For reprint orders, please contact:
reprints@futuremedicine.com

Global pharmacogenetics: genetic substructure of Eurasian populations and its effect on variants of drug-metabolizing enzymes

Oksana Makeeva[1], Vadim Stepanov[1], Valery Puzyrev[1] David B Goldstein[2] & Iris Grossman[2,3†]

†Author for correspondence
[1]Tomsk Research Institute of Medical Genetics of The Russian Academy of Medical Sciences, 10 Nab. Ushaiky, Tomsk, 634050 Russia
Tel.: +7 382 251 2228;
Fax: +7 382 251 3744;
E-mail: oksana.makeeva@medgenetics.ru
[2]IGSP Center for Population Genomics & Pharmacogenetics, Duke University, Durham, NC, USA
Tel.: +1 919 684 0896;
Fax: +1 919 668 6787
[3]Pharmacogenetics, Research and Development, GlaxoSmithKline, 5 Moore Drive, Research Triangle Park, Durham 27709, NC, USA
Tel.: +1 919 483 8006;
Fax: +1 919 315 4174;
E-mail: iris.x.grossman@gsk.com

Aims: To study the frequency distribution of cytochrome P450 (CYP) functional genetic variants in five Eurasian populations from the territory of Siberia in Russia.
Materials & Methods: Unrelated healthy Tuvinians, Buryats, Altaians, Yakuts and Russians (n = 87–88) were genotyped for *CYP2C9*2, CYP2C9*3, CYP2C19*2, CYP2C19*3, CYP3A5*3* and CYP3A5*6. Standard pairwise genetic distances, locus-specific and global Fst statistics were calculated. **Results:** CYP allele and genotype frequencies demonstrated significant variability. Overall, the degree of between-population variance displayed by CYP SNPs was lower than that recorded from neutral short tandem repeats and Alu-insertion polymorphism, indicating evolutionary conservation of CYP polymorphisms. CYP-based genetic distances were well correlated with the geographic distances across populations (r = 0.822, p = 0.008). **Conclusions:** Although the tested variants were present in the neighboring, yet secluded, populations at the expected range of frequencies, the observed frequencies were significantly variable across Eurasian populations, indicating potential relevance to clinical decision making.

Keywords: Alu insertion, Caucasus populations, drug-metabolizing enzymes, P450, personalized medicine, pharmacogenetics, SNP

The appropriate class, dose and treatment regimen that lead to optimal therapeutic safety and efficacy are presumed to be governed, at least partially, by genetic determinants. Thus, it is expected that, in the future, genetic markers will serve as diagnostics, informing and guiding clinical decision making. Based on this principle, pharmacogenetic research is aimed at identifying individuals who are expected to benefit from a certain pharmacotherapy, while dramatically reducing the suffering and costs associated with adverse drug reactions (ADRs). Indeed, it has been stated that most of the commonly used drugs are effective in only 25–60% of patients, and more than 2 million cases of ADRs occur annually in the USA, including 100,000 deaths [1].

From the early years up to contemporary pharmacogenetic research, genes encoding proteins that take part in the pharmacokinetics of drugs have been the focus of research. Among these genes the most prominently studied molecule group has indisputably been the drug-metabolizing enzymes (DMEs), predominantly the cytochrome P450 (CYP) family, since it is responsible for the oxidative metabolism of most drugs administered into the human body. A large body of evidence collected over the previous decades by various independent groups shows a direct impact of functional variations in DMEs on the pharmacokinetics of drugs metabolized by these enzymes [2,3]. The above observations have led many to argue that patients' genotypes with regard to the key DME genes should be incorporated into clinical decision making [4,5]. In fact, several drug labels already include recommendations for genotyping prior to treatment initiation and during follow-up (for example, the anticoagulant warfarin [6]).

Despite the growing progress in pharmacogenetic research and pioneering examples of its impact on daily clinical management, many countries worldwide still lack the appropriate resources for individual genotyping and incorporation of routine pharmacogenetic testing into their national healthcare systems. It has thus been suggested that American- and European-based genetic knowledge would be applicable to the health management of the developing world, until resources are identified for the development of nation-specific pharmacogenetic profiles in these countries [7]. Indeed, many countries rely on the US FDA/European Medicines Agency (EMEA) for safety and dosing guidance, disregarding specific consideration related to non-American and non-European populations. This is a crucial point, as reported allele and genotype frequencies for one population are not necessarily applicable to other, even similarly defined, populations (for example, allele frequencies of the glutathione-*S*-transferase isoenzyme θ gene [*GSTT1*] range from 13–28% amongst Caucasian populations [8]). It is thus essential to

10.2217/14622416.9.7.847 © 2008 Future Medicine Ltd ISSN 1462-2416

evaluate and demonstrate the utility of global genetic profiles in geographically and ethnically distinct populations.

Siberia possesses unique demographic characteristics of massive ethnic heterogeneity, resulting from a legacy of centuries of interface and communication between Caucasoids and Mongoloids. The disproportionate distribution of small populations across an extensive stretch of territory (average population density of approximately three individuals per square kilometer) lead some autochthonous groups to remain anthropologically, linguistically and genetically different from each other. Most of the modern population in Siberia is composed of Russians (Slavs), while the remaining indigenous populations, including Buryats, Tuvinians, Yakuts and Southern Altaians, belong to Mongoloid and Turkic-speaking groups, who have retained their native language and ethnic identity throughout history, resulting in minimal interethnic admixture. To this end, it is customary in the scientific literature to relate to Eurasian populations as roughly a unified entity, without distinction between subpopulations that might be subject to unique genetic pressures resulting in differential genetic profiles and, thus, potentially requiring uniquely tailored pharmacotherapies. We therefore set out to investigate whether known clinically relevant variations in DMEs from the CYP family exhibit unique genetic properties in five distinct Eurasian populations.

Materials & methods

Study population

A total of 438 blood samples from apparently healthy, unrelated individuals were collected from the following five autochthonous groups from the territory of Siberia in Russia: Southern Altaians (n = 87) residing in the village of Beshpeltir of the Altai republic; Tuvinians (n = 88) residing in Kyzil; Buryats (n = 88) from the Okinski and Kizhinginsky districts of the Buryatia Republic; Yakuts (n = 88) from the village of Cheriktei in the Sakha (Yakutia) Republic; and Russians (n = 87) residing in Tomsk (a city in which approximately 90% of the population is Russian). A map of Russia indicating the geographical residence of each of the study's populations is presented in Figure S1 in the Supplementary Material. DNA samples were obtained from the DNA bank of the Research Institute of Medical Genetics in Tomsk, Russia. Blood samples were collected from volunteers throughout the last 15 years. Equal proportions of men and women in each population were selected for the current study. In order to ensure the lack of cryptic relatedness amongst individuals, only samples from unrelated individuals as depicted in genealogical analysis of at least the last three generations were selected for genotyping. The study was approved by the Institutional Review Board of the Research Institute of Medical Genetics (Tomsk, Russia). Written informed consent was obtained from the volunteers before blood sampling.

Genotyping

Total genomic DNA was extracted from whole blood using the perchlorate/chloroform method [9]. DNA quantity and quality were assessed spectrophotometrically. All SNP variants were genotyped by TaqMan® fluorescence-based allelic discrimination [10]. Premade assays offered as Assay-by-Demand were used when available; otherwise primers were designed using the Applied Biosystems (ABI; CA, USA) Assay-by-Design tool. PCRs were carried out using the standard ABI protocol for a 5 µl reaction volume. Fluorescence spectra were detected by use of the 7900HT Fast Real Time PCR System, and the data was analyzed using the SDS software, version 2.2.2 (ABI). All genotype calls were determined independently by two researchers (Makeeva O and Grossman I), and ambiguous calls were regenotyped or discarded. Genotyping failure rates were approximately 1%. *Alu* and short tandem repeat (STR) genotyping assays and results have been published elsewhere [11–15] (available in English upon request). Characteristics of all polymorphic sites reported in this paper are provided in Table S1–S4 of the Supplementary Material.

Statistical analysis

Allele and genotype frequencies were calculated by direct counting, and Hardy–Weinberg equilibrium was evaluated by an exact test [16], as implemented in Arlequin version 3.1. [17]. Genetic variants for which deviations from Hardy–Weinberg equilibrium were identified were excluded from further analysis. Pairwise allele frequency comparisons were tested by Fisher's exact test and the Bonferroni correction was used for multiple tests. Pairwise and locus-specific Fst values and corresponding p-values were computed using Arlequin version 3.1. [18–20]. The Mantel test (10,000 permutations) [21] was implemented to evaluate the correlations between the Nei's standard pairwise

genetic distances based on CYP SNPs data, with the corresponding distance in kilometers between each of the population locations. Population relationships were also analyzed via principal component analysis, as implemented in STATISTICA 6.0. (StatSoft Inc., OK, USA, 2001).

Results

Allele & genotype frequency distribution

Allele and genotype frequencies observed in Tuvinians, Buryats, South Altaians, Yakuts and Russians are presented in Table 1. Upon testing for Hardy–Weinberg equilibrium, no departures from equilibrium were detected. The *CYP3A5*6* allele was not detected in this study, in accordance with reports for other Caucasoid populations worldwide. Analysis of the remaining five CYP functional SNPs [101] revealed a significantly high level of diversity across Siberian populations. Results for the corresponding pairwise comparisons of allele frequencies are summarized in Table 2.

As expected, Russians show a high similarity in allele frequencies of all the five polymorphic CYP variants to other European populations, exhibiting intermediate frequencies compared

Table 1. Genotype and allele frequencies for *CYP450* gene polymorphisms in five Siberian populations‡.

Population/ CYP polymorphism	N	Genotype			Allele frequencies (%)	
		Wt/Wt	*Wt/Vr*	*Vr/Vr*	*Wt*	*Vr*
CYP2C9*2		***CC***	***CT***	***TT***	***C***	***T***
Tuvinians	88	86	2	0	98.9	1.1
Buryats	88	84	4	0	97.7	2.3
Altaians	87	77	10	0	94.3	5.7
Yakuts	88	86	2	0	98.9	1.1
Russians	87	67	19	1	87.9	12.1
CYP2C9*3		***CC***	***CT***	***TT***	***C***	***T***
Tuvinians	88	79	9	0	94.9	5.1
Buryats	88	85	3	0	98.3	1.7
Altaians	87	72	14	1	90.8	9.2
Yakuts	88	87	1	0	99.4	0.6
Russians	74	65	8	1	93.2	6.8
CYP2C19*2		***GG***	***GA***	***AA***	***G***	***A***
Tuvinians	88	63	24	1	85.2	14.8
Buryats	88	54	31	3	79.0	21.0
Altaians	87	64	20	3	85.1	14.9
Yakuts	88	54	27	7	76.7	23.3
Russians	82	64	16	2	87.8	12.2
CYP2C19*3		***GG***	***GA***	***AA***	***G***	***A***
Tuvinians	88	84	4	0	97.7	2.3
Buryats	88	77	10	1	93.2	6.8
Altaians	87	80	7	0	96.0	4.0
Yakuts	87	79	8	0	95.4	4.6
Russians	87	87	0	0	100.0	0.0
CYP3A5*3		***AA***	***AG***	***GG***	***A***	***G***
Tuvinians	87	2	21	64	14.4	85.6
Buryats	87	5	21	61	17.8	82.2
Altaians	85	0	18	67	10.6	89.4
Yakuts	88	0	13	75	7.4	92.6
Russians	83	0	14	69	8.4	91.6

‡CYP3A5*6 *was monomorphic in all the five studied Siberian groups*

N: Number of individuals studied in each population; Vr: Variant; Wt: Wild-type.

Table 2. Pairwise comparison of allele frequencies for *CYP2C9*, *CYP2C19* and *CYP3A5* in the five tested populations (Fisher's exact test).

Population/ CYP polymorphism	Tuvinians (n = 88)	Buryats (n = 88)	Altays (n = 87)	Yakuts (n = 88)	Russians (n = 87)	Locus Fst (p value)
CYP2C9*2						
Tuvinians	–	NS	0.020	NS	**0.00002**	
Buryats	–	–	NS	NS	**0.0003**	0.045 (p = 0.000)
Altays	–	–	–	0.020	NS	
Yakuts	–	–	–	–	**0.00002**	
Russians	–	–	–	–	–	
CYP2C9*3						
Tuvinians	–	NS	NS	0.020	NS	
Buryats	–	–	**0.002**	NS	0.024	0.024 (p = 0.001)
Altays	–	–	–	**0.0001**	NS	
Yakuts	–	–	–	–	**0.003**	
Russians	–	–	–	–	–	
CYP2C19*2						
Tuvinians	–	NS	NS	NS	NS	
Buryats	–	–	NS	NS	0.030	0.010 (p = 0.039)
Altays	–	–	–	NS	NS	
Yakuts	–	–	–	–	**0.011**	
Russians	–	–	–	–	–	
CYP2C19*3						
Tuvinians	–	NS	NS	NS	NS	
Buryats	–	–	NS	NS	**0.0004**	0.014 (p = 0.010)
Altays	–	–	–	NS	**0.015**	
Yakuts	–	–	–	–	**0.007**	
Russians	–	–	–	–	–	
CYP3C5*3						
Tuvinians	–	NS	NS	NS	NS	
Buryats	–	–	NS	**0.004**	**0.016**	0.012 (p = 0.016)
Altays	–	–	–	NS	NS	
Yakuts	–	–	–	–	NS	
Russians	–	–	–	–	–	

Only significant comparisons are presented (p < 0.05). Tests that remain significant after Bonferroni correction are marked in bold (p < 0.0167). NS: Not significant.

with previously published results: *CYP2C9*2*: 12% (ranging from 10% in Belgians to 17% in Croatians); *CYP2C9*3*: 7% (ranging from 5% in British to 13% in Spanish); *CYP2C19*2*: 12% (from 9% in Belgians to 17% in Swedish) (Figure 1–3). The *CYP2C19*3* variant was absent in Russians, as well as in most other European populations (Figure 4). Our findings closely resemble the two previous publications reporting frequency information on the *CYP2C9* and *CYP2C19* polymorphisms in Russians [22,23]. An overall high frequency of the *CYP3A5*3* variant was detected (92%), which is typical in other Europeans (87–94%) (Figure 5).

Highly significant differences in allele frequencies for *CYP2C9*2* were detected between Russians and all native Siberian populations, except for Altaians (Tables 1 & 2): 12% in Russians versus 1% in Tuvinians and Yakuts (p = 0.00002) and 2% in Buryats (p = 0.0003). The frequency of *CYP2C9*3* was highest in Altaians, who had an allelic frequency of 9% of the variant allele, versus 1.7% in Buryats (p = 0.002) and 0.6% in Yakuts (p = 0.0001) (Tables 1 & 2). The *CYP2C19*2* and **3* variants are known to be more common in Asian than in European populations, where the **3* polymorphism is rare or absent (Figure 4, [24–33]). We report here that the *CYP2C19*3* variant

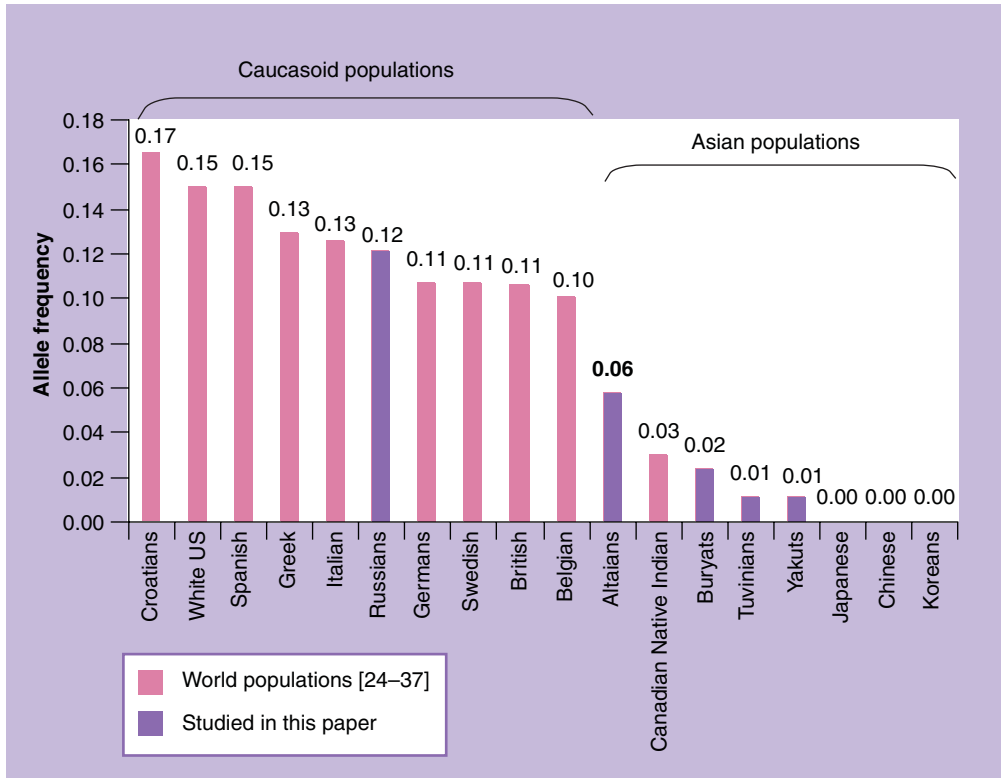
exhibits a frequency of 2.3% in Tuvinians, 4.0% in Altaians, 4.6% in Yakuts and 6.8% in Buryats. All populations, except Tuvinians, significantly differed from Russians, who lacked the *3 polymorphism allele ($p < 0.015$, Table 2). No significant differences were found between indigenous Siberian populations in *CYP2C19*2* frequency (15–23%), while Russians had lower *CYP2C19*2* frequency (12%) when compared with Yakuts (23%, $p = 0.011$) and Buryats (21%, $p = 0.030$). The frequency of the variant *CYP3A5*3* allele varied from 82% in Buryats to 93% in Yakuts.

To further analyze the CYP polymorphism frequency distribution in our five Eurasian populations we compared the genetic versus geographic distances between each pair of populations. Two matrices of pairwise genetic (Nei's standard genetic distance) and geographic distances (in km) were obtained, and their correlation was evaluated using Mantel test statistics (Figure 6). The analyses revealed that the CYP-based genetic distances exhibit a high and significant correlation with geographic distances between populations with a Mantel correlation coefficient, $r = 0.822$, $p = 0.008$.

Genetic relationships of Siberian populations amongst themselves & across other Caucasoid & Asian populations: principal component analysis

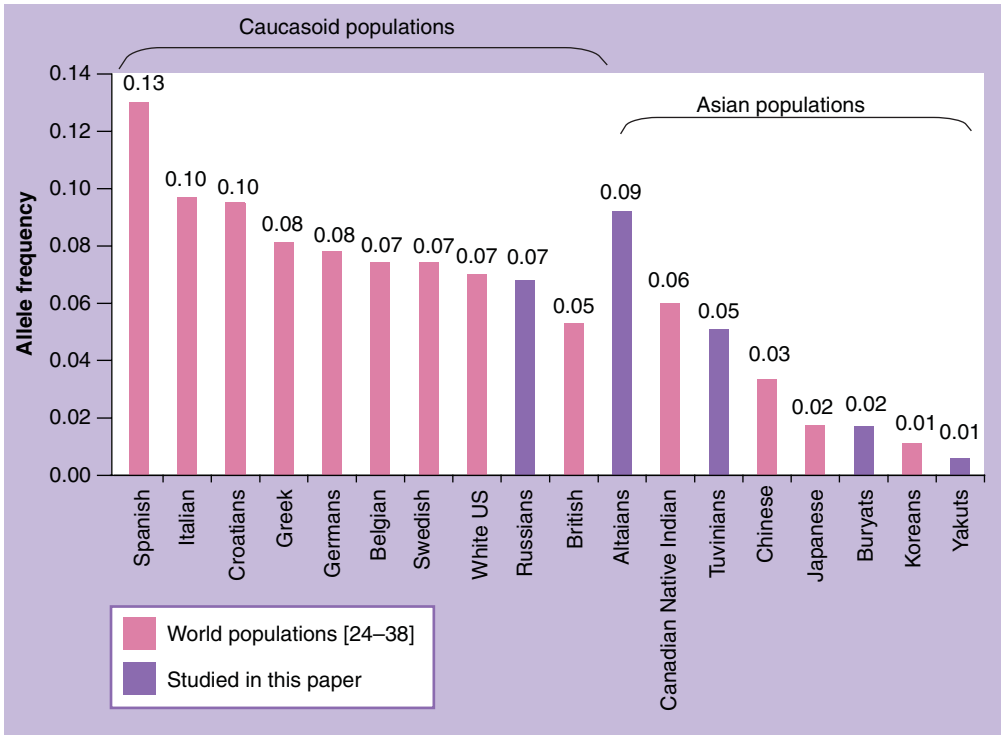
We assessed a genetic relationship among the populations under study, as well as their relatedness to other world populations (both Caucasoid and Asian) through principal component analysis (PCA; Figure 7). Analysis of principal components showed a very compact cluster of several European populations, with Russians positioned in the middle. Native Siberian populations showed proximity to Europeans (especially Altaians and Tuvinians) when compared with other Asian populations analyzed (Koreans, Chinese and Japanese). Two principal components accounted for 99.7% of the genetic variance observed in CYP SNPs frequency.

Figure 1. *CYP2C9*2* allele frequencies in worldwide Caucasoid and Asian populations.



Tables S5–S9 in the Supplementary Material lists the source population sizes, reported allele frequencies and references. When several independent papers reported on the same population, the data was pooled together and calculated based on the total counts.

Figure 2. *CYP2C9*3* allele frequencies in worldwide Caucasoid and Asian populations.



Tables S5–S9 in the Supplementary Material lists the source population sizes, reported allele frequencies and references. When several independent papers reported on the same population, the data was pooled together and calculated based on the total counts.

Fst statistics: CYP SNPs versus Alu repeats & STR polymorphisms

To examine whether the functional CYP450 variant frequency spectrum was shaped by selection forces distinct from the evolution of neutral genetic regions, we calculated pairwise, locus-specific and global (over all studied SNPs) Fst statistics based on the above tested functional CYP450 variants, six neutral STR and seven randomly selected Alu polymorphisms in a subset of the samples. Distances of each Alu and STR marker from known genes are presented in Tables S3 & S4 of the Supplementary Materials (their frequency data was published previously [11–15], and is available in English upon request). Alu and STR polymorphisms are widely used in population genetics and evolutionary studies, especially when both nonfunctional systems are analyzed simultaneously: Alu repeats exhibit a low mutation rate (stable evolution) and always display an ancestral state of an absence of insertion. STRs, on the other hand, are characterized by a high mutation rate and multiple alleles, useful in differentiation between closely related populations.

Table 3 summarizes single-locus and global Fst estimates for the CYP, Alu and STR polymorphisms. Fst single-locus values for the CYP polymorphisms ranged from 0.010 for *CYP2C19*2* to 0.045 for *CYP2C9*2*. The global Fst for the 5 CYP polymorphisms was 0.021; whereas for STRs and Alu insertions it was 0.050 and 0.095, respectively. Thus, Fst estimates demonstrate a low level of between-population diversity (and probably fixation of allele frequencies) in CYP450 functional SNPs.

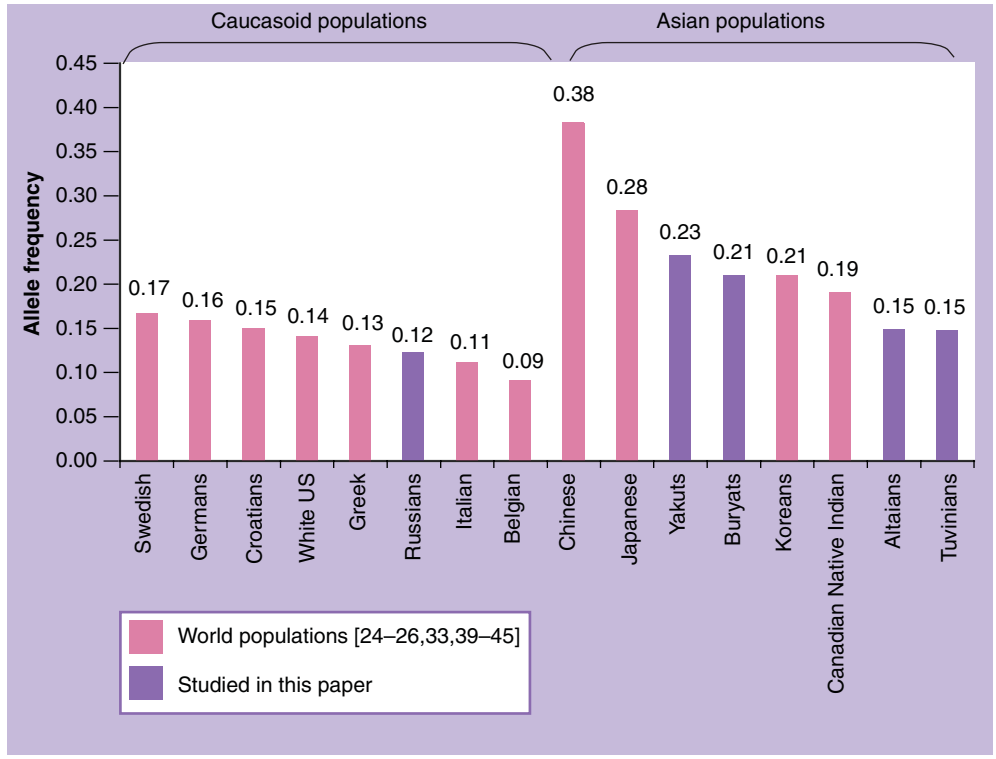
Pairwise Fsts results show that for the CYP450 genes, only 18 (36%) out of 50 pairwise tests (five populations and five SNPs) were significant ($p < 0.05$), with Fst values of 0.017–0.088 (Table 4). In comparison, Fst values for Alu polymorphisms found 38 (54%) out of 70 implemented pairwise tests to be significant, exhibiting higher average Fst values (up to 0.423, Table 5). Pairwise Fst calculations over six neutral STRs were significant, and showed the highest differentiation between Yakuts and Buryats (0.081) and Yakuts and Russians (0.073, Table 6).

Discussion

We report for the first time, to the best of our knowledge, the distribution of pharmacogenetically relevant allele and genotype frequencies in several native Siberian populations. The subset of polymorphic alleles that have been chosen for the analysis was based on recommendations by the FDA and other working groups for measurement in Caucasians and Asians [34]. We focused on this area of the globe as it displays unique demographic characteristics of large ethnic heterogeneity, along with low population density, which allows many autochthonous groups to remain genetically different from each other. The latter attributes have lead researchers to question whether such subpopulations can be adequately represented by large umbrella projects, such as the HapMap [35], claiming to account for most of the common genetic variation worldwide based on four representative populations (i.e., Caucasians, Africans, Japanese and Chinese). While multiple groups [36,37] have proven the generalizablity of HapMap-selected tag SNPs, the current report is the first to evaluate the generalizability and transferability of functional genetic variants in clinically relevant DMEs among subpopulations residing in close geographic proximity to each other. Evolutionary forces acting upon functional polymorphisms are expected to behave differently to those shaping other properties of the genome [38], and thus deserve a special focus before ascertainment of clinical applicability through programs such as the PharmacoGenetics for Every Nation Initiative (PGENI) [7]. This was launched with the purpose of constructing a complete database on gene variants that affect the efficacy or toxicity of drugs, together with information on frequencies of these variants in world populations (especially outside the USA and Europe). This database is expected to be useful to international organizations, such as the WHO, in targeting treatment recommendations to specific populations [7].

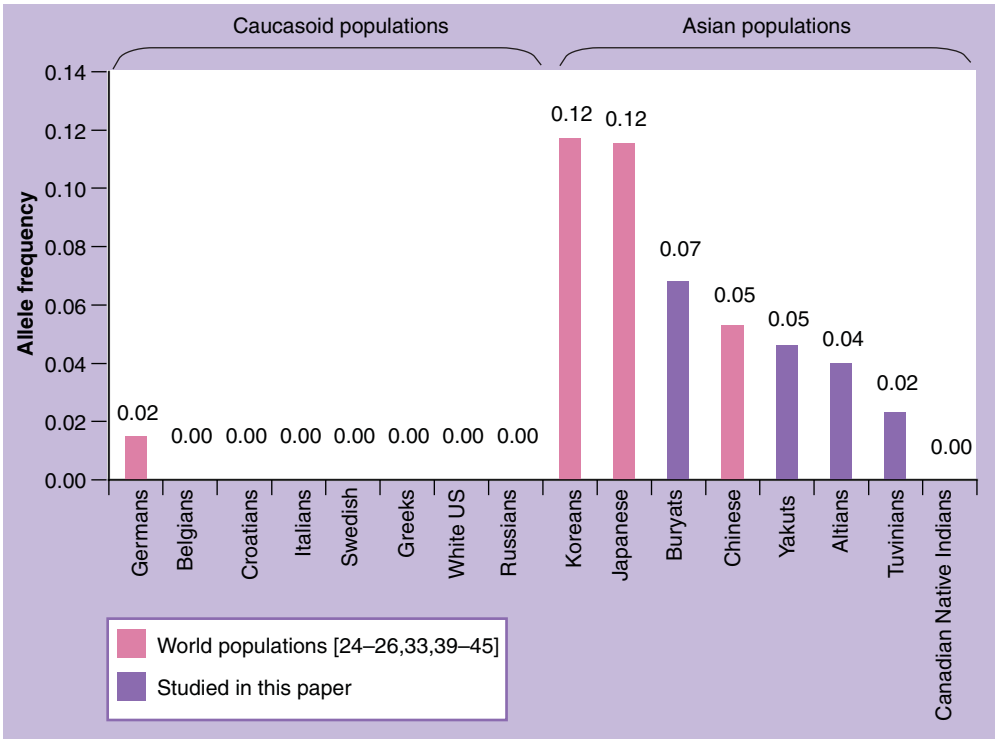
Comprehensive comparison of frequencies between the current investigated samples and other populations throughout the world

Figure 3. *CYP2C19*2* allele frequencies in worldwide Caucasoid and Asian populations.



Tables S5–S9 in the Supplementary Material lists the source population sizes, reported allele frequencies and references. When several independent papers reported on the same population, the data was pooled together and calculated based on the total counts.

Figure 4. *CYP2C19*3* allele frequencies in worldwide Caucasoid and Asian populations.



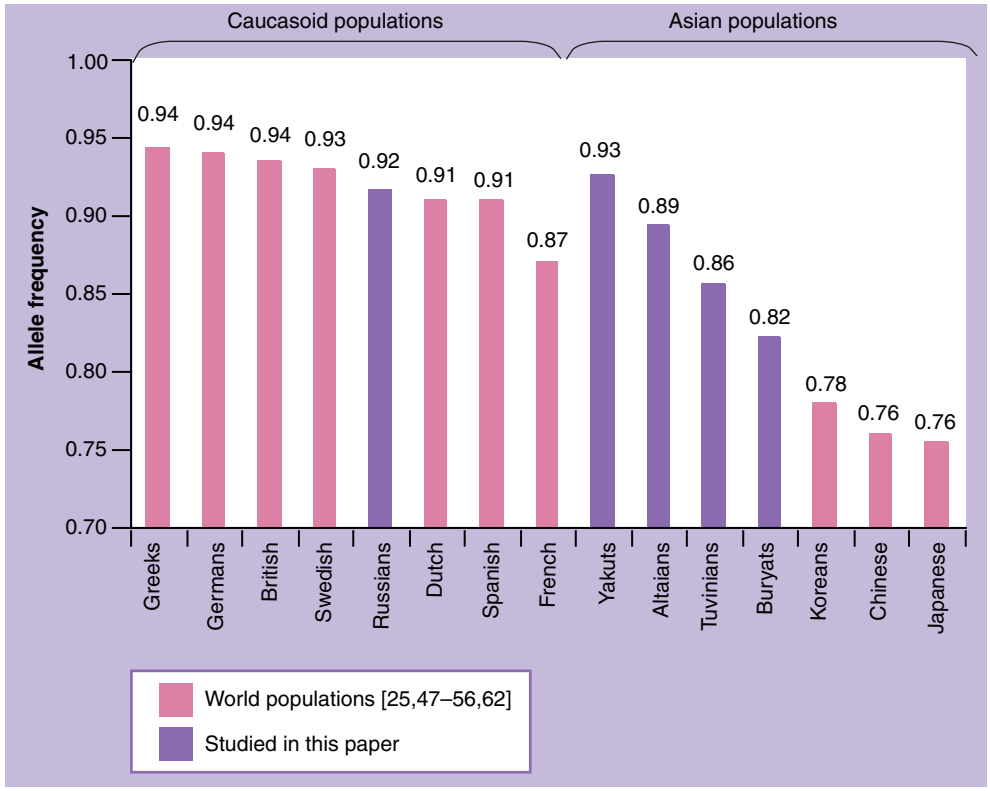
Tables S5–S9 in the Supplementary Material lists the source population sizes, reported allele frequencies and references. When several independent papers reported on the same population the data was pooled together and calculated based on the total counts.

indicates that, as expected, Russians were very close to other Europeans in allele frequencies of all the studied CYP450 polymorphisms. At the same time, some highly significant differences between Russians and indigenous Siberian populations were observed in allele frequencies of *CYP2C9*2*, *CYP2C9*3*, *CYP3A5*3*, *CYP2C19*2* and *CYP2C19*3* (Tables 1 & 2). In general, indigenous Siberian populations (Tuvinians, Buryats, Southern Altaians and Yakuts) exhibited allele frequency patterns characteristic of Asian populations (Korean, Chinese, Japanese and native Americans, who originated from Siberia, as was shown by many studies): high frequency of *CYP2C19*2* (from 15% in Tuvinians and Altaians to 23% in Yakuts) and low frequency of *CYP2C19*3* (up to 7% in Buryats). However, Siberian populations remained closer to Europeans when compared with Chinese and Japanese, as was demonstrated by pricipal component analysis (Figure 7). The Mantel test revealed that CYP-based genetic distances exhibit a significant correlation with geographic distances.

It was previously shown that some of the genes of DMEs are likely targets of selective pressure, and thus show unusual patterns of variation frequency across human populations. This was specifically demonstrated for the *CYP3A5* locus [39]. The *CYP3A5*3* variant affects the conversion of cortisol to 6β-hydroxicortisol in the kidney, leading to higher reabsorption of sodium and water retention (and thus connected with the clinical phenotype of salt-sensitive hypertension). This functional effect confers a selective advantage when populations experience water shortages [40].

Fixation indicies (F) statistics are used to assess differentiation among populations, as depicted by signatures in patterns of genetic variation [41–43]. Estimation of neutral Fst values was carried out based on a set of seven unlinked Alu repeats and six neutral STR loci (Table 3). When the mode of selection is local adaptation, populations with similar selective regimes will acquire similar allele frequencies (yielding low Fst), whereas in others selection will drive allele frequencies apart (yielding high Fst). Global Fst

Figure 5. *CYP3A5*3* allele frequencies in worldwide Caucasoid and Asian populations.



Tables S5–S9 in the Supplementary Material lists the source population sizes, reported allele frequencies and references. When several independent papers reported on the same population, the data was pooled together and calculated based on the total counts.

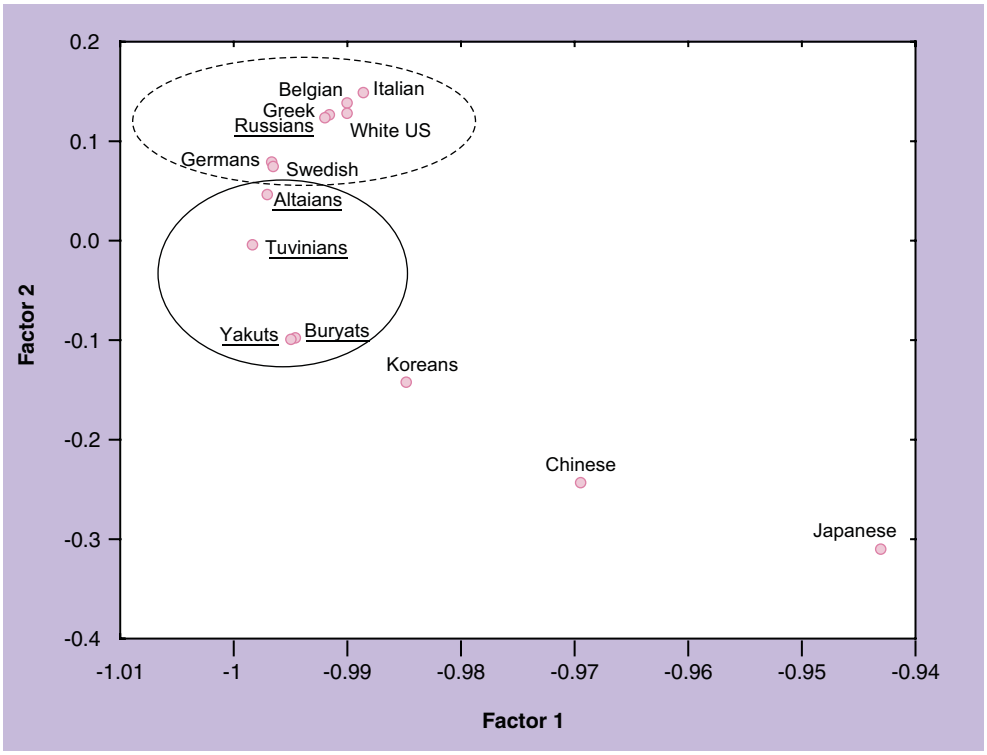
estimations demonstrated low between-population diversity (and probably fixation of allele frequencies) for CYP450 functional SNPs versus neutral STRs and Alu repeats (Fst = 0.021, 0.050, and 0.095, respectively). Consequently, pairwise Fst calculations detected larger variability between pairs of populations when analyzing randomly selected Alu repeats in comparison to analysis of functional variants in CYP genes (Tables 4–6).

Figure 6. Matrices of genetic (standard Nei's) and geographic (km) distances between the studied populations.

Matrix of geographic distances (km)		Tuvinians	Buryats	Altaians	Yakuts	Russians	Matrix of genetic distances (standard Nei's)
	Tuvinians	0	0.0018	0.0012	0.0034	0.0039	
	Buryats	1065	0	0.004	0.0025	0.0078	
	Altaians	617	1660	0	0.0039	0.0017	
	Yakuts	2514	1750	2960	0	0.007	
	Russians	3700	4538	3181	4880	0	

Matrix of genetic distances (Nei's standard genetic distance) – above the diagonal.
Matrix of geographic distances (km) – below the diagonal.
Significance of testing – 10,000 permutations.
The Mantel test indicated that there is a significant correlation between the pairwise genetic and geographic distance matrices with a correlation coefficient = 0.822, p = 0.008.

Figure 7. Plot of the first two principal components resulting from analysis of the CYP gene allele frequencies in 14 world populations.



Populations studied in this paper are underlined. Table S10 in the Supplementary Material lists the allele frequency input data used in this analysis

Multiple independent reports indicating the potential clinical utility of CYP functional variant screening for the purpose of blood-level monitoring and patient response prediction have been published in recent decades. In this report we focused on *CYP2C9*, *CYP2C19* and *CYP3A5*, for which common deficient alleles have been described, that have prominent potential to convey significant clinical utility. CYP2C9 hydroxylates approximately 16% of drugs in current clinical use [44]. Of special interest are those with a narrow therapeutic index, such as *S*-warfarin, tolbutamide and phenytoin, where impairment in CYP2C9 metabolic activity may require dose adjustment, as well as lead to severe toxicity. CYP3A5 has been indicated to have clinical relevance related to a range of therapeutics, including imatinib (an anticancer drug) and tacrolimus [45] (an immunodepressant). *CYP2C19* poor metabolizers, comprising 3–4% of Caucasians and African–Americans, and 14–21% of Asians, may require dose adjustment for some tricyclic antidepressants, moclobemide, citalopram, diazepam and omeprazole [46]. The clinical outcome of duodenal ulcer treated with proton pump inhibitor-based, anti-*Helicobacter pylori* regimens have been reported to vary according to *CYP2C19* genotype as well [47]. The FDA has recently approved, for the first time, a genetic test, the AmpliChip CYP450 Test, prompting the *CYP2C19* (and *CYP2D6*) gene variants to be prime candidates for diagnostic pharmacogenetic testing in clinical practice [46]. However, it should be mentioned, that there is currently not enough evidence to prove the utility of these tests in the clinical practice of several classes of antipsychotics and antidepressants [48,49].

Future perspective

The promise and utility of pharmacogenetic research to daily clinical care has been recognized by regulatory agencies both in the USA and in Europe. These developments are currently being integrated into recommendations and warnings included in the package inserts of marketed therapeutic formularies, as well as report recommendations and requirements for newly developed drugs in research and development phases [50,51]. However, these proceedings apply to American

Table 3. Per locus and global Fst-values in the three studied marker systems (CYP polymorphisms, randomly selected Alu insertions and neutral STR polymorphisms).

Locus name	Fst-values	p-value
CYP gene polymorphisms		
*CYP2C9*2*	0.045	0.000
*CYP2C9*3*	0.025	0.001
*CYP2C19*2*	0.010	0.039
*CYP2C19*3*	0.014	0.010
*CYP3A5*3*	0.012	0.016
Global Fst over all CYP SNPs	0.021	–
Randomly selected Alu insertion polymorphisms		
A25	0.005	0.066
ACE	0.033	0.000
APOA1	0.067	0.000
CD4	0.161	0.000
F13B	0.166	0.000
PLAT	0.020	0.000
PV92	0.217	0.000
Global Fst over all Alu insertion polymorphisms	0.095	–
Neutral STR loci		
D4S397	0.031	0.000
D5S393	0.038	0.000
D8S514	0.074	0.000
D9S161	0.057	0.000
D11S1358	0.049	0.000
D13S173	0.053	0.000
Global Fst over all STR loci	0.050	–

and European countries only, creating the need for assessment of the impact and utility of these new tests in the rest of the world, where independent genotyping is currently impractical and financially prohibitive.

It should be stressed out that the proven utility of genotyping results need to be tested in clinical settings [51], and in the final decision-making process, cost–effectiveness must also be considered. Under the assumption that such global

Executive summary

Pharmacogenetics may serve as a globally applicable diagnostic tool

- It has been extensively demonstrated that there is a high frequency variability in functional CYP gene variants across various populations and ethnic groups.
- It is important to collect information on these clinically relevant pharmacogenetics variants in different areas of the globe.
- This genetic profile can assist healthcare organizations in the prioritization and selection of essential medicines on a national level, before individual pharmacogenetic testing will become feasible.
- Implementation of global pharmacogenetic plans, such as the PharmacoGenetics for Every Nation Initiative (PGENI; http://pgeni.unc.edu/), at this time can promote drug treatment safety and efficacy throughout the world.

Frequency distribution of clinically relevant CYP polymorphisms in different populations

- While studying CYP polymorphism frequency distribution in five Eurasian populations, it has been shown that CYP-based genetic distances are well correlated with their geographic distances.
- However, between-population variability was lower when compared with neutral genetic variants, and thus some fixation of the CYP alleles may be assumed, which can be due to selective pressure acting on these specific genome regions.
- Interpretation of drug-metabolizing enzymes genotyping results in one population cannot rely solely on screening of another representative population, but rather necessitates further exploration in subpopulations, specifically where variants are monomorphic in one population but not in others (as in the case of *CYP2C19*3*).

pharmacogenetic profiles may be ultimately assembled, which capture the majority of clinically relevant genetic variation applicable to the entire world population, initiatives such as PGENI [7] have been launched. Our findings indicate that, although allele frequencies may differ substantially across Eurasian subpopulations, adopting American/European-based pharmacogenetic profiles may indeed be of adequate utility as surrogates to individual genotype screening in the present day. We therefore believe that implementation of global pharmacogenetics plans at this time can promote drug treatment safety and efficacy throughput the world, laying the foundations for routine incorporation of pharmacogenetics testing into daily clinical management concerning all regions of the globe.

Acknowledgements

We thank M Golubenko, V Tadinova, N Maksimova, K Puzyrev, A Rudko, I Khitrinskaia, M Spiridonova, V Kharkov and E Eremina for their assistance in assembling the population collections analyzed in this study. We thank the reviewers for their insightful comments and suggestions throughout the review process. Special thanks are extended to the individuals who volunteered anonymous blood samples.

Financial & competing interests disclosure

This work was supported by the Russian Agency for Science and Innovations (#02.444.11.7008 and #02.444.11.7282) and by the IGSP Center of Population Genomics and Pharmacogenetics, Duke University, USA. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Table 4. Pairwise Fst values among the five Siberian populations tested within the three studied CYP genes.

Population	Locus name	Buryats	Altays	Yakuts	Russians
Tuvinians	*CYP2C9*2*	0	0.026*	0	0.088***
	*CYP2C9*3*	0.012	0.007	0.0318*	0
	*CYP2C19*2*	0.008	0	0.017	0
	*CYP2C19*3*	0.018	0	0.002	0.017
	*CYP3A5*3*	0	0.001	0.019*	0.011
Buryats	*CYP2C9*2*	–	0.010	0	0.064***
	*CYP2C9*3*	–	0.048**	0	0.027*
	*CYP2C19*2*	–	0.007	0	0.022
	*CYP2C19*3*	–	0.002	0	0.062**
	*CYP3A5*3*	–	0.015	0.043**	0.017*
Altays	*CYP2C9*2*	–	–	0.026*	0.187*
	*CYP2C9*3*	–	–	0.072***	0
	*CYP2C19*2*	–	–	0.017	0
	*CYP2C19*3*	–	–	0	0.035*
	*CYP3A5*3*	–	–	0	0
Yakuts	*CYP2C9*2*	–	–	–	0.088***
	*CYP2C9*3*	–	–	–	0.051**
	*CYP2C19*2*	–	–	–	0.035*
	*CYP2C19*3*	–	–	–	0.040*
	*CYP3A5*3*	–	–	–	0

*$p < 0.05$

**$p < 0.01$

***$p < 0.0001$

A total of 18 (36%) out of 50 pairwise Fst tests were significant ($p < 0.05$), where Fst values varied from 0.017 to 0.088. A total of 17 out of 50 (34%) pairwise tests Fst = 0.

Table 5. Pairwise Fst values among the five Siberian populations tested within randomly selected Alu insertion polymorphisms.

Population	Locus name	Buryats	Altays	Yakuts	Russians
Tuvinians	*A25*	0	0.012	0.018*	0
	ACE	0.025***	0.036***	0.099***	0
	APOA1	0.075***	0.041***	0.077***	0.123***
	PLAT	0.005	0.017*	0.042***	0.046***
	F13B	0	0.022*	0	0.255***
	PV92	0.001	0.034**	0	0.363***
	CD4	0	0.009	0.016*	0.224*
Buryats	*A25*	–	0.002	0.007	0
	ACE	–	0	0.025*	0.030**
	APOA1	–	0.001	0	0.005
	PLAT	–	0	0.011	0.015*
	F13B	–	0.035**	0	0.278***
	PV92	–	0.057***	0	0.420*
	CD4	–	0.022***	0.006	0.220***
Altays	*A25*		–	0	0.004
	ACE	–	–	0.014	0.042**
	APOA1	–	–	0.003	0.025*
	PLAT	–	–	0	0.002
	F13B	–	–	0.029*	0.141***
	PV92	–	–	0.056***	0.123***
	CD4	–	–	0.049**	0.234***
Yakuts	*A25*	–	–	–	0.010
	ACE	–	–	–	0.110***
	APOA1	–	–	–	0
	PLAT	–	–	–	0
	F13B	–	–	–	0.261***
	PV92	–	–	–	0.423***
	CD4	–	–	–	0.237***

*$p < 0.05$
**$p < 0.01$
***$p < 0.0001$
38 (54%) out of 70 pairwise Fst tests were significant ($p < 0.05$), where Fst values varied from 0.017 to 0.423. A total of 17 out of 70 (24%) pairwise tests Fst = 0.

Table 6. Pairwise Fst values among the five Siberian populations tested in six neutral STR polymorphisms (D4S397, D5S393, D8S514, D9S161, D11S1358 and D13S173).

Population	Buryats	Altays	Yakuts	Russians
Tuvinians	0.0189	0.0548	0.0625	0.0292
Buryats	–	0.0517	0.0810	0.0462
Altays	–	–	0.0583	0.0378
Yakuts	–	–	–	0.0731

Bibliography

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Wilkinson GR: Drug metabolism and variability among patients in drug response. *N. Engl. J. Med.* 352, 2211–2221 (2005).
2. Meyer UA, Zanger UM: Molecular mechanisms of genetic polymorphisms of drug metabolism. *Annu. Rev. Pharmacol. Toxicol.* 37, 269–296 (1997).
3. Gaedigk A, Gotschall RR, Forbes NS, Simon SD, Kearns GL, Leeder JS: Optimization of cytochrome P4502D6 (CYP2D6) phenotype assignment using a genotyping algorithm based on allele frequency data. *Pharmacogenetics* 9, 669–682 (1999).
4. Kirchheiner J, Nickchen K, Bauer M *et al.*: Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response. *Mol. Psychiatry* 9, 442–473 (2004).
5. de Leon J, Armstrong SC, Cozza KL: The dosing of atypical antipsychotics. *Psychosomatics* 46, 262–273 (2005).
6. Aquilante CL, Langaee TY, Lopez LM *et al.*: Influence of coagulation factor, vitamin K epoxide reductase complex subunit 1, and cytochrome P450 2C9 gene polymorphisms on warfarin dose requirements. *Clin. Pharmacol. Ther.* 79, 291–302 (2006).
7. Marsh S, van Booven DJ, McLeod HL: Global pharmacogenetics: giving the genome to the masses. *Pharmacogenomics* 7(4), 625–631 (2006).

•• Important paper describing the PharmacoGenetics for Every Nation Initiative as a first step to making pharmacogenetics globaly applicable.

8. Garte S, Gaspari L, Alexandrie AK *et al.*: Metabolic gene polymorphism frequencies in control populations. *Cancer Epidemiol. Biomarkers Prev.* 10, 1239–1248 (2001).
9. Johns MB, Paulus-Thomas JE: Purification of human genomic DNA from whole blood using sodium perchlorate in place of phenol: *Anal. Biochem.* 180, 276–278 (1989).
10. Livak KJ: Allelic discrimination using fluorogenic probes and the 5´ nuclease assay. *Genet. Anal.* 14, 143–149 (1999).
11. Stepanov VA, Spiridonova MG, Tadinova VN, Puzyrev VP: Analysis of genetic diversity of North Eurasian populations using autosomal microsatellite loci. *Russ. J. Genet.* 39(10), 1168–1174 (2003).
12. Khitrinskaia IIu, Stepanov VA, Puzyrev VP *et al.*: Genetic peculiarity of the Yakut population as infered from the autosomal loci data. *Mol. Biol. (Mosk).* 37(2), 205–209 (2003).
13. Khitrinskaia IIu, Stepanov VA, Pyzyrev VP: Analysis of *Alu* polymorphism in Buryat populations. *Russ. J. Genet.* 37(11), 1306–1311 (2001).
14. Khitrinskaia IIu, Stepanov VA, Puzyrev VP, Spiridonova MG, Voevoda MI: Genetic differentiation of the population of central Asia inferred from autosomal markers. *Russ. J. Genet.* 39(10), 1175–1184 (2003).
15. Stepanov VA, Spiridonova MG, Puzyrev VP: Comparative phylogenetic study of native north Eurasian populations using a panel of autosomal microsatellite loci. *Russ. J. Genet.* 11(39), 1326–1333 (2003).
16. Guo S, Thompson E: Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics* 48, 361–372 (1992).
17. Excoffier L, Laval G, Schneider S: Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1, 47–50 (2005).
18. Weir BS, Cockerham CC: Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370 (1984).
19. Excoffier L, Smouse P, Quattro J: Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491 (1992).
20. Weir BS: *Genetic Data Analysis II: Methods for discrete population genetic data*. Sinauer Assoc, Inc.: MA, USA, 1–445 (1996).
21. Smouse P, Long J, Sokal RR: Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst. Zool.* 35, 627–632 (1986).
22. Gaikovitch EA, Cascorbi I, Mrozikiewicz PM *et al.*: Polymorphisms of drug-metabolizing enzymes CYP2C9, CYP2C19, CYP2D6, CYP1A1, NAT2 and of P-glycoprotein in a Russian population. *Eur. J. Clin. Pharmacol.* 59, 303–312 (2003).
23. Pchelina SN, Sirotkina OV, Taraskina AE, Vavilova TV, Shwarzman AL, Schwartz EI: The frequency of cytochrome P450 2C9 genetic variants in the Russian population and their associations with individual sensitivity to warfarin therapy. *Thromb. Res.* 115(3), 199–203 (2005).
24. Allabi AC, Gala JL, Desager JP, Heusterspreute M, Horsmans Y: Genetic polymorphisms of *CYP2C9* and *CYP2C19* in the Beninese and Belgian populations. *J. Clin. Pharmacol.* 56, 653–657 (2003).
25. Arvanitidis K, Ragia G, Iordanidou M *et al.*: Genetic polymorphisms of drug-metabolising enzymes *CYP2D6*, *CYP2C9*, *CYP2C19* and *CYP3A5* in the Greek population. *Fundam. Clin. Pharmacol.* 21(4), 419–426 (2007).
26. Bozina N, Granic P, Lalic Z, Tramišak I, Lovric M, Stavljenic-Rukavina A: Genetic polymorphisms of cytochromes P450: *CYP2C9*, *CYP2C19*, and *CYP2D6* in Croatian population. *Croat. Med. J.* 44(4), 425–428 (2003).
27. Scordo MG, Caputi AP, D'Arrigo C, Fava G, Spina E: Allele and genotype frequencies of *CYP2C9*, *CYP2C19* and *CYP2D6* in an Italian population. *Pharmacol. Res.* 50, 195–200 (2004).
28. Chang M, Dahl ML, Tybring G, Gotharson E, Bertilsson L: Use of omeprazole as a probe drug for CYP2C19 phenotype in Swedish Caucasians: comparison with S-mephenytoin hydroxylation phenotype and *CYP2C19* genotype. *Pharmacogenetics* 5(6), 358–363 (1995).
29. Nowak MP, Sellers EM, Tyndale RF: Canadian Native Indians exhibit unique *CYP2A6* and *CYP2C19* mutant allele frequencies. *Clin. Pharmacol. Ther.* 64, 378–383 (1998).
30. Roh HK, Dahl ML, Tybring G, Yamada H, Cha YN, Bertilsson L: *CYP2C19* genotype and phenotype determined by omeprazole in a Korean population. *Pharmacogenetics* 6, 547–551 (1996).
31. Shimizu T, Ochiai H, Asell F *et al.*: Bioinformatics research on inter-racial difference in drug metabolism I. Analysis of frequencies of mutant alleles and poor metabolizers on CYP2D6 and CYP2C19. *Drug Metab. Pharmacokinet.* 18(1), 48–70 (2003).
32. Takakubo F, Kuwano A, Kondo I: Evidence that poor metabolizers of (S)-mephenytoin could be identified by haplotypes of *CYP2C19* in Japanese. *Pharmacogenetics* 6, 265–267 (1996).
33. Yamada S, Onda M, Kato S *et al.*: Genetic differences in *CYP2C19* single nucleotide polymorphisms among four Asian populations. *J. Gastroenterol.* 36, 669–672 (2001).

34. Andersson T, Flockhart DA, Goldstein DB *et al.*: Drug-metabolizing enzymes: evidence for clinical utility of pharmacogenomic tests. *Clin. Pharmacol. Ther.* 78(6), 559–581 (2005).
• Useful paper (commentary) describing the most medically significant drug-metabolizing enzyme (DME) genes, polymorphic sites, their clinical relevance and recommendations as to which polymorphic alleles to evaluate in various ethnic/racial groups.
35. International HapMap Consortium: A haplotype map of the human genome. *Nature* 437, 1299–1320 (2005).
36. Gu S, Pakstis AJ, Li H, Speed WC, Kidd JR, Kidd KK: Significant variation in haplotype block structure but conservation in tagSNP patterns among global populations. *Eur. J. Hum. Genet.* 15(3), 302–312 (2007).
37. González-Neira A, Ke X, Lao O *et al.*: The portability of tagSNPs across populations: a worldwide survey. *Genome Res.* 16(3), 323–330 (2006).
38. Mitchell-Olds T, Willis JH, Goldstein DB: Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nat. Rev.* 8, 845–856 (2007).
39. Kuehl P, Zhang J, Lin Y *et al.*: Sequence diversity in *CYP3A* promoters and characterization of the genetic basis of polymorphic *CYP3A5* expression. *Nat. Genet.* 27, 383–391 (2001).
40. Weir BS, Cockerham CC: Estimating F-statistics for the analysis of population structure. *Evolution* 38(6), 1358–1370 (1984).
41. Weir BS, Cardon LR, Anderson AD, Nielsen DM, Hill WG: Measures of human population structure show heterogeneity among genomic regions. *Genome Res.* 15, 1468–1476 (2005).
42. Akey JM, Zhang G, Zhang K, Jin L, Shriver MD: Interrogating a high-density SNP map for signatures of natural selection. *Genome Res.* 12, 1805–1814 (2002).
•• Research paper identifying a number of genomic regions that are the target of natural selection. The data supports the hypothesis that selection has influenced extensive patterns of human genetic variation.
43. Schwarz UI: Clinical relevance of genetic polymorphisms in the human *CYP2C9* gene. *Eur. J. Clin. Invest.* 33(Suppl. 2), 23–30 (2003).
44. Zhang X, Chen BC, Xiang FL *et al.*: Influence of *CYP3A5* and *MDR1* polymorphisms on tacrolimus concentration in the early stage after renal transplantation. *Clin. Transplant.* 19(5), 638–643 (2005).
45. de Leon J: AmpliChip CYP450 test: personalized medicine has arrived in psychiatry. *Expert Rev. Mol. Diagn.* 6, 277–286 (2006).
46. Lehmann DF, Medicis JJ, Franklin PD: Polymorphisms and the pocketbook: the cost–effectiveness of cytochrome P450 2C19 genotyping in the eradication of *Helicobacter pylori* infection associated with duodenal ulcer. *J. Clin. Pharmacol.* 43(12), 1316–1323 (2003).
47. Matchar DB, Thakur ME, Grossman I *et al.*: Testing for cytochrome P450 polymorphisms in adults with non-psychotic depression treated with selective serotonin reuptake inhibitors (SSRIs). *Evid. Rep. Technol. Assess (Full Rep.)* 146, 1–77 (2007).
48. Grossman I, Liu Y, Walley N *et al.*: Pharmacogenetic analysis of antipsychotics: comprehensive analysis of pharmacokinetic variants. Presented at: *56th Annual Meeting of the American Society of Human Genetics (ASHG)*. New Orleans, LA, USA, 9–13 October 2006.
49. Bulletin Board: FDA releases final guidance for pharmacogenomic data. *Pharmacogenomics* 6(3), 209 (2005).
50. Grossman I: Routine pharmacogenetic testing in clinical practice – dream or reality? *Pharmacogenomics* 8(10), 1449–1459 (2007).
51. Cascorbi I, Ackermann E, Sachse C, Brockmoller J, Roots I: A novel *CYP2C9* intron 2T/C transition and linkage to mutations Leu359 and Cys144. *Clin. Pharmacol. Ther.* 63, 198 (1998).
52. Taube J, Halsall D, Baglin T: Influence of cytochrome P-450 *CYP2C9* polymorphisms on warfarin sensitivity and risk over-anticoagulation in patients on long-term treatment. *Blood* 96, 1816–1819 (2000).
53. Garcia-Martin E, Martinez C, Ladero JM, Gamito FJ, Agundez JA: High frequency of mutations related to impaired CYP2C9 metabolism in a Caucasian population. *Eur. J. Clin. Pharmacol.* 57, 47–49 (2001).
54. Dorado P, Berecz R, Norberto MJ, Yasar U, Dahl ML, LLerena A: *CYP2C9* genotypes and diclofenac metabolism in Spanish healthy volunteers. *Eur. J. Clin. Pharmacol.* 59, 221–225 (2003).
55. Yasar U, Eliasson E, Dahl ML, Johansson I, Ingelman-Sundberg M, Sjöqvist F: Validation of methods for *CYP2C9* genotyping: frequencies of mutant alleles in a Swedish. *Biochem. Biophys. Res. Commun.* 254(3), 628–631 (1999).
56. Gaedigk A, Casley WL, Tyndale RF, Sellers EM, Jurima-Romet M, Leeder JS: Cytochrome P4502C9 (*CYP2C9*) allele frequencies in Canadian Native Indian and Inuit populations. *Can. J. Physiol. Pharmacol.* 79, 841–847 (2001).
57. Wang SL, Huang J, Lai MD, Tsai JJ: Detection of *CYP2C9* polymorphism based on the polymerase chain reaction in Chinese. *Pharmacogenetics* 5, 37–42 (1995).
58. Nakai K, Habano W, Nakai K *et al.*: Ethnic differences in *CYP2C9*2* (Arg144Cys) and *CYP2C9*3* (Ile359Leu) genotypes in Japanese and Israeli populations. *Life Sci.* 78, 107–111 (2005).
59. Nasu K, Kubota T, Ishizaki T: Genetic analysis of *CYP2C9* polymorphism in a Japanese population. *Pharmacogenetics* 7, 405–409 (1997).
60. Yoon YR, Shon JH, Kim MK *et al.*: Frequency of cytochrome P450 2C9 mutant alleles in a Korean population. *Br. J. Clin. Pharmacol.* 51, 277–280 (2001).
61. Aynacioglu AS, Sachse C, Bozkurt A *et al.*: Low frequency of defective alleles of cytochrome P450 enzymes 2C19 and 2D6 in the Turkish population. *Clin. Pharmacol. Ther.* 66, 185–192 (1999).
62. van Schaik RHN, van der Heiden IP, van den Anker JN, Lindemans J: *CYP3A5* variant allele frequencies in Dutch Caucasians. *Clin. Chem.* 48, 1668–1671 (2002).
63. Quaranta S, Chevalier D, Bourgarel-Rey V *et al.*: Identification by single-strand conformational polymorphism analysis of known and new mutations of the *CYP3A5* gene in a French population. *Toxicol. Lett.* 164(2), 177–184 (2006).
64. King BP, Leathart JB, Mutch E, Williams FM, Daly AK: CYP3A5 phenotype–genotype correlations in a British population. *Br. J. Clin. Pharmacol.* 55(6), 625–629 (2003).
65. Gervasini G, Vizcaino S, Gasiba C, Carrillo JA, Benitez J: Differences in *CYP3A5*3* genotype distribution and combinations with other polymorphisms between Spaniards and other Caucasian populations. *Ther. Drug Monit.* 27(6), 819–821 (2005).
66. Mirghani RA, Sayi J, Aklillu E *et al.*: *CYP3A5* genotype has significant effect on quinine 3-hydroxylation in Tanzanians, who have lower total CYP3A activity than a Swedish population. *Pharmacogenet. Genomics* 16(9), 637–645 (2006).

67. Thompson EE, Kuttab-Boulos H, Witonsky D, Yang L, Roe BA, Di Rienzo A: *CYP3A5* variation and the evolution of salt-sensitivity variants. *Am. J. Hum. Genet.* 75(6), 1059–1069 (2004).
• Interesting research devoted to investigation of the *CYP3A5* gene and the remarkable interpopulation allele frequency differences it displays; the study aims to show that genetic variation in DMEs can be under the influence of selective pressure.
68. Hustert E, Haberl M, Burk O *et al.*: The genetic determinants of the *CYP3A5* polymorphism. *Pharmacogenetics* 11(9), 773–779 (2001).
69. Balram C, Zhou Q, Cheung YB, Lee EJ: *CYP3A5*3* and **6* single nucleotide polymorphisms in three distinct Asian populations. *Eur. J. Clin. Pharmacol.* 59(2), 123–126 (2003).
70. Wang HP, Xie JJ, Zhang ZY *et al.*: Study on polymorphisms of *CYP3A5* gene and their clinical role. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 22(4), 423–426 (2005).
71. Saeki M, Saito Y, Nakamura T *et al.*: Single nucleotide polymorphisms and haplotype frequencies of *CYP3A5* in a Japanese population. *Hum. Mutat.* 21(6), 653 (2003).
72. Hiratsuka M, Takekuma Y, Endo N *et al.*: Allele and genotype frequencies of *CYP2B6* and *CYP3A5* in the Japanese population. *Eur. J. Clin. Pharmacol.* 58(6), 417–421 (2002).
73. Fukuen S, Fukuda T, Maune H *et al.*: Novel detection assay by PCR-RFLP and frequency of the *CYP3A5* SNPs, *CYP3A5*3* and **6*, in a Japanese population. *Pharmacogenetics* 12(4), 331–334 (2002).

Website

101. Home page of the human cytochrome P450 allele nomenclature committee http://www.cypalleles.ki.se/

Supplementary Material

The supplementary material includes:
- Figure S1
- Tables S1–10

Figure S1. Map of Russia denoting the location of residence for the studied populations.

1: Tuvinians; 2: Buryats; 3: Southern Altaians; 4: Yakuts; 5: Russians.

Table S1. The CYP450 gene polymorphisms studied, their identifiers and the surrounding sequence.

Gene	Variant	SNP rs#	ABI Assay-by-Demand #	Sequence
CYP2C9	**2*	rs1799853	C_25625805_10	GATGGGGAAGAGGAGCATTGAGGAC[C/T]GTGT TCAAGAGGAAGCCCGCTGCCT
CYP2C9	**3*	rs1057910	C_27104892_10	TGTGGTGCACGAGGTCCAGAGATAC[C/A]TTGAC CTTCTCCCCACCAGCCTGCC
CYP2C19	**2*	rs4244285	C_25986767_70	TTCCCACTATCATTGATTATTTCCC[A/G]GGAACCC ATAACAAATTACTTAAAA
CYP2C19	**3*	rs4986893	C_27861809_10	ACATCAGGATTGTAAGCACCCCCTG[A/G]ATCCA GGTAAGGCCAAGTTTTTTGC
CYP3A5	**3*	rs776746	C_26201809_30	ATGTGGTCCAAACAGGGAAGAGATA[T/C]TGAAA GACAAAAGAGCTCTTTAAAG
CYP3A5	**6*	rs10264272	C_30203950_10	CTAAGAAACCAAATTTTAGGAACTT[C/T]TTAGTG CTCTCCACAAAGGGGTCTT

ABI: Applied biosystems.

Table S2. The genetic models and their functional annotations.

Gene	Variant	Marker	Allele function
CYP2C9	**2* Arg144Cys	rs1799853	T decreased activity
	**3* Ile359Leu	rs1057910	C severely decreased activity
CYP2C19	**2* Pro227Pro	rs4244285	A inactive
	**3* Trp212Ter	rs4986893	A inactive
CYP3A5	**3*	rs776746	G severely decreased activity
	**6* Lys208Lys	rs10264272	A decreased activity

Table S3. Characteristics of the STR polymorphisms tested.

N	STR ID	Chr	Location start	Location end	Closest gene	Gene start	Gene end	Distance from nearest gene	Primer F	Primer R
1	D4S379	4	141,087,371	141,087,580	*MAML3*	140,859,927	141,031,580	55,791	CACATAACTTCCCTTGCTGG	ACTGTTGTCAAATCAGGCTC
2	D5S393	5	135,729,237	135,729,402	*TRPC7*	135,377,022	135,720,974	8,263	TTCTACCTGACCTTTCCTCT	CATTCCTCATTCCTCATTCC
3	D8S514	8	123,811,415	123,811,633	*ZHX2*	123,863,082	124,055,936	51,449	CCAGTTGGCAAGCATTGT	CTGAACCCAGTAGAGTTAGGAGA
4	D9S161	9	27,622,327	27,622,447	*C9orf72*	27,536,544	27,563,864	58,463	TGCTGCATAACAAATTACCAC	CATGCCTAGACTCCTGATCC
5	D11S1358	11	90,005,698	90,005,841	*CHORDC1*	89,573,832	89,595,827	409,871	ACAACCTGGATGAACCC	ACTTTCTGTCTTTATGATTTTGATT
6	D13S173	13	106,604,948	106,605,119	*Q5W0J2_HUMAN*	106,620,319	107 317 084	15,200	CCCTGTTCCAGTAATGATGACC	GTCTCTGGCTGCTCTCAAGACTAT

Chromosomal locations are based on Ensembl Release 41 October 2006.
Chr: Chromosome; F: Forward; R: Reverse.

Table S4. Characteristics of the Alu insertion polymorphisms tested.

N	Alu ID	Chr	Location start	Location end	Closest gene	Gene start	Gene end	Distance from nearest gene	Primer F	Primer R
1	A25	8	88,534,626	88,534,917	*CNBD1*	87,947,840	88,435,135	99,491	CCACAAATAGGCTCATGTAGAAC	TATAATATGGCCTGGATTATACC
2	ACE (rs4646994)	17	58,919,548	58,919,833	*ACE*	58,908,166	58,938,721	Inside	CTGGAGACCACTCCCATCCTTTCT	GATGTGGCCATCACATTCGCTAGAT
3	APOA1	11	116,217,638	116,218,049	*Q5M9N1*	116,219,328	116,474,347	1,279	AAGTGCTGTAGGCCATTTAGATTAG	AGTCTTCGATGACAGCGTATACAGA
4	CD4 (rs4646985)	12	6,791,268	6,791,524	*CD4*	6,769,005	6,800,233	Inside	AGGCCTTGTAGGGTTGGTCTGATA	TGCAGCTGCTGAGTGAAAGAACTG
5	F13	1	195,278,252	195,278,689	*F13B*	195,274,944	195,303,020	Inside	TCAACTCCATGAGATTTTCAGAAGT	CTGGAAAAAATGTATTCAGGTGAGT
6	PLAT (rs4646972)	8	42,154,434	42,154,745	*PLAT*	42,151,912	42,184,351	Inside	GTAAGAGTTCCGTAACAGGACAGCT	CCCCACCCTAGGAGAACTTCTCTTT
7	PV92 (rs3138523)	16	81,451,610	81,451,869	*CDH13*	81,449,460	82,387,705	Inside	AACTGGGAAAATTTGAAGAGAAAGT	TGAGTTCTCAACTCCTGTGTGTTAG

Chromosomal locations are based on Ensembl Release 41 October 2006.
F: Forward; R: Reverse.

Table S5. Source information for Figure 1 listing the displayed populations, the CYP allele frequencies, the number of subjects studied and references: *CYP2C9*2*.

Population	Allele frequency (%)	N	Ref.
Germans	0.107	367	[51]
Belgian	0.1	121	[24]
Croatians	0.165	200	[26]
British	0.106	561	[52]
Italian	0.125	360	[27]
Spanish	0.143	157	[53]
Spanish	0.156	102	[54]
Spanish pooled	0.149	259	[53,54]
Swedish	0.107	430	[55]
Greek	0.129	283	[25]
White US	0.15	325	[56]
Chinese	0	102	[56]
Chinese	0	115	[57]
Chinese pooled	0	217	[56,57]
Japanese	0	147	[58]
Japanese	0	218	[59]
Japanese pooled	0	365	[58,59]
Koreans	0	574	[60]
Canadian Native Indian	0.03	114	[56]

When several independent papers reported on the same population, the data were pooled together and calculated based on the total counts.

Table S6. Source information for Figure 1 listing the displayed populations, the CYP allele frequencies, the number of subjects studied and references: *CYP2C9*3*.

Population	Allele frequency (%)	N	Ref.
Germans	0.078	367	[51]
Belgian	0.074	121	[24]
Croatians	0.095	200	[26]
British	0.053	561	[52]
Italian	0.097	360	[27]
Spanish	0.162	157	[53]
Spanish	0.098	102	[54]
Spanish pooled	0.137	259	[53,54]
Swedish	0.074	430	[55]
Greek	0.081	283	[26]
White US	0.07	325	[56]
Chinese	0.05	102	[56]
Chinese	0.017	115	[57]
Chinese pooled	0.03	217	[56,57]
Japanese	0.014	147	[59]
Japanese	0.021	218	[59]
Japanese pooled	0.0175	365	[58,59]
Koreans	0.011	574	[60]
Canadian Native Indian	0.06	114	[56]

When several independent papers reported on the same population, the data were pooled together and calculated based on the total counts.

Table S7. Source information for Figure 1 listing the displayed populations, the CYP allele frequencies, the number of subjects studied and references: *CYP2C19*2*.

Population	Allele frequency (%)	N	Ref.
Europeans	0.145	1659	[31]
Germans	0.159	237	[61]
Belgian	0.091	121	[24]
Croatians	0.15	200	[26]
Italian	0.111	360	[27]
Swedish	0.166	160	[28]
Greek	0.131	283	[25]
White US	0.141	556	[31]
Chinese	0.31	516	[31]
Chinese	0.455	121	[33]
Chinese pooled	0.38	637	[31,33]
Japanese	0.294	843	[31]
Japanese	0.274	217	[32]
Japanese pooled	0.28	1060	[32,33]
Koreans	0.209	103	[30]
Canadian Native Indian	0.191	115	[29]

When several independent papers reported on the same population, the data were pooled together and calculated based on the total counts.

Table S8. Source information for Figure 1 listing the displayed populations, the CYP allele frequencies, the number of subjects studied and references: *CYP2C19*3*.

Population	Allele frequency (%)	N	Ref.
Europeans	0.001	707	[31]
Germans	0.015	237	[61]
Belgian	0	121	[24]
Croatians	0	200	[26]
Italian	0	360	[27]
Swedish	0	160	[28]
Greek	0	283	[25]
White US	0	556	[31]
Chinese	0.061	516	[31]
Chinese	0.045	121	[33]
Chinese pooled	0.053	637	[33,31]
Japanese	0.123	843	[31]
Japanese	0.108	217	[32]
Japanese pooled	0.116	1060	[31,32]
Koreans	0.117	103	[30]
Canadian Native Indian	0	115	[29]

When several independent papers reported on the same population, the data were pooled together and calculated based on the total counts.

Table S9. Source information for Figure 1 listing the displayed populations, the CYP allele frequencies, the number of subjects studied and references: *CYP3A5*3*.

Population	Allele Frequency (%)	N	Ref.
Dutch	0.91	500	[62]
French	0.87	114	[63]
British	0.935	100	[64]
Spanish	0.91	177	[65]
Greek	0.94	283	[25]
Swedish	0.93	136	[66]
Italian	0.95	36	[67]
German	0.94	432	[68]
Chinese	0.76	108	[69]
Chinese	0.761	180	[70]
Chinese pooled	0.760	288	[69,70]
Japanese	0.759	187	[71]
Japanese	0.74	265	[72]
Japanese	0.767	200	[73]
Japanese pooled	0.755	652	[71–73]

When several independent papers reported on the same population, the data were pooled together and calculated based on the total counts.

Table S10. Frequency data used in the principal component analysis.

	Germans	Belgian	Italian	Swedish	Greek	White US	Tuvinians	Buryats	Altaians	Yakuts	Russians	Japanese	Chinese	Koreans
*CYP2C9*2*	0.107	0.1	0.125	0.107	0.129	0.15	0.011	0.023	0.057	0.011	0.121	0	0	0.01
*CYP2C9*3*	0.078	0.074	0.097	0.074	0.081	0.07	0.051	0.017	0.092	0.006	0.068	0.0335	0.018	0.011
*CYP2C19*2*	0.159	0.091	0.111	0.166	0.131	0.141	0.148	0.21	0.149	0.233	0.122	0.383	0.284	0.209
*CYP2C19*3*	0.015	0	0	0	0	0	0.023	0.068	0.04	0.046	0	0.053	0.116	0.117
*CYP3A5*3*	0.94	0.91	0.95	0.95	0.944	0.91	0.856	0.822	0.894	0.926	0.916	0.7605	0.755	0.78