACTGAATCTGCAGGAGGGCGAATTCCTGCTGCAGGCTCTGAACGGGTTT
GTGCTGGTGGTCACCACAGACGCTCTGGTCTTCTACGCATCTAGTACCATCCAGGACTATCTGGGATTTTCAGCAG
TCAGATGTGATCCACCAGAGCGTCTATGAGCTGATTTCATACCCGAGGATCGAGCCGAATTCAGCGGCAGCTGCAT
TGGGCTCTGAATCCCTCTCAGTGCACAGAGAGTGGGCAGGGAATCGAGGAAGCCACTGGACTGCCTCAGACCGTG
GTCTGTTACAATCCAGACCAGCTGCCCCCTGAGAACAGCCCTCTGATGGAAAGGTGCTTTATTTGTGCGCTGCGA
TGCCTGCTGGATAAATCAAGCGGCTTCTGGCCATGAACTTTTCAGGGGAAACTGAAGTATCTGCACGGACAGAAG
AAAAAGGGAAAGGACGGCTCTATCCTGCCACCCCAGCTGGCTCTGTTTCGCCATTGCTACACCTCTGCAGCCTCCA
AGCATCCTGGAGATTTCGCACAAAGAACTTCATCTTTTCGAACTAAGCATAAGCTGGACTTTACCCCAATCGGATGC
GATGCAAAGGGCAGAATTGTGCTGGGGTACACCGAGGCCGAATGTGCACAAGGGGCTCTGGCTACCAGTTCATC
CACGCAGCCGATATGCTGTATTGTGCCGAGAGCCATATCAGAATGATCAAGACCCGGGAATCCGGAAATGATCGTG
TTCCGGCTGCTGACTAAGAACAATAGATGGACCTGGGTCCAGTCCAATGCTAGGCTGCTGTACAAGAACGGCAGA
CCCGACTATATCATTGCAACACAGAGGCCTCTGACTGATGAGGAAGGGACCGAGCACCTGCGGAAAAGAAAACACT
AAGCTGCCCTTCATGTTTACTACCGGCGAGGCCGTGCTGTATGAAGCTACCAATCCATTCCCCGCAATCATGGAC

CCTCTGCCACTGCGCACCAAAAACGGCACAAGCGGGAAGGACTCCGCCACAACCTTCTACCCTGAATAAGGATAGT
CTGAACCCATCCTCTCTGCTGGCTGCAATGATGCAGCAGGATGAGTCCATCTACCTGTATCCAGCTAGTTCAACT
AGCTCCACCGCACCTTCGAAAACAATTTCTTTAATGAGAGCATGAACGAATGCCGGAATTGGCAGGACAACACA
GCTCCCATGGGCAATGATACTATCCTGAAGCACGAGCAGATTGACCAGCCCAGGATGGCAACAGCTTCGCAGGA
GGCCATCCTGGGCTGTTTTAGGACAGTAAAAATTCAGATCTGTACTCCATCATGAAGAACCTGGGAATCGACTTC
GAGGATATCAAGCACATGCAGAATGAAAAGTTCTTTGAAAACGACTTCAGCGGGCAGGTTGGATTTTCGGGACATC
GATCTGACAGACGAAATTCGACTTACGTCCAGGATAGTCTGTCAAAAAGCCCCTTCATCCCCTCCGACTATCAG
CAGCAGCAGTCTCTGGCCCTGAATTTCTAGTTGTATGGTGCAGGAGCACCTGCAGCTGGAACAGCAGCAGCAGCAG
CAGCACCATCAGAAGCAGATTGTGGTCGAGCCCCAGCAGCAGCTGTGCCAGAAAATGAAGCATATGCAAGTGAAT
GGCATGTTGAAAACCTGGAATAGTAACCAGTTCGTCCCATTTTCATGTCCCAGCAGGACCCTCAGCAGTACAAC
GTGTTACCGATCTGCACGGGATCAGCCAGGAGTTTTCTTACAAAAGTGAATGGACTCAATGCCATATACACAG
AACTTCATCTCTTGCAACCAGCCTGTGCTGCCACAGCATAGTAAGTGTACTGAGCTGGACTACCCAATGGGGAGC
TTCGAACCCTCCCCTTATCCAACCACATCAAGCCTGGAGGATTTTGTGACTTGCCTGCAGCTGCCCGAAAATCAG
AAACACGGGCTGAACCCACAGTCCGCCATCATTACACCCCAGACTTGCTACGCAGGAGCCGTGTCTATGTATCAG
TGTCAGCCCAGCCTCAGCACACCCATGTGGGACAGATGCAGTGTAAACCCGTCCTGCCTGGACAGCAGGCCTTC
CTGAATAAGTTTTCAGAACGGCGTGCTGAATGAGACCTACCCTGTGAACTGAACAATATCAACAATACACAGACT
ACCACACATCTGCAGCCCCTGCACCATCCAAGCGAAGCAAGACCATTTCCCGACCTGACATCATCTGGGTTCTCTG
TAA

Codon optimized human AHR cDNA sequence

ATGAACAGCAGCAGCGCTAACATCACCTACGCTAGCCGAAAACGGCGGAAGCCCGTGCAGAAAACCGTGAAGCCC
ATCCCCGCTGAAGGCATCAAAAGCAACCCCAGCAAACGGCACCGGGATCGGCTGAACACCGAACTGGATCGGCTG
GCTAGCCTGCTCCCCTTCCCCCAGGATGTGATCAACAAGCTGGATAAACTGAGCGTGCTGCGGCTCAGCGTGAGC
TACCTGCGGGCCAAGAGCTTCTTCGATGTGGCTCTGAAAAGCAGCCCCACCGAACGGAACGGCGGCCAGGATAAC
TGCCGGGCTGCTAACTTCCGGGAAGGCCTGAACCTGCAGGAAGGCGAATTCCTGCTCCAGGCTCTGAACGGCTTC
GTGCTGGTGGTGACCACCGATGCTCTGGTGTCTACGCTAGCAGCACCATCCAGGATTACCTGGGCTTCCAGCAG
AGCGATGTGATCCACCAGAGCGTGTACGAACTGATCCACACCGAAGATCGGGCTGAATTCAGCGGCAGCTGCAC
TGGGCTCTGAACCCAGCCAGTGCACCGAAAAGCGGCCAGGGCATCGAAGAAGCTACCGGCCTCCCCAGACCGTG
GTGTGCTACAACCCGATCAGATCCCCCCCCGAAAACAGCCCCCTGATGGAACGGTGCTTCATCTGCCGGCTGCGG
TGCCTGCTCGATAACAGCAGCGGCTTCTGGCCATGAACTTCAGGGCAAGCTGAAGTACCTGCACGGCCAGAAA
AAGAAAGGCAAGATGGCAGCATCCTGCCCCCCCAGCTGGCTCTGTTCGCTATCGCTACCCCCCTGCAGCCCCC
AGCATCTCGAAAATCCGGACCAAAAACCTTCACTTCCGGACCAACAAAACAACTGGACTTCACCCCCCTCGGCTGC
GATGTAAAGGCGGATCGTGTGGCTACACCAAGCTGAACTGTGCACACGGGGCAGCGGCTACCAGTTCATC
CACGCTGCTGATATGCTGTACTGCGCCGAGAGCCATCCGGATGATCAAGACCGGCGAAAAGCGGCATGATCGTG
TTCCGGCTGCTCACCAAAAACAACCGGTGGACCTGGGTGCAGAGCAACGCTCGGCTGCTCTACAAAACCGCCGG
CCCGATTACATCATCGTGACCCAGCGGCCCTGACCGATGAAGAAGGCACCGAGCACCTGCGGAAAACGGAACACC
AAGCTGCCCTTCATGTTACACCACCGGCGAAGCTGTGCTGTACGAAGCCACCAACCCCTTCCCCGCTATCATGGAT
CCCCTGCCCCTGCGGACCAAAAACGGCACAGCGGCAAAGATAGCGCTACCACCAGCACCTGAGCAAGGATAGC
CTCAACCCAGCAGCCTCCTGGCTGCCATGATGCAGCAGGATGAGAGCATCTACCTCTACCCCGCTAGCAGCACC
AGCAGCACCGCCCCCTTCGAAAACAATTTCTTCAACGAAAGCATGAACGAATGCCGGAACTGGCAGGATAACACC
GCCCCATGGGCAACGATAACCATCCTGAAACACGAGCAGATCGACCAGCCCCAGGATGTGAACAGCTTCGCTGGC
GGCCACCCCGGCCTCTTCCAGGATAGCAAAAACAGCGACCTGTACAGCATCATGAAAAACCTGGGCATCGATTTT
GAAGACATCCGGCACATGCAGAACGAAAATTTCTTCCGGAACGATTTTCAGCGGGCAAGTGGACTTCCGGGACATC
GACCTGACCGATGAAATCCTGACCTACGTGCAGGATAGCTGAGCAAGAGCCCCTTCATCCCCAGCGATTACCAG
CAGCAGCAGAGCCTGGCTCTGAACAGCAGCTGCATGGTGCAGGAACACCTGCACCTGGAACAGCAGCAGCAGCAG
CACCAGAAAACAGGTGGTGGTGAACCCAGCAGCAGCTGTGCCAGAAGATGAAGCACATGCAGGTGAACGGCATG
TTCGAAAACCTGGAACAGCAACCAGTTCGTGCCCTTCAACTGCCCCAGCAGGACCCCGCAGTACAACGTGTTT
ACCGATCTGCACGGCATCAGCCAGGAGTTCCCTTACAAAAGCGAAAATGGATAGCATGCCCTACACCCAGAATTC
ATCAGCTGCAACCAGCCCGTGTGCCCCAGCACAGCAAATGCACCGAGCTGGATTACCCCATGGGCAGCTTCGAA
CCCAGCCCCTACCCACCACCAGCAGCCTGGAAGATTTCTGTGACCTGCCTGCAGCTGCCCGAAAACAGAAAAC
GGCCTGAACCCCCAGAGCGCTATCATCACCCCCAGACCTGCTACGCTGGCGCTGTGAGCATGTACCAGTGCCAG
CCCGAACCCAGCACACCCACGTGGGCCAGATGCAGTACAACCCCGTGTGCCCCGGCCAGCAGGCCTTCCTGAAC
AAGTTCCAGAACGGCGTGCTGAACGAAACCTACCCCGTGAACCTGAACAACATCAACAACACCCAGACCACCACC
CACCTGCAGCCCCTGCACCACCCAGCGAAGCTCGGCCCTTCCCGATCTGACCAGCAGCGGCTTCCTGTGA

Supplementary fig. 5. Codon optimized cDNA sequences for the green monkey, chimpanzee and human AHR.