Online Supporting Information

Analysis of Human *CYP1A1* and *CYP1A2* Genes and Their Shared Bidirectional Promoter in Eight World Populations

Lucia F. Jorge-Nebert, Zhengwen Jiang[†], Ranajit Chakraborty, Joanna Watson, Li Jin[†],
Stephen T. McGarvey[‡], Ranjan Deka, and Daniel W. Nebert*

Department of Environmental Health, and Center for Environmental Genetics (**CEG**)
University of Cincinnati Medical Center, Cincinnati, Ohio 45267-0056

[†]Current address**:** Chinese Academy of Sciences and Max Planck Society Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

[‡]International Health Institute
Brown University, Providence, Rhode Island 02912

*To whom correspondence should be addressed**:**
Daniel W. Nebert
Univ Cincinnati Med Center
Environmental Health
PO Box 670056
Cincinnati
Ohio
45267-0056
United States
Phone: 513-821-4664
Fax: 513-821-4664
Email: dan.nebert@uc.edu

SUPP. TABLE S1. One Hundred Six SNPs across the *CYP1A1_CYP1A2* Locus Genes, Found during Our Studies of Eight World Populations

SNP #	SNP Name	Relative Position	NCBI Assignment	SNP Property	5' Flanking reference sequence
	XRE II 1	CYP1A1 +4732			
1	g.72798247C>T	CYP1A1 +4245	**rs17861080**	3' END	CGTCTCAAAAAAAAAAAGACCTGGAAATCT
2	g.72798250G>A	CYP1A1 +4242	**rs17861081**	3' END	CTCAAAAAAAAAAAGACCTGGAAATCTCGC
3	g.72798264A>G	CYP1A1 +4228	**rs17861082**	3' END	GACCTGGAAATCTCGCGGTGCCCACTCCCC
4	g.72798694A>G	CYP1A1 +3798	**rs4646903**	3 'END	GAGGCTGAGGTGGGAGAATCGTGTGAGCCC
5	g.72799072G>T	CYP1A1 +3420	**rs17861084**	3' UTR	GCTCTTATGCAAGCATGCAAGCTCAATGCA
6	g.72799170C>A	CYP1A1 +3322	**rs17861085**	3' UTR	GGGAGACCAATAGAAGGTAATTGAAATACC
7	g.72799175C>G	CYP1A1 +3317	**rs17861086**	3' UTR	ACCAATAGAAGGTAATTGAAATACCCCCCC
8	g.72799260A>G	CYP1A1 +3232	**rs4986884**	3' UTR	TCCCCGCAGACACAAGTCCCCAGCCCCTCC
	g.72799270T>C	CYP1A1 +3222	**rs4986883**	3' UTR	CACAAGTCCCCAGCCCCTCCAGGACAGCAA
9	g.72799524T>C	CYP1A1 +2968	**rs4986882**	3' UTR	GTTCTAACTTGCTGAAGCCAGTCAGCACCC
10	g.72799540G>C	CYP1A1 +2952	**rs4986881**	3' UTR	GCCAGTCAGCACCCTCACAGAGCCAGCTAG
11	g.72799697G>A	CYP1A1 +2795	**rs4986880**	3' UTR	GTTTTACCTGTTGTCTCTGGAGGGTGTGCA
12	g.72800038T>C	CYP1A1 +2454	**rs1048943**	p.Ile462Val	CCAGGAAGAGAAAGACCTCCCAGCGGGCAA
13	g.72800040G>T	CYP1A1 +2452	**rs1799814**	p.Thr461Asn	AGGAAGAGAAAGACCTCCCAGCGGGCAATG
	XRE II 2	CYP1A1 +1069			
14	g.72801524C>T	CYP1A1 +968	**ss159100334**	INTRON 2	AAAGTTCTATTTCCCTGCCAAGGAAGAAGA
15	g.72802170C>A	CYP1A1 +322	**ss159100335**	p.Asp108Tyr	CATTACTGATGAGGGTGAAGGTGTAGAGGT
16	g.72802230C>T	CYP1A1 +262	**ss159100336**	p.Gly88Ser	GCACCAGGGCCTGCCGGATGGTGTCCAGGC
17	g.72802259A>G	CYP1A1 +233	**rs17861094**	p.Ile78Thr	CCGCTCAGCACCACCACGGGTGTGGAGCCA
18	g.72802358C>T	CYP1A1 +134	**rs4646422**	p.Gly45Asp	AGGGTCAGCATGTGCCCAATCAGAGGCCAG
19	g.72803245G>A	CYP1A1 –754	**rs4646421**	INTRON 1	TTCCCTGGAAATTTCCAGCTGCTGTCACAT
20	g.72803860A>C	CYP1A1 –1369	**rs17861097**	INTRON 1	ATTTCAACATCTGAGAGAGGGTAAGACAAG
21	g.72804229C>A	CYP1A1 –1738	**rs2606345**	INTRON 1	TTCCAGCTCCCTGCAGTTGGCAATCTGTCA
22	g.72805066G>A	CYP1A1 –2575	**ss159100337**	PROMOTER	TCTCTGGGATTGGAGAGAAAGGTGGCGGAG
23	g.72805087G>A	CYP1A1 –2596	**ss159100338**	PROMOTER	GTGGCGGAGGGCGGCGGGGGTGGGGGGTCG
24	g.72805088G>A	CYP1A1 –2597	**ss159100339**	PROMOTER	TGGCGGAGGGCGGCGGGGGTGGGGGGTCGG

SNP #	SNP Name	Relative Position	NCBI Assignment	SNP Property	5' Flanking reference sequence
25	g.72805175G>A	CYP1A1 –2684	**ss159100340**	PROMOTER	CTGGGAACACCTGGAAGTCCCAATTCCAAG
	new AHRE	CYP1A1 –2825			
26	g.72805393C>T	CYP1A1 –2902	**rs17861099**	PROMOTER	GGCGGGGGCGGGGCTGCCCCGTGGTGACCT
	AHRE1	CYP1A1 –2928			
	AHRE2	CYP1A1 –3323			
	g.72805870C>T	CYP1A1 –3379	**rs4646418**	PROMOTER	GGGGGCAGAGGTCAGGCGGCCCGTCCCCGC
	AHRE3	CYP1A1 –3411			
27	g.72805943A>C	CYP1A1 –3452	**rs3826041**	PROMOTER	GGACCGGAGGCCCGCGCAGCCACCCAGCCG
28	g.72805958C>T	CYP1A1 –3467	**rs3826042**	PROMOTER	CGCAGCCACCCAGCCGACCCATTCCCCGGC
	AHRE4	CYP1A1 –3492			
29	g.72805984G>C	CYP1A1 –3493	**rs4646417**	PROMOTER	GGCCCGGCGCGGGGGCTCGCAGTCACGCGA
	g.72805985G>A	CYP1A1 –3494	**rs17861104**	PROMOTER	GCCCGGCGCGGGGGCTCGCAGTCACGCGAG
	AHRE5	CYP1A1 –3810			
30	g.72806502C>T	CYP1A1 –4011	**rs2470893**	PROMOTER	AACTACTATGTGCCAGGAGCTGTTCGGAGG
31	g.72806896T>C	CYP1A1 –4405	**rs7495708**	PROMOTER	CCTGGCCAATATGGCAAAACCCCATCTCTA
	AHRE6	CYP1A1 –4553			
32	g.72807071G>A	CYP1A1 –4580	**rs17861107**	PROMOTER	GCCACTGCACTCCAGCCTGGGTGACAGAGC
33	g.72807300C>A	CYP1A1 –4809	**rs17861108**	PROMOTER	GCAGGGAGCACCCCATTGAGAGACATGAGA
34	g.72807303C>A	CYP1A1 –4812	**rs17861109**	PROMOTER	GGGAGCACCCCATTGAGAGACATGAGACTT
	AHRE7	CYP1A1 –5504			
35	g.72809516C>G	CYP1A1 –7025	**rs17861110**	PROMOTER	TATAGCCCCGACCCAACCTGCCACTTCTTT
	new AHRE	CYP1A1 –7504			
36	g.72810325A>C	CYP1A1 –7834	**rs17861111**	PROMOTER	ACAGACTCAGCCCCTGGGCCAGGGCCCCTC
	AHRE8	CYP1A1 –8040			
	XRE II 3	CYP1A1 –8045			
37	g.72810752A>G	CYP1A1 –8261	**rs936225**	PROMOTER	CAAGCAAGATCCACAGCAAACCCCTGCCCC
38	g.72811935G>A	CYP1A1 –9444	**rs17861113**	PROMOTER	AGCAGATCTGGGAGATGAGGGGAGGAAGCG
39	g.72812040C>T	CYP1A1 –9549	**rs17861114**	PROMOTER	CTGTTAAAAAACCAATATCACCATCACCTC
40	g.72812273C>G	CYP1A1 –9782	**ss159100341**	PROMOTER	CAGGCCCCAAGAAGACCTACCCAGGTTGAA
41	g.72812385C>T	CYP1A1 –9894	**rs17861115**	PROMOTER	TGGGAGGCAAATTGTATCTCCAGAAGACCC

SNP #	SNP Name	Relative Position	NCBI Assignment	SNP Property	5' Flanking reference sequence
42	g.72812867T>C	CYP1A1 –10376	**rs12441817**	PROMOTER	AGTCTGGAGAGACAAGTGGAAAGCATCTAG
43	g.72813041C>T	CYP1A1 –10550	**rs4886605**	PROMOTER	GTTACATGCATAGATTGCAAAGTGGTCAGG
44	g.72813202G>A	CYP1A1 –10711	**rs17861118**	PROMOTER	CAGTATGCCTTCTCCTATAGTTTTGTTTTT
45	g.72813680C>A	CYP1A1 –11189	**rs11631784**	PROMOTER	CCGATTATTTAAAAGTATGTGGCACCTCCC
46	g.72814232A>G	CYP1A1 –11741	**rs17861120**	PROMOTER	TCCAAGACAGTCAGCTGAATGGATCCCCTC
47	g.72814485G>A	CYP1A1 –11994	**rs17861121**	PROMOTER	TTAAAGTCTCCAGGTTCCTCCCAACAGGAT
	XRE II 4	CYP1A2 –12783			
	AHRE9	CYP1A2 –11993			
	AHRE10	CYP1A2 –11822			
48	g.72818979C>T	CYP1A2 –10154	**rs17861122**	PROMOTER	AGGTGTGGATGATGTGATGGCCAGAGCTAC
49	g.72818980G>A	CYP1A2 –10153	**rs11632547**	PROMOTER	GGTGTGGATGATGTGATGGCCAGAGCTACC
50	g.72819195A>T	CYP1A2 –9938	**rs17861124**	PROMOTER	GAGAAACTAACCCAGGATAGTGTAGTACCA
51	g.72819208G>A	CYP1A2 –9925	**rs17861125**	PROMOTER	AGGATAGTGTAGTACCAATGGACTAGTAAC
52	g.72819500C>T	CYP1A2 –9633	**rs11632706**	PROMOTER	TCTAGAGGACAAAGGATTTTCAGCACTACC
53	g.72819544G>A	CYP1A2 –9589	**rs11632814**	PROMOTER	TCTTTTTCATCTGTTTGCCTGCATATTGGC
54	g.72819640G>A	CYP1A2 –9493	**rs16972208**	PROMOTER	GCACAACTTACGCTCAAAGAATATGGTTGA
55	g.72819929C>T	CYP1A2 –9204	**rs11633861**	PROMOTER	GCTGCAGTGAGCCATGATGGTGCCACTATA
	AHRE11	CYP1A2 –9145			
56	g.72820071A>C	CYP1A2 –9062	**ss159100342**	PROMOTER	ATGGCATCTTGGTGTCATACCAGGAGAGAG
57	g.72820286G>T	CYP1A2 –8847	**ss159100343**	PROMOTER	ATGGGGAAGAAGTGGTTGGAAGGAAGGTTG
58	g.72820374G>T	CYP1A2 –8759	**rs17861130**	PROMOTER	ATTGTCCCCACTGGCTGGTGGGAAAATTGT
59	g.72820453A>G	CYP1A2 –8680	**rs2472299**	PROMOTER	TCACATGTTGAGCTGAGGAGTTCAAGCTTT
60	g.72820546G>A	CYP1A2 –8587	**rs17861132**	PROMOTER	CAACTTCAGAGACTTCACTTGAGTGGCAGC
61	g.72820629G>A	CYP1A2 –8504	**rs17861133**	PROMOTER	TGGCTACTTAAATTTAAAATTAGGCCAGGC
62	g.72820857G>A	CYP1A2 –8276	**rs17861134**	PROMOTER	CCCATGAGGTCAAGGTTGCAGTGAGCCAAG
63	g.72820915C>T	CYP1A2 –8218	**rs17861135**	PROMOTER	ACAGAGTGAGACCCTTTTTTAAAAACAAAA
64	g.72820962A>G	CYP1A2 –8171	**rs2472300**	PROMOTER	ACCTCACAGGGTGGCTATAAAGCTAAAACA
65	g.72821129T>C	CYP1A2 –8004	**rs17861137**	PROMOTER	GGGCAGGTGAGCATTCTTACTTGTGAACAC
66	g.72821229G>T	CYP1A2 –7904	**rs17861138**	PROMOTER	CCCCACTAGCCATGGATACCAAAATCTATG
	XRE II 5	CYP1A2 –7623			

SNP #	**SNP Name**	**Relative Position**	**NCBI Assignment**	**SNP Property**	**5' Flanking reference sequence**
67	g.72821686A>G	CYP1A2 –7447	**rs17861139**	PROMOTER	AAAAAAAAAAAAAGATTTGTTAGTTATACT
	XRE II 6	CYP1A2 –7290			
68	g.72821975G>A	CYP1A2 –7158	**rs17861140**	PROMOTER	GCAGTATGGAGGCTGTGGGAACTTAACTGC
69	g.72822042C>T	CYP1A2 –7091	**rs17861141**	PROMOTER	GCAGGGTATAAAATGATGAGCCAAAGTGGC
70	g.72822176G>A	CYP1A2 –6957	**rs17861142**	PROMOTER	AGTGAACAGAAACTGCTGCTGAAGCCTTTG
	g.72822189G>A	CYP1A2 –6944	**rs17861143**	PROMOTER	TGCTGCTGAAGCCTTTGAGTGCTAAGCAAA
71	g.72822309T>C	CYP1A2 –6824	**rs2445618**	PROMOTER	AAAGATCAAAAATATCCAGTGGTGGCCAGG
72	g.72822437A>G	CYP1A2 –6696	**rs17861145**	PROMOTER	AACTCTGTCCCTCCTAAAGATACAAAAATT
	AHRE12	CYP1A2 –5597			
	XRE II 7	CYP1A2 –4575			
	XRE II 8	CYP1A2 –4103			
	E-BOX6	CYP1A2 –3813			
	E-BOX5	CYP1A2 –3670			
	AHRE13	CYP1A2 –3390			
	E-BOX4	CYP1A2 –3380			
	E-BOX3	CYP1A2 –3098			
	E-BOX2	CYP1A2 –2924			
	E-BOX1	CYP1A2 –2294			
73	g.72827329A>G	CYP1A2 –1804	**rs2069524**	PROMOTER	CTTGAACACCTGATCTCAGGTGATCCGCCC
74	g.72827425T>C	CYP1A2 –1708	**rs2069525**	PROMOTER	TTTTTAACAGAACCATTCAAAAGGAGGTTG
75	g.72828394T>G	CYP1A2 –739	**rs2069526**	INTRON 1	AAGACGGGGAGCCTGGGCTAGGTGTAGGGG
76	g.72828400G>C	CYP1A2 –733	**rs28399417**	INTRON 1	GGGAGCCTGGGCTAGGTGTAGGGGTCCTGA
77	g.72828784C>G	CYP1A2 –349	**rs17861149**	INTRON 1	CTTTGGTACAATACCCAGCATGCATGCTGTG
78	g.72828955G>A	CYP1A2 –178	**rs17861150**	INTRON 1	CCAGCTCTCAGATTCTGTGATGCTCAAAGG
79	g.72828970C>A	CYP1A2 –163	**rs762551**	INTRON 1	TGTGATGCTCAAAGGGTGAGCTCTGTGGGC
80	g.72829185C>G	CYP1A2 +53	**rs17861152**	p.Ser18Cys	CCTTCTCGGCCACAGAGCTTCTCCTGGCCT
81	g.72829381G>T	CYP1A2 +249	**rs17861153**	Syn.	GGACGTCCTGCAGATCCGCATTGGCTCCAC
82	g.72829438C>T	CYP1A2 +306	**rs17861154**	Syn.	CACCATCCGGCAGGCCCTGGTGCGGCAGGG
83	g.72829747C>T	CYP1A2 +615	**rs17861155**	Syn.	AGTGGCCAACGTCATTGGTGCCATGTGCTT
84	g.72830334C>T	CYP1A2 +1202	**rs4646425**	INTRON 2	TGTCACGTTGCTTCCCTGTGTTCACACTAA

SNP #	SNP Name	Relative Position	NCBI Assignment	SNP Property	5' Flanking reference sequence
85	g.72830564T>C	CYP1A2 +1432	**ss159100344**	INTRON 2	CTTTGACCTTGGAAGTGCCAGAGGTGCCCC
86	g.72830645C>A	CYP1A2 +1513	**rs17861157**	p.Ser298Arg	CAAGCACAGCAAGAAGGGGCCTAGAGCCAG
87	g.72830646G>A	CYP1A2 +1514	**rs35796837**	p.Gly299Ser	AAGCACAGCAAGAAGGGGCCTAGAGCCAGC
88	g.72830947C>A	CYP1A2 +1815	**ss159100345**	INTRON 3	TCCCCATCTGAACAATAAGGTGTTCTCTGG
89	g.72831290G>A	CYP1A2 +2159	**rs2472304**	INTRON 4	CCCCAACCCTATAGCCAGGAGAAGCCTTGA
80	g.72831453G>C	CYP1A2 +2321	**rs3743484**	INTRON 4	AGTCTTACATAAGAGTGACATGGGGTATAA
91	g.72831746G>A	CYP1A2 +2614	**rs17861159**	INTRON 5	GCCCTGGCTCAGCATCTCCTTCCCGACCTC
92	g.72832260G>C	CYP1A2 +3128	**ss159100346**	INTRON 5	GAAATTCCCCGGGGCTGCCCCAGGCTGCCT
93	g.72832387C>A	CYP1A2 +3255	**ss159100347**	INTRON 5	AGCCAGGCGCAAAGAGAAGTTTAGTAAATA
94	g.72832745T>C	CYP1A2 +3613	**rs4646427**	INTRON 6	AGGCTGTTTGTCCCTGCTAGGAACTGTTTA
95	g.72833457T>C	CYP1A2 +4325	**ss159100348**	INTRON 6	GTAGGATGAAGAAACGTATGTCTTTCTTTC
96	g.72833665A>G	CYP1A2 +4533	**ss159100349**	INTRON 6	ATTTAAAGTAAATTAATATTAAAAGGCACT
97	g.72834479C>T	CYP1A2 +5347	**rs2470890**	Syn.	TGTCCAGGCGCGGCTGCGCTTCTCCATCAA
98	g.72834527C>T	CYP1A2 +5395	**ss159100350**	3' UTR	TTCTGAGGCCAGGGAGCGAGTGGGGGCCAG
	XRE II 10	CYP1A2 +6223			
99	g.72835806C>G	CYP1A2 +6674	**rs17861162**	3' UTR	GTGCTGGCATTACAGGTGTGAGCCACGGTG
100	g.72835817A>G	CYP1A2 +6685	**rs17861163**	3' UTR	ACAGGTGTGAGCCACGGTGCCCGGCCCACA
101	g.72836881G>A	CYP1A2 +7749	**rs17861164**	3' END	GATGCCTCACCCCTGTAATTCCAGCACTTC
102	g.72836996G>A	CYP1A2 +7864	**rs2960192**	3' END	TACAAAAATTAGCCAGGCATGGTGGCGGGC

Green background: SNPs detected in a large panel of subjects and not included in the genetic analysis. *Yellow background***:** 60 SNPs found in resequencing of the Oceanian/Amerindian samples (current study). ***Yellow background with red letters*****:** 21 SNPs found only during the Oceanian/Amerindian resequencing (current study). Of these, 17 are new and unreported and have been assigned dbSNP ss numbers. ***Yellow background with black letters*****:** 39 SNPs found in the resequencing of all populations—AF, AS, CA, OC, AM (Jiang et al., 2005) plus two Oceanians and three Amerindian populations (current study). **In clear background with black letters:** 46 SNPs found only during the resequencing of AF, AS, CA, OC, AM (Jiang et al., 2005). Relative positions listed are calculated with respect to the translation start codon. To transform these to positions relative to TSS, for *CYP1A1***:** a) add 2,439 if position ≥ +1; b) add 2,440 if –1 ≥ position ≥ –2,439; c) add 2,439 if position < –2,439. For *CYP1A2*: a) add 896 if position ≥ +1; b) add 897 if –1 ≥ position ≥ –896; c) add 896 if position < –896. Only 102 SNPs included in the genetic analysis have received a SNP number.

SUPP. TABLE S2. Two Hundred Ninety-Five SNPs across the *CYP1A1_CYP1A2* Locus—278 SNPs Previously Reported plus the 17 New SNPs Found in the Present Study

CHROMOSOMAL POSITION	REFERENCE SEQUENCE NUMBER	RELATIVE POSITION	ROLE	In a masked area
		CYP1A1		
chr15:72789002	rs28693826	13490 G/T	3′ flank	
chr15:72789886	rs35669097	12606 A/G	3′ flank	
chr15:72789925	rs34160031	12567 A/G	3′ flank	
chr15:72790104	rs1456432	12388 A/G	3′ flank	
chr15:72790126	rs34465004	12366 C/G	3′ flank	
chr15:72790130	rs35646069	12362 A/G	3′ flank	
chr15:72790305	rs28371819	12187 A/G	3′ flank	
chr15:72790561	rs6495121	11931 G/T	3′ flank	
chr15:72792628	rs28689146	9864 A/T	3′ flank	
chr15:72792678	rs12915975	9814 A/G	3′ flank	
chr15:72794607	rs35570076	7885 A/G	3′ flank	
chr15:72794702	rs35857345	7790 C/T	3′ flank	
chr15:72794814	rs2472308	7678 G/T	3′ flank	
chr15:72794835	rs36098749	7657 C/G	3′ flank	
chr15:72794853	rs2472309	7639 G/T	3′ flank	
chr15:72798054	rs28661955	4438 C/T	3′ flank	
chr15:72798247	rs17861080	4245 C/T	3′ flank	
chr15:72798250	rs17861081	4242 A/G	3′ flank	
chr15:72798264	rs17861082	4228 A/G	3′ flank	
chr15:72798694	rs4646903	3798 C/T	3′ flank	
chr15:72798907	rs2515900	3585 A/G	3′ flank	
chr15:72798908	rs2472307	3584 C/T	3′ flank	
chr15:72798965	rs28427336	3527 C/G	3'-UTR	
chr15:72799072	rs17861084	3420 G/T	3'-UTR	
chr15:72799169	rs28399432	3323 C/G	3'-UTR	
chr15:72799170	rs17861085	3322 A/C	3'-UTR	
chr15:72799175	rs17861086	3317 C/G	3'-UTR	
chr15:72799260	rs4986884	3232 C/T	3'-UTR	
chr15:72799270	rs4986883	3222 A/G	3'-UTR	
chr15:72799288	rs1800031	3204 C/T	3'-UTR	
chr15:72799417	rs35427048	3075 C/T	3'-UTR	
chr15:72799524	rs4986882	2968 A/G	3'-UTR	
chr15:72799535	rs28399431	2957 C/T	3'-UTR	
chr15:72799540	rs4986881	2952 C/G/T	3'-UTR	
chr15:72799697	rs4986880	2795 C/T	3'-UTR	
chr15:72799798	rs2606346	2694 C/T	3'-UTR	
chr15:72799947	rs28399430	2545 C/G	Coding exon	
chr15:72799978	rs28399429	2514 A/G	Coding exon	
chr15:72800014	rs36121583	2478 G/T	Coding exon	
chr15:72800034	rs2278970	2458 C/G	Coding exon	
chr15:72800038	rs1048943	2454 A/G	Coding exon	
chr15:72800040	rs1799814	2452 A/C	Coding exon	
chr15:72800277	rs28399428	2215 C/T	Intron (boundary)	
chr15:72800616	rs2856833	1876 A/C	Coding exon	
chr15:72801080	rs4987133	1412 C/T	Coding exon	
chr15:72801102	rs34260157	1390 C/T	Coding exon	

chr15:72801524	ss159100334	968 C/T	Intron
chr15:72801622	rs4986879	870 A/G	Intron (boundary)
chr15:72801974	rs28399427	518 C/G	Coding exon
chr15:72801979	rs28399426	513 A/C	Coding exon
chr15:72802098	rs35196245	394 C/G	Coding exon
chr15:72802170	ss159100335	322 C/A	Coding exon
chr15:72802171	rs35267501	321 C/T	Coding exon
chr15:72802215	rs2229150	277 C/T	Coding exon
chr15:72802230	ss159100336	262 C/T	Coding exon
chr15:72802259	rs17861094	233 A/G	Coding exon
chr15:72802296	rs35035798	196 A/G	Coding exon
chr15:72802358	rs4646422	134 A/G	Coding exon
chr15:72802524	rs28399425	-33 C/T	Intron (boundary)
chr15:72802711	rs34331219	-220 A/G	Intron
chr15:72802832	rs2606344	-341 G/T	Intron
chr15:72802905	rs35374834	-414 C/T	Intron
chr15:72802932	rs4986885	-441 A/G	Intron
chr15:72802960	rs35766947	-469 A/G	Intron
chr15:72803043	rs34268526	-552 A/C	Intron
chr15:72803245	rs4646421	-754 C/T	Intron
chr15:72803260	rs35992445	-769 C/T	Intron
chr15:72803355	rs34445263	-864 C/G	Intron
chr15:72803478	rs4646420	-987 C/T	Intron
chr15:72803523	rs34513002	-1032 C/T	Intron
chr15:72803702	rs8031941	-1211 G/T	Intron
chr15:72803860	rs17861097	-1369 A/C	Intron
chr15:72804229	rs2606345	-1738 A/C	Intron
chr15:72804513	rs2856844	-2022 C/G	Intron
chr15:72804537	rs35892217	-2046 C/T	Intron
chr15:72804678	rs35089928	-2187 A/G	Intron
chr15:72804680	rs34726614	-2189 A/G	Intron
chr15:72805066	ss159100337	-2575 G/A	Promoter
chr15:72805087	ss159100338	-2596 G/A	Promoter
chr15:72805088	ss159100339	-2597 G/A	Promoter
chr15:72805175	ss159100340	-2684 G/A	Promoter
chr15:72805249	rs35701229	-2758 C/T	Promoter
chr15:72805271	rs36099343	-2780 C/T	Promoter
chr15:72805366	rs35339247	-2875 C/T	Promoter
chr15:72805383	rs35686934	-2892 C/T	Promoter
chr15:72805393	rs17861099	-2902 C/T	Promoter
chr15:72805760	rs28588839	-3269 A/G	Promoter
chr15:72805870	rs4646418	-3379 A/G	Promoter
chr15:72805943	rs3826041	-3452 A/C	Promoter
chr15:72805958	rs3826042	-3467 C/T	Promoter
chr15:72805966	rs2856832	-3475 C/G	Promoter
chr15:72805968	rs2856831	-3477 C/G	Promoter
chr15:72805984	rs4646417	-3493 C/G	Promoter
chr15:72805985	rs17861104	-3494 A/G	Promoter
chr15:72805989	rs2472296	-3498 A/G	Promoter
chr15:72805993	rs2445619	-3502 A/G	Promoter
chr15:72806191	rs35141552	-3700 C/G	Promoter
chr15:72806238	rs36090160	-3747 A/G	Promoter
chr15:72806342	rs3809585	-3851 C/G	Promoter
chr15:72806347	rs36074335	-3856 C/T	Promoter
chr15:72806396	rs36018288	-3905 A/G	Promoter
chr15:72806491	rs34077361	-4000 A/G	Promoter
chr15:72806502	rs2470893	-4011 A/G	Promoter

chr15:72806798	rs34458627	-4307 C/T	Promoter	
chr15:72806896	rs7495708	-4405 C/T	Promoter	
chr15:72806938	rs34069996	-4447 A/G	Promoter	
chr15:72807050	rs35236831	-4559 C/G	Promoter	
chr15:72807071	rs17861107	-4580 A/G	Promoter	
chr15:72807300	rs17861108	-4809 A/C	Promoter	
chr15:72807303	rs17861109	-4812 A/C	Promoter	
chr15:72808964	rs7180066	-6473 G/T	Promoter	x
chr15:72808976	rs7179590	-6485 A/G	Promoter	x
chr15:72808984	rs7181062	-6493 A/T	Promoter	x
chr15:72809084	rs7179952	-6593 C/T	Promoter	x
chr15:72809183	rs28621690	-6692 C/T	Promoter	x
chr15:72809216	rs8039800	-6725 A/T	Promoter	x
chr15:72809225	rs7180012	-6734 A/T	Promoter	x
chr15:72809516	rs17861110	-7025 C/G	Promoter	
chr15:72810325	rs17861111	-7834 A/C	Promoter	
chr15:72810752	rs936225	-8261 A/G	Promoter	
chr15:72810989	rs7182554	-8498 C/T	Promoter	x
chr15:72811935	rs17861113	-9444 A/G	Promoter	
chr15:72812040	rs17861114	-9549 C/T	Promoter	
chr15:72812199	rs2470892	-9708 C/G	Promoter	
chr15:72812273	ss159100341	-9782 C/G	Promoter	
chr15:72812385	rs17861115	-9894 C/T	Promoter	
chr15:72812386	rs2470891	-9895 C/T	Promoter	
chr15:72812464	rs35434501	-9973 A/G	Promoter	
chr15:72812542	rs35245678	-10051 A/G	Promoter	
chr15:72812690	rs35042613	-10199 A/C	Promoter	
chr15:72812867	rs12441817	-10376 C/T	Promoter	
chr15:72813041	rs4886605	-10550 C/T	Promoter	
chr15:72813202	rs17861118	-10711 A/G	Promoter	
chr15:72813333	rs7496533	-10842 A/T	Promoter	
chr15:72813480	rs7496395	-10989 C/T	Promoter	
chr15:72813680	rs11631784	-11189 A/C	Promoter	
chr15:72814232	rs17861120	-11741 A/G	Promoter	
chr15:72814485	rs17861121	-11994 A/G	Promoter	
		CYP1A2		
chr15:72818574	rs2472298	-10559 A/G	Promoter	
chr15:72818785	rs34769904	-10348 A/G	Promoter	
chr15:72818808	rs35109618	-10325 C/T	Promoter	
chr15:72818979	rs17861122	-10154 C/T	Promoter	
chr15:72818980	rs11632547	-10153 A/G	Promoter	
chr15:72818995	rs35290120	-10138 A/G	Promoter	
chr15:72819013	rs34033635	-10120 C/T	Promoter	
chr15:72819195	rs17861124	-9938 A/T	Promoter	
chr15:72819208	rs17861125	-9925 A/G	Promoter	
chr15:72819480	rs16972198	-9653 A/G	Promoter	
chr15:72819500	rs11632706	-9633 C/T	Promoter	
chr15:72819544	rs11632814	-9589 A/G	Promoter	
chr15:72819640	rs16972208	-9493 A/G	Promoter	
chr15:72819929	rs11633861	-9204 C/T	Promoter	
chr15:72820071	ss159100342	-9062 A/C	Promoter	
chr15:72820286	ss159100343	-8847 G/T	Promoter	
chr15:72820374	rs17861130	-8759 G/T	Promoter	
chr15:72820448	rs35104183	-8685 C/T	Promoter	
chr15:72820453	rs2472299	-8680 A/G	Promoter	

chr15:72820525	rs35900399	-8608 C/T	Promoter	
chr15:72820546	rs17861132	-8587 A/G	Promoter	
chr15:72820629	rs17861133	-8504 A/G	Promoter	
chr15:72820857	rs17861134	-8276 A/G	Promoter	
chr15:72820915	rs17861135	-8218 C/T	Promoter	
chr15:72820962	rs2472300	-8171 A/G	Promoter	
chr15:72821129	rs17861137	-8004 C/T	Promoter	
chr15:72821229	rs17861138	-7904 G/T	Promoter	
chr15:72821686	rs17861139	-7447 A/G	Promoter	
chr15:72821899	rs7172415	-7234 C/G	Promoter	
chr15:72821975	rs17861140	-7158 A/G	Promoter	
chr15:72822042	rs17861141	-7091 C/T	Promoter	
chr15:72822176	rs17861142	-6957 A/G	Promoter	
chr15:72822189	rs17861143	-6944 A/C	Promoter	
chr15:72822309	rs2445618	-6824 C/T	Promoter	
chr15:72822437	rs17861145	-6696 A/G	Promoter	
chr15:72822835	rs28454554	-6298 A/G	Promoter	x
chr15:72823367	rs2472301	-5766 A/G	Promoter	x
chr15:72823369	rs2472302	-5764 C/T	Promoter	x
chr15:72825243	rs2069512	-3890 A/G	Promoter	x
chr15:72825244	rs2069513	-3889 A/G	Promoter	x
chr15:72825273	rs2069514	-3860 A/G	Promoter	x
chr15:72825440	rs2069515	-3693 A/G	Promoter	x
chr15:72825456	rs2069516	-3677 A/G	Promoter	x
chr15:72825508	rs2069517	-3625 C/T	Promoter	x
chr15:72825535	rs2069519	-3598 G/T	Promoter	x
chr15:72825539	rs2069520	-3594 G/T	Promoter	x
chr15:72825876	rs2470894	-3257 G/T	Promoter	x
chr15:72825878	rs2472303	-3255 A/T	Promoter	x
chr15:72826020	rs2069521	-3113 A/G	Promoter	x
chr15:72826224	rs34341603	-2909 G/T	Promoter	x
chr15:72826286	rs2069522	-2847 C/T	Promoter	x
chr15:72826299	rs34633714	-2834 C/T	Promoter	x
chr15:72826325	rs12592480	-2808 A/C	Promoter	x
chr15:72826356	rs35590132	-2777 A/G	Promoter	x
chr15:72827030	rs34834171	-2103 A/G	Promoter	x
chr15:72827189	rs35083748	-1944 A/C	Promoter	x
chr15:72827288	rs2069523	-1845 A/C	Promoter	
chr15:72827329	rs2069524	-1804 A/C/G/T	Promoter	
chr15:72827425	rs2069525	-1708 A/C/G/T	Promoter	
chr15:72827850	rs28622071	-1283 A/G	Promoter	
chr15:72828082	rs28399416	-1051 C/T	Promoter	
chr15:72828394	rs2069526	-739 G/T	Intron	
chr15:72828400	rs28399417	-733 C/G	Intron	
chr15:72828404	rs12720461	-729 C/T	Intron	
chr15:72828471	rs35198905	-662 G/T	Intron	
chr15:72828784	rs17861149	-349 C/G	Intron	
chr15:72828955	rs17861150	-178 A/G	Intron	
chr15:72828970	rs762551	-163 A/C	Intron	
chr15:72828972	rs35858607	-161 A/C	Intron	
chr15:72829185	rs17861152	53 C/G	Coding exon	
chr15:72829262	rs3743482	130 A/G	Coding exon	
chr15:72829354	rs35407132	222 C/T	Coding exon	
chr15:72829381	rs17861153	249 G/T	Coding exon	
chr15:72829438	rs17861154	306 C/T	Coding exon	

chr15:72829442	rs34067076	310	A/G	Coding exon	
chr15:72829525	rs34153904	393	A/G	Coding exon	
chr15:72829747	rs17861155	615	C/T	Coding exon	
chr15:72829982	rs34954622	850	C/T	Intron (boundary)	
chr15:72830026	rs34182516	894	C/T	Intron (boundary)	
chr15:72830334	rs4646425	1202	C/T	Intron	
chr15:72830412	rs35407489	1280	A/G	Intron	
chr15:72830484	rs34264399	1352	A/G	Intron (boundary)	
chr15:72830564	ss159100344	1432	T/C	Intron	
chr15:72830645	rs17861157	1513	A/C	Coding exon	
chr15:72830646	rs35796837	1514	A/G	Coding exon	
chr15:72830681	rs34058039	1549	C/T	Coding exon	
chr15:72830691	rs28399418	1559	A/G	Coding exon	
chr15:72830722	rs28399419	1590	C/T	Intron (boundary)	
chr15:72830947	ss159100345	1815	C/A	Intron	
chr15:72831291	rs2472304	2159	A/G	Intron (boundary)	
chr15:72831301	rs34711134	2169	A/G	Intron (boundary)	
chr15:72831364	rs34440552	2232	C/T	Intron	
chr15:72831453	rs3743484	2321	C/G	Intron (boundary)	
chr15:72831702	rs28399420	2570	A/G	Intron (boundary)	
chr15:72831729	rs34980540	2597	A/G	Intron (boundary)	
chr15:72831746	rs17861159	2614	A/G	Intron	
chr15:72831777	rs35902286	2645	C/G	Intron	
chr15:72831778	rs28399421	2646	C/T	Intron	
chr15:72831826	rs28399422	2694	A/C	Intron	
chr15:72832201	rs34931285	3069	C/G	Intron	
chr15:72832260	ss159100346	3128	G/C	Intron	
chr15:72832300	rs2472305	3168	A/G	Intron	
chr15:72832339	rs2960195	3207	C/T	Intron	
chr15:72832341	rs2960194	3209	C/T	Intron	
chr15:72832387	ss159100347	3255	C/A	Intron	
chr15:72832411	rs36043490	3279	C/G	Intron	
chr15:72832419	rs2472306	3287	A/G	Intron	
chr15:72832479	rs4886607	3347	C/T	Intron (boundary)	
chr15:72832620	rs2960193	3488	C/T	Coding exon	
chr15:72832737	rs34356615	3605	A/G	Intron (boundary)	
chr15:72832745	rs4646427	3613	C/T	Intron (boundary)	
chr15:72833457	ss159100348	4325	T/C	Intron	
chr15:72833665	ss159100349	4533	A/G	Intron	
chr15:72833730	rs12593808	4598	G/T	Intron	x
chr15:72833928	rs12904742	4796	A/G	Intron	x
chr15:72834142	rs28399423	5010	C/T	Intron (boundary)	
chr15:72834222	rs28399424	5090	C/T	Coding exon	
chr15:72834300	rs34151816	5168	C/T	Coding exon	
chr15:72834365	rs4646428	5233	A/T	Coding exon	
chr15:72834474	rs35647699	5342	A/G	Coding exon	
chr15:72834479	rs2470890	5347	C/T	Coding exon	
chr15:72834527	ss159100350	5395	C/T	3'-UTR	
chr15:72834653	rs11636419	5521	A/G	3'-UTR	x
chr15:72834756	rs28465265	5624	A/C	3'-UTR	x
chr15:72835022	rs1139497	5890	A/G	3'-UTR	
chr15:72835806	rs17861162	6674	C/G	3'-UTR	
chr15:72835817	rs17861163	6685	A/G	3'-UTR	
chr15:72836615	rs7403788	7483	A/T	3’ flank	
chr15:72836705	rs35562753	7573	C/T	3’ flank	
chr15:72836881	rs17861164	7749	A/G	3’ flank	
chr15:72836996	rs2960192	7864	C/T	3’ flank	

chr15:72837253	rs12912043	8121	A/C	3′ flank
chr15:72837400	rs12912491	8268	A/G	3′ flank
chr15:72837410	rs12912250	8278	C/T	3′ flank
chr15:72837426	rs12912259	8294	C/T	3′ flank
chr15:72837560	rs7402578	8428	C/T	3′ flank
chr15:72837907	rs12591078	8775	C/T	3′ flank
chr15:72838061	rs11072501	8929	A/G	3′ flank
chr15:72838672	rs8038819	9540	A/C	3′ flank
chr15:72839115	rs1350194	9983	A/T	3′ flank
chr15:72839548	rs12903896	10416	C/T	3′ flank
chr15:72839824	rs11854147	10692	C/T	3′ flank
chr15:72839873	rs11072502	10741	A/G	3′ flank
chr15:72839948	rs11072503	10816	C/T	3′ flank
chr15:72839964	rs11072504	10832	C/T	3′ flank
chr15:72839967	rs12439187	10835	C/T	3′ flank
chr15:72839980	rs11857376	10848	A/G	3′ flank
chr15:72839988	rs11072505	10856	A/G	3′ flank
chr15:72840047	rs11072506	10915	A/G	3′ flank
chr15:72840485	rs11637417	11353	C/G	3′ flank
chr15:72840495	rs11632725	11363	C/T	3′ flank
chr15:72841919	rs11072507	12787	C/G	3′ flank
chr15:72842060	rs28610543	12928	C/T	3′ flank
chr15:72843411	rs11631682	14279	A/G	3′ flank
chr15:72843934	rs11632822	14802	C/T	3′ flank
chr15:72844256	rs4886406	15124	A/C	3′ flank
chr15:72844800	rs4886606	15668	A/T	3′ flank

The 106 SNPs on *grey background* were found during our studies [current study plus (Jiang et al., 2005)] of eight world populations. Relative positions listed are calculated with respect to translation start codon. To transform these to positions relative to the TSS, see Supp. Table S1 legend. SNPs marked with an “x” (in a masked area) were not surveyed in our studies, because they are located in regions with more than 85% highly-repetitive DNA, which were excluded by *RepeatMasker*.

.

SUPP. TABLE S3. Comparison of Number of SNPs Found vs SNPs Not Found in Our Studies

Total SNPs previously reported**:**	278	
New SNPs found in the present study**:**	17	
Grand total**:**	295	
Reported SNPs located outside our resequenced area**:**		
SNPs downstream of *CYP1A1* +5000	15	
SNPs downstream of *CYP1A2* +8000	26	
SNPs in masked areas**:**	33	
Total**:**	74	
SNPs previously reported and newly found in the resequenced area**:**	221	(100%)
SNPs found in the resequencing of same area**:**	106	(48%)
SNPs not found in the resequencing of same area**:**	115	(52%)

The SNPs reported previously include all SNPs found in the present study, plus those found previously by us (Jiang et al., 2005), plus those in dbSNP.

SUPP. TABLE S4. Frequency of 102 Individual SNPs across the *CYP1A1_CYP1A2* Locus in Eight World Populations

SNP #	SNP Name	Relative Position	NCBI Assignment	A1	A2	Ancestral	P_{A2} in Africans	P_{A2} in Caucasians	P_{A2} in Asians	P_{A2} in Samoans	P_{A2} in New Guineans	P_{A2} in Pehuenche	P_{A2} in Teribe	P_{A2} in Dogrib
1	g.72798247C>T	*CYP1A1* +4245	**rs17861080**	C	T	C	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
2	g.72798250G>A	*CYP1A1* +4242	**rs17861081**	G	A	G	3/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
3	g.72798264A>G	*CYP1A1* +4228	**rs17861082**	A	G	G*	3/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
4	g.72798694A>G	*CYP1A1* +3798	**rs4646903**	A	G	A	3/12	1/12	2/6	14/24	14/14	7/8	3/8	5/16
5	g.72799072G>T	*CYP1A1* +3420	**rs17861084**	G	T	G	2/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
6	g.72799170C>A	*CYP1A1* +3322	**rs17861085**	C	A	A*	3/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
7	g.72799175C>G	*CYP1A1* +3317	**rs17861086**	C	G	C	0/12	0/14	0/8	1/24	0/16	0/8	0/8	0/16
8	g.72799260A>G	*CYP1A1* +3232	**rs4986884**	A	G	G*	3/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
9	g.72799524T>C	*CYP1A1* +2968	**rs4986882**	T	C	C*	3/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
10	g.72799540G>C	*CYP1A1* +2952	**rs4986881**	G	C	C*	3/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
11	g.72799697G>A	*CYP1A1* +2795	**rs4986880**	G	A	G	3/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
12	g.72800038T>C	*CYP1A1* +2454	**rs1048943**	T	C	T	0/12	0/14	1/8	2/24	0/16	7/8	3/8	5/16
13	g.72800040G>T	*CYP1A1* +2452	**rs1799814**	G	T	G	0/12	1/14	0/8	0/24	0/16	0/8	0/8	0/16
14	g.72801524C>T	*CYP1A1* +968	**ss159100334**	C	T	C	0/12	0/14	0/8	0/24	0/16	0/8	0/8	1/16
15	g.72802170C>A	*CYP1A1* +322	**ss159100335**	C	A	C	0/12	0/14	0/8	1/24	0/16	0/8	0/8	0/16
16	g.72802230C>T	*CYP1A1* +262	**ss159100336**	C	T	C	0/12	0/14	0/8	0/24	0/16	1/8	0/8	0/16
17	g.72802259A>G	*CYP1A1* +233	**rs17861094**	A	G	A	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16

SNP #	**SNP Name**	**Relative Position**	**NCBI Assignment**	**A1**	**A2**	**Ancestral**	**P_{A2} in Africans**	**P_{A2} in Caucasians**	**P_{A2} in Asians**	**P_{A2} in Samoans**	**P_{A2} in New Guineans**	**P_{A2} in Pehuenche**	**P_{A2} in Teribe**	**P_{A2} in Dogrib**
18	g.72802358C>T	*CYP1A1* +134	**rs4646422**	C	T	C	0/12	0/14	2/8	5/24	0/16	0/8	0/8	0/16
19	g.72803245G>A	*CYP1A* –754	**rs4646421**	G	A	G	1/8	1/14	3/8	13/24	15/16	7/8	3/8	4/14
20	g.72803860A>C	*CYP1A1* –1369	**rs17861097**	A	C	A	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
21	g.72804229C>A	*CYP1A1* –1738	**rs2606345**	C	A	C	1/12	8/14	0/8	1/24	0/16	1/8	3/8	2/16
22	g.72805066G>A	*CYP1A1* –2575	**ss159100337**	G	A	G	0/12	0/14	0/8	0/22	0/16	0/8	0/8	1/16
23	g.72805087G>A	*CYP1A1* –2596	**ss159100338**	G	A	G	0/12	0/14	0/8	0/24	1/16	0/8	0/8	0/16
24	g.72805088G>A	*CYP1A1* –2597	**ss159100339**	G	A	G	0/12	0/14	0/8	0/24	2/16	0/8	0/8	0/16
25	g.72805175G>A	*CYP1A1* –2684	**ss159100340**	G	A	G	0/12	0/14	0/8	1/22	0/16	0/8	0/8	0/16
26	g.72805393C>T	*CYP1A1* –2902	**rs17861099**	C	T	C	0/12	1/14	0/8	0/24	0/16	0/8	0/8	0/16
27	g.72805943A>C	*CYP1A1* –3452	**rs3826041**	A	C	C*	8/10	2/14	4/8	15/24	11/12	7/8	3/8	7/16
28	g.72805958C>T	*CYP1A1* –3467	**rs3826042**	C	T	C	0/10	1/14	1/8	2/24	0/12	0/8	0/8	2/16
29	g.72805984G>C	*CYP1A1* –3493	**rs4646417**	G	C	G	2/10	0/14	0/8	0/24	0/16	0/8	0/8	0/16
30	g.72806502C>T	*CYP1A1* –4011	**rs2470893**	C	T	C	0/10	1/10	0/8	0/24	0/16	0/8	0/8	0/16
31	g.72806896T>C	*CYP1A1* –4405	**rs7495708**	T	C	C*	9/12	2/12	4/8	15/24	9/10	7/8	3/8	7/16
32	g.72807071G>A	*CYP1A1* –4580	**rs17861107**	G	A	G	2/6	0/12	0/6	0/24	0/16	0/8	0/8	0/16
33	g.72807300C>A	CYP1A1 –4809	**rs17861108**	C	A	C	3/10	0/14	0/8	0/24	0/16	0/8	0/8	0/16
34	g.72807303C>A	*CYP1A1* –4812	**rs17861109**	C	A	C	3/10	1/12	3/8	13/24	15/16	7/8	3/8	3/14
35	g.72809516C>G	*CYP1A1* –7025	**rs17861110**	C	G	C	0/12	0/14	1/8	2/24	0/16	7/8	3/8	4/16
36	g.72810325A>C	*CYP1A1* –7834	**rs17861111**	A	C	A	1/12	0/8	0/8	0/24	0/16	0/8	0/8	0/16

SNP #	SNP Name	Relative Position	NCBI Assignment	A1	A2	Ancestral	P_{A2} in Africans	P_{A2} in Caucasians	P_{A2} in Asians	P_{A2} in Samoans	P_{A2} in New Guineans	P_{A2} in Pehuenche	P_{A2} in Teribe	P_{A2} in Dogrib
37	g.72810752A>G	*CYP1A1* –8261	**rs936225**	A	G	A	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
38	g.72811935G>A	*CYP1A1* –9444	**rs17861113**	G	A	G	0/12	0/14	1/8	0/24	0/16	0/8	0/8	0/16
39	g.72812040C>T	*CYP1A1* –9549	**rs17861114**	C	T	C	0/12	1/14	0/8	0/24	0/16	0/8	0/8	0/16
40	g.72812273C>G	*CYP1A1* –9782	**ss159100341**	C	G	C	0/12	0/14	0/8	0/24	0/16	0/8	0/8	2/16
41	g.72812385C>T	*CYP1A1* –9894	**rs17861115**	C	T	C	0/12	0/14	2/8	2/24	0/16	7/8	3/8	5/16
42	g.72812867T>C	*CYP1A1* –10376	**rs12441817**	T	C	T	1/12	1/14	5/8	16/24	15/16	7/8	3/8	5/16
43	g.72813041C>T	*CYP1A1* –10550	**rs4886605**	C	T	T*	9/12	2/14	6/8	18/24	15/16	7/8	3/8	7/14
44	g.72813202G>A	*CYP1A1* –10711	**rs17861118**	G	A	G	0/12	1/14	0/8	0/24	0/16	0/8	0/8	0/16
45	g.72813680C>A	*CYP1A1* –11189	**rs11631784**	C	A	C	4/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
46	g.72814232A>G	*CYP1A1* –11741	**rs17861120**	A	G	A	1/12	1/14	4/8	15/22	15/16	7/8	3/8	5/16
47	g.72814485G>A	*CYP1A1* –11994	**rs17861121**	G	A	G	2/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
48	g.72818979C>T	*CYP1A2* –10154	**rs17861122**	C	T	C	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
49	g.72818980G>A	*CYP1A2* –10153	**rs11632547**	G	A	G	1/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16
50	g.72819195A>T	*CYP1A2* –9938	**rs17861124**	A	T	T*	2/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16
51	g.72819208G>A	*CYP1A2* –9925	**rs17861125**	G	A	A*	2/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16
52	g.72819500C>T	*CYP1A2* –9633	**rs11632706**	C	T	T*	2/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16
53	g.72819544G>A	*CYP1A2* –9589	**rs11632814**	G	A	A*	2/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16
54	g.72819640G>A	*CYP1A2* –9493	**rs16972208**	G	A	G	2/12	0/14	2/8	10/24	0/16	7/8	3/8	5/16
55	g.72819929C>T	*CYP1A2* –9204	**rs11633861**	C	T	C	1/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16

SNP #	**SNP Name**	**Relative Position**	**NCBI Assignment**	**A1**	**A2**	**Ancestral**	**P_{A2} in Africans**	**P_{A2} in Caucasians**	**P_{A2} in Asians**	**P_{A2} in Samoans**	**P_{A2} in New Guineans**	**P_{A2} in Pehuenche**	**P_{A2} in Teribe**	**P_{A2} in Dogrib**
56	g.72820071A>C	*CYP1A2* –9062	**ss159100342**	A	C	A	0/12	0/14	0/8	2/24	0/16	0/8	0/8	0/16
57	g.72820286G>T	*CYP1A2* –8847	**ss159100343**	G	T	G	0/12	0/14	0/8	0/24	3/16	0/8	0/8	0/16
58	g.72820374G>T	*CYP1A2* –8759	**rs17861130**	G	T	G	0/10	0/14	1/8	0/24	0/16	0/8	0/8	0/16
59	g.72820453A>G	*CYP1A2* –8680	**rs2472299**	A	G	G*	8/10	9/14	6/8	19/24	13/16	8/8	3/8	9/16
60	g.72820546G>A	*CYP1A2* –8587	**rs17861132**	G	A	G	2/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
61	g.72820629G>A	*CYP1A2* –8504	**rs17861133**	G	A	G	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
62	g.72820857G>A	*CYP1A2* –8276	**rs17861134**	G	A	G	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
63	g.72820915C>T	*CYP1A2* –8218	**rs17861135**	C	T	C	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
64	g.72820962A>G	*CYP1A2* –8171	**rs2472300**	A	G	G*	9/12	9/14	6/8	19/24	13/16	8/8	3/8	9/16
65	g.72821129T>C	*CYP1A2* –8004	**rs17861137**	T	C	T	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
66	g.72821229G>T	*CYP1A2* –7904	**rs17861138**	G	T	G	1/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16
67	g.72821686A>G	*CYP1A2* –7447	**rs17861139**	A	G	?	2/12	1/14	1/8	0/22	0/16	1/8	0/8	2/16
68	g.72821975G>A	*CYP1A2* –7158	**rs17861140**	G	A	G	4/12	0/14	2/8	9/22	0/16	7/8	3/8	5/16
69	g.72822042C>T	*CYP1A2* –7091	**rs17861141**	C	T	T*	2/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16
70	g.72822176G>A	*CYP1A2* –6957	**rs17861142**	G	A	G	2/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16
71	g.72822309T>C	*CYP1A2* –6824	**rs2445618**	T	C	C*	9/12	9/14	6/8	19/24	13/16	8/8	3/8	9/16
72	g.72822437A>G	*CYP1A2* –6696	**rs17861145**	A	G	A	2/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16
73	g.72827329A>G	*CYP1A2* –1804	**rs2069524**	A	G	A	1/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16
74	g.72827425T>C	*CYP1A2* –1708	**rs2069525**	T	C	C*	2/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16

SNP #	SNP Name	Relative Position	NCBI Assignment	A1	A2	Ancestral	P_{A2} in Africans	P_{A2} in Caucasians	P_{A2} in Asians	P_{A2} in Samoans	P_{A2} in New Guineans	P_{A2} in Pehuenche	P_{A2} in Teribe	P_{A2} in Dogrib
75	g.72828394T>G	*CYP1A2* –739	**rs2069526**	T	G	G*	2/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16
76	g.72828400G>C	*CYP1A2* –733	**rs28399417**	G	C	G	0/12	0/14	0/8	2/24	0/16	0/8	0/8	0/16
77	g.72828784C>G	*CYP1A2* –349	**rs17861149**	C	G	C	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
78	g.72828955G>A	*CYP1A2* –178	**rs17861150**	G	A	G	2/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
79	g.72828970C>A	*CYP1A2* –163	**rs762551**	C	A	A*	9/12	9/14	6/8	19/24	13/16	8/8	3/8	9/16
80	g.72829185C>G	*CYP1A2* +53	**rs17861152**	C	G	C	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
81	g.72829381G>T	*CYP1A2* +249	**rs17861153**	G	T	G	0/12	0/14	1/8	0/24	0/16	0/8	0/8	0/16
82	g.72829438C>T	*CYP1A2* +306	**rs17861154**	C	T	C	0/12	1/14	0/8	0/24	0/16	0/8	0/8	0/16
83	g.72829747C>T	*CYP1A2* +615	**rs17861155**	C	T	C	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
84	g.72830334C>T	*CYP1A2* +1202	**rs4646425**	C	T	C	0/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16
85	g.72830564T>C	*CYP1A2* +1432	**ss159100344**	T	C	T	0/12	0/14	0/8	1/24	0/16	0/8	0/8	0/16
86	g.72830645C>A	*CYP1A2* +1513	**rs17861157**	C	A	C	2/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
87	g.72830646G>A	*CYP1A2* +1514	**rs35796837**	G	A	G	0/12	0/14	0/8	2/24	7/16	0/8	0/8	0/16
88	g.72830947C>A	*CYP1A2* +1815	**ss159100345**	C	A	C	0/12	0/14	0/8	0/24	1/16	0/8	0/8	0/16
89	g.72831290G>A	*CYP1A2* +2159	**rs2472304**	G	A	G	0/12	8/14	3/8	3/24	0/16	0/8	0/8	2/16
90	g.72831453G>C	*CYP1A2* +2321	**rs3743484**	G	C	G	0/12	0/14	0/8	7/24	13/16	0/8	0/8	0/16
91	g.72831746G>A	*CYP1A2* +2614	**rs17861159**	G	A	G	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
92	g.72832260G>C	*CYP1A2* +3128	**ss159100346**	G	C	G	0/12	0/14	0/8	5/24	6/16	0/8	0/8	0/16
93	g.72832387C>A	*CYP1A2* +3255	**ss159100347**	C	A	C	0/12	0/14	0/8	1/24	0/16	0/8	0/8	0/16

SNP #	**SNP Name**	**Relative Position**	**NCBI Assignment**	**A1**	**A2**	**Ancestral**	**P_{A2} in Africans**	**P_{A2} in Caucasians**	**P_{A2} in Asians**	**P_{A2} in Samoans**	**P_{A2} in New Guineans**	**P_{A2} in Pehuenche**	**P_{A2} in Teribe**	**P_{A2} in Dogrib**
94	g.72832745T>C	*CYP1A2* +3613	**rs4646427**	T	C	T	1/12	1/14	1/8	0/24	0/16	1/8	0/8	2/16
95	g.72833457T>C	*CYP1A2* +4325	**ss159100348**	T	C	T	0/12	0/14	0/8	0/24	1/16	0/8	0/8	0/16
96	g.72833665A>G	*CYP1A2* +4533	**ss159100349**	A	G	A	0/12	0/14	0/8	0/24	1/16	0/8	0/8	0/16
97	g.72834479C>T	*CYP1A2* +5347	**rs2470890**	C	T	C	0/12	8/14	3/8	2/24	0/16	0/8	0/8	2/16
98	g.72834527C>T	*CYP1A2* +5395	**ss159100350**	C	T	C	0/12	0/14	0/8	0/24	0/16	0/8	0/8	1/16
99	g.72835806C>G	*CYP1A2* +6674	**rs17861162**	C	G	C	2/12	1/14	1/8	7/24	12/14	1/8	0/8	2/16
100	g.72835817A>G	*CYP1A2* +6685	**rs17861163**	A	G	A	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
101	g.72836881G>A	*CYP1A2* +7749	**rs17861164**	G	A	G	1/12	0/14	0/8	0/24	0/16	0/8	0/8	0/16
102	g.72836996G>A	*CYP1A2* +7864	**rs2960192**	G	A	G	0/12	8/14	3/8	2/24	0/16	0/8	0/8	2/16

Frequencies were obtained from resequencing data [current study plus (Jiang et al., 2005)] of 106 chromosomes: 12 African, 14 Caucasian, 8 Asian, 24 Samoan, 16 New Guinean, 8 Pehuenche, 8 Teribe and 16 Dogrib. A1 and A2 refer to the sequence of the reference and the alternative/mutant alleles of each SNP, respectively. The 21 segregating sites (*denoted* in red) represent those where the predominant allele in human differed from that in a nonhuman primate. The ancestral allele at most sites was inferred from the predominant homologous sequence of chimpanzee, except for SNP *CYP1A2* +1815, whose sequence corresponds to that of the orangutan. If the predominant sequence in human differed from that of chimpanzee, we searched for the sequence in orangutan or gorilla (see Supp. Table S6). In those sites where the sequence information was available, the sequence predominant in at least two species was considered ancestral; otherwise, the sequence in chimpanzee was considered ancestral. A total of 19 sites in humans (marked * in the ancestral column) seem to have diverged from the ancestral state. The sequence of SNP *CYP1A2* –7447 seems to be exclusively human and is not present in the other primates. P_{A2} is the frequency of the alternative/mutant allele A2. Relative positions listed are calculated with respect to the translation start codon. To transform these to positions relative to the TSS and for color code, see Supp. Table S1 legend.

SUPP. TABLE S5. Predicted AHRE Motifs in the BDP Region of *CYP1A1_CYP1A2* Locus*

Strand	**Position relative to *CYP1A1* TTS**	**Position relative to *CYP1A2* TSS**	**Position relative to translation start site**	**AHRE sequence (+) strand**	**AHRE sequence (-) strand**	**E- value**	**Reported previously**
+	–386	–22921	*CYP1A1* –2825	GGGCGTGC		0.094	present study
+	–489	–22818	*CYP1A1* –2928	TCGCGTGA		0.041	(Jaiswal et al., 1985; Kubota et al., 1991)
+	–884	–22423	*CYP1A1* –3323	TTGCGTGC		0.018	(Jaiswal et al., 1985; Kubota et al., 1991)
+	–972	–22335	*CYP1A1* –3411	TTGCGTGA		0.036	(Jaiswal et al., 1985; Kubota et al., 1991)
–	–1053	–22261	*CYP1A1* –3492	TCACGCGA	TCGCGTGA	0.041	(Jaiswal et al., 1985; Kubota et al., 1991)
+	–1371	–21936	*CYP1A1* –3810	TTGCGTGC		0.018	(Jaiswal et al., 1985; Kubota et al., 1991)
–	–2114	–21200	*CYP1A1* –4553	TCACGCCA	TGGCGTGA	0.084	(Jaiswal et al., 1985; Kubota et al., 1991)
+	–3065	–20242	*CYP1A1* –5504	TGGCGTGA		0.084	(Kubota et al., 1991)
+	–5601	–17706	*CYP1A1* –8040	TTGCGTGA		0.036	(Corchero et al., 2001)
+	–7504	–15803	*CYP1A1* –9943	GTGCGTGA		0.075	present study
–	–12210	–11097	*CYP1A2* –11993	TCACGCCA	TGGCGTGA	0.084	(Corchero et al., 2001)
+	–12381	–10926	*CYP1A2* –11822	TGGCGTGA		0.084	(Corchero et al., 2001)
+	–15058	–8249	*CYP1A2* –9145	TAGCGTGC		0.061	(Corchero et al., 2001)

Strand	**Position relative to *CYP1A1* TTS**	**Position relative to *CYP1A2* TSS**	**Position relative to translation start site**	**AHRE sequence (+) strand**	**AHRE sequence (-) strand**	**E- value**	**Reported previously**
–	–18606	–4701	*CYP1A2* –5597	TCACGCCA	TGGCGTGA	0.084	(Corchero et al., 2001)
–	–20813	–2494	*CYP1A2* –3390	ACACGCAC	GTGCGTGT	0.089	(Quattrochi and Tukey, 1989)

*Location (TSS = transcription start site) and characteristics of 15 AHREs predicted by the software MAPPER. The entire sequence of the BDP of the *CYP1A1_CYP1A2* locus (23,306 bp) was considered for the query, whilst model M00235 from Transfac (TTGCGTGC) provided the consensus sequence for the AHRE, *i.e.* binding-site of the AHR-ARNT heterodimer. The software found 15 hits having E-values below 0.10 and 27 hits with E-values >0.10 (the latter not shown). E-values refer to the statistical significance of hits, *i.e.* the probability of finding a “good hit” by chance alone in a non-homologous sequence.

SUPP. TABLE S6. Sequences of 21 Divergent Segregating Sites (between Human and Chimpanzee, Orangutan or Gorilla)

				Human						
SNP #	**SNP Name**	**Relative Position**	**NCBI Assignment**	**A1**	**A2**	**Chimpanzee**	**Orangutan**	**Gorilla**	**Ancestral**	**Commentary**
3	g.72798264A>G	*CYP1A1* +4228	**rs17861082**	A	G	G	T		–	Recurrent mutation
6	g.72799170C>A	*CYP1A1* +3322	**rs17861085**	C	**A**	**A**	A		A	
8	g.72799260A>G	*CYP1A1* +3232	**rs4986884**	A	G	G	–	T	–	Recurrent mutation
9	g.72799524T>C	*CYP1A1* +2968	**rs4986882**	T	**C**	**C**	C		C	
10	g.72799540G>C	*CYP1A1* +2952	**rs4986881**	G	**C**	**C**	–	C	C	
27	g.72805943A>C	*CYP1A1* –3452	**rs3826041**	A	**C**	**C**	C		C	
31	g.72806896T>C	*CYP1A1* –4405	**rs7495708**	T	**C**	**C**	C		C	
39	g.72812040C>T	*CYP1A1* –9549	**rs17861114**	C	T	T*	C		C	
42	g.72812867T>C	*CYP1A1* –10376	**rs12441817**	T	C	C*	T		T	
43	g.72813041C>T	*CYP1A1* –10550	**rs4886605**	C	**T**	**T**	T		T	
50	g.72819195A>T	*CYP1A2* –9938	**rs17861124**	A	**T**	**T**	T		T	
51	g.72819208G>A	*CYP1A2* –9925	**rs17861125**	G	A	A	–		–	Exclusively human/chimp
52	g.72819500C>T	*CYP1A2* –9633	**rs11632706**	C	**T**	**T**	T		T	
53	g.72819544G>A	*CYP1A2* –9589	**rs11632814**	G	**A**	**A**	A		T	
59	g.72820453A>G	*CYP1A2* –8680	**rs2472299**	A	**G**	**G**	G		G	
64	g.72820962A>G	*CYP1A2* –8171	**rs2472300**	A	**G**	**G**	G		G	

				Human						
SNP #	**SNP Name**	**Relative Position**	**NCBI Assignment**	**A1**	**A2**	**Chimpanzee**	**Orangutan**	**Gorilla**	**Ancestral**	**Commentary**
69	g.72822042C>T	*CYP1A2* –7091	**rs17861141**	C	T	T	**–**		**–**	Exclusively human/chimp
71	g.72822309T>C	*CYP1A2* –6824	**rs2445618**	T	**C**	**C**	C		C	
74	g.72827425T>C	*CYP1A2* –1708	**rs2069525**	T	**C**	**C**	C		C	
75	g.72828394T>G	*CYP1A2* –739	**rs2069526**	T	**G**	**G**	G		G	
79	g.72828970C>A	*CYP1A2* –163	**rs762551**	C	A	A	**–**		**–**	Exclusively human/chimp

A1 is the reference/predominant human allele and A2, the alternative/mutant one. There were 21 segregating sites, in which the predominant allele was different between human and chimpanzee (divergent sites). These were further analyzed to try to define their ancestral state**;** this was done by searching for the predominant allele in orangutan and, when the sequence was not present in this species, we also searched in gorilla. The ancestral state was chosen from the sequence predominant in two species**;** sites where it was not possible to ascertain an ancestral state were either found exclusively in human and chimp, or apparently have undergone recurrent mutations in nonhuman primates. In most cases (14 SNPs), humans diverged from the ancestral state**;** however, there were two exceptions (*****), in which chimpanzees diverged from the ancestral state.