**PROCEEDINGS OF
THE ROYAL SOCIETY B**

BIOLOGICAL SCIENCES

Beyond drugs: the evolution of genes involved in human response to medications

Journal:	*Proceedings B*
Manuscript ID	Draft
Article Type:	Review
Date Submitted by the Author:	n/a
Complete List of Authors:	Fuselli, Silvia; University of Ferrara, Life Sciences and Biotechnologies
Subject:	Evolution < BIOLOGY, Genetics < BIOLOGY, Health and Disease and Epidemiology < BIOLOGY
Keywords:	Pharmacogenomics, drug metabolizing enzymes, natural selection, evolution, rare genetic variants
Proceedings B category:	Genetics & Genomics

**SCHOLARONE™
Manuscripts**

Author-supplied statements

Relevant information will appear here if provided.

Ethics

Does your article include research that required ethical approval or permits?:
This article does not present research with ethical considerations

Statement (if applicable):
CUST_IF_YES_ETHICS :No data available.

Data

It is a condition of publication that data, code and materials supporting your paper are made publicly available. Does your paper present new data?:
My paper has no data

Statement (if applicable):
CUST_IF_YES_DATA :No data available.

Conflict of interest

I/We declare we have no competing interests

Statement (if applicable):
CUST_STATE_CONFLICT :No data available.

Authors' contributions

I am the only author on this paper

Statement (if applicable):
CUST_AUTHOR_CONTRIBUTIONS_TEXT :No data available.

REVIEW

Beyond drugs: the evolution of genes involved in human response to medications

Silvia Fuselli*

Department of Life Sciences and Biotechnology
University of Ferrara, Ferrara, Italy

*corresponding author: Silvia Fuselli, fss@unife.it

1

Abstract

The genetic variation of our species reflects human demographic history and adaptation to diverse local environments. Part of this genetic variation affects individual responses to exogenous substances, such as food, pollutants and drugs, and plays an important role in drug efficacy and safety. This review provides a synthesis of the evolution of loci implicated in human pharmacological response and metabolism, interpreted within the theoretical framework of population genetics and molecular evolution. In particular, I review and discuss key evolutionary aspects of different pharmacogenes in humans and other species, such as the relationship between type of substrates and rate of evolution; the selective pressure exerted by landscape variables or dietary habits; expected and observed patterns of rare genetic variation. Finally, I discuss how this knowledge can be translated, directly or after the implementation of specific studies, into practical guidelines.

1. Introduction

Evolutionary history and adaptation of humans to diverse local environments left a strong signature on the genetic variation of our species. A subset of this genetic variation carried by pharmacogenes implicated in response to medications (hereafter PGx genes), affects drug efficacy and safety, and varies extensively between individuals and across populations [1–3]. PGx genes encode for molecules such as drug metabolizing enzymes (DMEs) responsible for drug degradation or activation; drug transporters, regulating the movement of drug molecules across the cell membrane; and drug targets, to which the drug directly binds to produce a therapeutic effect [1,4,5]. Established PGx genotype–phenotype relationships have been translated into successful clinical implementation in cases such as clopidogrel (gene: *CYP2C219*), simvastatin (*SLCO1B1*), warfarin (*CYP2C9* and *VKORC1*), thiopurines (*TPMT*) and tacrolimus (*CYP3A5*)[2,6]. In recent years, developments in DNA sequencing technologies enabled the transition from the candidate-gene approach typical of pharmacogenetics to a genome-wide pharmacogenomic approach [2,7]. Yet, the translation of extensive pharmacogenetic and genomic knowledge into clinical practice has been slower than expected. This has prompted the establishment of scientific networks and programs aimed at overcoming the barriers that have thus far hindered clinical implementations [8,9]. A better understanding of specific evolutionary aspects concerning variation in PGx genes could improve the accuracy in predicting drug response phenotypes. For example, the extreme level of genetic polymorphism shown by DMEs [3,10] is related to their ability to detoxify the body from molecules of environmental origin that vary significantly with diet, climate, and lifestyle. Thus, the wide

3

range of variation in DMEs genes, which are mediators between the organism and the environment, makes evolutionary sense: more alleles allow a wider range of responses, both individually and as a species. Regarding the spatial structure of this genetic variation, populations living in similar environmental conditions should share adaptive genetic variants or allele combinations resulting in similar phenotypes. Conversely, marked differences are expected between populations for locally adaptive alleles [11,12]. These genetic differences between populations are particularly relevant in the age of personalized pharmacogenomics.

The value of evolutionary principles in medicine is increasingly being recognized in public health [13]. Here, I mention three examples with a clear applicative relevance: a) the knowledge on HIV evolution is essential for the design of preventive and therapeutic measures [14]; b) understanding the link between the altered symbiotic community of microorganisms living in our body and the onset of specific diseases [15] can guide the development of disease prevention protocols; c) evolutionary-based models of human disease architecture are crucial in designing and interpreting genome-wide association studies [16]. Similarly, improving our understanding of the evolution of PGx loci can enhance the potential of personalized therapy for patients.

This review provides a synthesis of the current knowledge on PGx loci evolution in humans and in other species, and indicates how this knowledge can be translated, directly or after the implementation of specific studies, into practical guidelines. Particular emphasis will be given to the evolutionary history of rare variants in PGx genes that may have important implications for future pharmacogenetics study design and clinical implementation.

2. Insights into the molecular evolution of PGx genes

Comparisons between different species can assist in identifying specific genes or gene families characterized by frequent or rare changes along their evolutionary history.

Genes encoding for DMEs -such as cytochrome P450 (CYP450), UDP-glucuronosyltransferase (UGT) or N-acetyltransferase (NAT)- but also for transporters such as ATP-binding cassette (ABC) show extreme levels of single nucleotide variation (SNP) and copy number variation (CNV) in animal genomes [17–19]. Most DME and transporter genes are part of gene families subject to a birth-and-death model of evolution, whereby new genes are created by repeated duplications and remain active in the genome, whereas others are inactivated or lost [20]. A central question is whether the extent of diversification of DME and transporter genes within and between species is stochastic (i.e. selectively neutral) or driven by adaptation. In this second case, the rate of gene evolution should be related to the function of the encoded protein.

The analysis of 50 to 80 CYP450 genes in ten nearly complete vertebrate genomes[21] showed that most of the genes characterized by frequent gene duplications and losses encode enzymes that function primarily as xenobiotic detoxifiers (henceforth *exogenous* genes). One possible explanation is that these genes have evolved under positive selection for copy number changes, which may have increased the ability of the species to coevolve with the environment. Signatures of positive selection for amino acid changes observed in this group of *exogenous* genes further support this interpretation. On the other hand, genes with few changes in copy number over evolutionary time were found to encode enzymes

with known functions in development and physiology (henceforth *endogenous* genes) [21].

In this case, changes in gene number, especially gene losses, may be relatively more

deleterious and, therefore, kept at low frequency by purifying selection. Evidence obtained

by comparing CYP450 genes in different *Drosophila* species, as well as glutathione S

transferase (GSTs) genes in mammals further supports the *exogenous-endogenous* model of

evolution [22,23]. Within humans, a recent analysis showed that CYP genes metabolizing

exogenous molecules have a much greater proportion of SNPs highly differentiated

between populations compared to any other category of PGx genes [3]. Similarly, the

different selective regime inferred for the two human NAT genes, namely purifying

selection on *NAT1* and positive selection for amino acid diversification on *NAT2,* could be

explained in terms of substrate's nature, since only *NAT1* seems to metabolize endogenous

molecules[24,25]. In conclusion, even if exceptions have been identified [17], the

exogenous-endogenous model of evolution seems to hold for different species and DME

genes. Prediction based on this model may be useful for pharmacogenetic applications

(Figure 1A). For example, patterns of genetic variation of gluthatione S-transferases

(GSTs) superfamily enzymes with a prevalent exogenous metabolism that includes

carcinogens, may be important for cancer detection and progression, and to predict the

outcome of cancer treatment [26]. In terms of evolution and practical applications it would

be interesting to verify if the *exogenous-endogenous* model can be extended to transporter

genes, many of which are part of actively evolving gene families, such as ABCs or Solute

Carriers (SLCs) genes[18,27,28].

3. Environmental factors influencing the evolution of PGx genes

Loci coding for proteins that interact with molecules of environmental origin, including PGx genes, often experience some sort of natural selection. Classical examples are the loci involved in pathogen defense, such as those coding for the major histocompatibility complex (MHC) of vertebrates, which typically evolve under long-term balancing selection, or the lysozyme of primates, evolving under positive selection in specific phylogenetic lineages[29,30]. Interestingly, signatures of natural selection have been detected at several PGx genes, mostly DMEs, with specific genetic variants or CNV pattern conferring a selective advantage in a specific environment.

(a) PGx loci and examples of adaptation in non-human species

Instructive examples of adaptation by natural selection involving PGx genes have been identified in non-human species. Recently, a comparative study of carnivore, omnivore, and herbivore mammalian genomes has revealed a general relationship between diet and rate of evolution of *UGT 1* and *2* gene families[31]. While the two families appear markedly contracted (i.e. few gene copies per genome) in carnivores, their copy number is expanded in herbivores and omnivores in which the dietary content of plant-derived toxicants is higher. Among carnivores, loss of function mutations inactivating the gene *UGT1A6* are more common in groups consuming a hypercarnivore diet (i.e. a diet comprising of more than 70% meat)[32]. Since the rate of nucleotide substitutions in genes is inversely related to the stringency of the functional constraints acting upon them[33], genes that code for enzymes that no longer have a metabolic role are expected to accumulate mutations and

7

become pseudogenes[34]. Given the low dietary content of plant-derived phenolic

intoxicants in carnivores, the relaxation of natural selection is probably responsible for the

pseudogenization of the otherwise broadly conserved *UGT1A6* gene[32].

In birds, whole-genome comparisons of 48 species with different feeding habits,

geographic distribution, and migratory behavior revealed the influence of these variables on

the CYP2 enzyme family evolution[35]. Specific CYP2 enzymes in different groups of

birds showed striking signatures of diversifying selection at substrate recognition sites, as

well as difference in copy number. Different CYP2 molecules are probably involved in the

adaptation of a specific avian group to a distinct ecological niche. In most of the reported

examples adaptation seems to be achieved by variation in the number of coding genes or

gene copies, likely leading to an increased or a decreased metabolic effect[36]. The

CYP450 enzymes in the Bactrian camel adapted to the arid conditions of the desert support

this conclusion. Compared to the genome of most vertebrates, the *CYP4A* and *CYP4F*

families in this species are contracted, whereas the *CYP2E* and *CYP2J* families are

expanded. Since all these CYP450s are involved in the arachidonic acid pathway, this

specific copy number pattern may contribute to the camels' ability to take in a large amount

of salt, apparently without developing hypertension[37].

(b) Coevolution between PGx loci and different lifestyles in humans

Human populations subsisted by hunting wild animals and gathering wild plants for about

90 percent of their history. Between 10,000 and 4,000 years ago, domestic plants and

animals appeared independently in several regions of the world. This so called Neolithic

transition represents a turning point in the evolution of our species in terms of subsistence

mode, with consequent dramatic changes in human lifestyle, population size, and diet [38].

All these changes left a footprint in our genome [39–41]. Several studies have investigated

the effect of the Neolithic transition on the evolution of PGx genes by comparing present-

day hunter-gatherers with different types of food-producing populations (Supplementary

Table 1, Figure 1B). Most of these studies focused on the gene *NAT2* coding for the Phase

II DME N-acetyltransferase 2. Detailed information on the *NAT2* genetic variation is

available for all geographic regions across the world [42]. Interestingly, when *NAT2* alleles

are translated into predicted phenotype activity, gene variants coding for slow-function

NAT2 isoforms represent more than 50% of the total variation worldwide, with the

exception of East Asia [43]. Furthermore, *NAT2* genetic diversity shows an association

with distinct subsistence mode in all continents except America, whereby food-producing

populations tend to be slower metabolizers compared to hunter-gatherers [43]. From a

neutral evolutionary perspective, dietary modification brought about by the emergence of

farming could have relaxed the selective constraint on the detoxification enzyme NAT2.

However, signature of a selective sweep, such as that exhibited by the slow-function

*NAT2*5B* in western and central Eurasian populations [25], or evidence of genetic variation

actively maintained by balancing selection [44,45] invoke a non-neutral explanation for a

reduction in the NAT2 enzymatic activity associated to the Neolithic transition. A low

metabolic activity may have been selected to avoid the bioactivation of toxic substances

(e.g. NAT2 activates heterocyclic carcinogens found in well-cooked meat)[24].

A similar trend was observed in one of the most important DME in humans, the

Cytochrome P450 2D6. By controlling for environment and demography, a comparison of

9

the entire *CYP2D6* genetic variation between present-day hunter-gatherers and food-producers revealed in the latter a significant excess of both slower metabolizers and intermediate-frequency non-synonymous polymorphisms [11] (Figure 1B). However, akin to *NAT2,* the identification of the selective agent responsible for diet-related patterns of evolution in *CYP2D6* (e.g. the presence/absence of a specific substrate or a different concentration of substrates) has yet to be identified. Further insights into genetic changes related to the transition from hunter-gatherer to farming lifestyles may be obtained from ancient genomes across different time points and environments [41]. Ancient DNA allows measuring the action of natural selection directly by detecting allele frequency changes through time, in particular before, during, and after a specific event requiring adaptation. Such direct measurements improve estimates of the strength and timing of selection, as in the case of lipid metabolism and lactose tolerance in humans[46].

(c) PGx loci and examples of recent human adaptation

Understanding the role played by positive selection in the evolution of human populations has been a central issue in evolutionary genetics[47]. Genome scans in humans have occasionally identified PGx loci as target of recent positive selection, including signals at the *SLC24A5* locus in non-African populations, possibly due to its association with skin pigmentation[48,49], or at the *CYP3A4* and *CYP3A5* loci[50]. The example of the *CYP3A* cluster shows how specific and identifiable landscape variables can shape patterns of genetic variation with important phenotypic outcomes. The cluster includes four genes - *CYP3A4*, *CYP3A5*, *CYP3A7*, and *CYP3A43*- located on chromosome 7. *CYP3A4* and *CYP3A5* are important Phase I DMEs and are responsible for the metabolism of many

10

commonly used drugs (pharmgkb.org, cpicpgx.org). A whole-gene resequencing study

conducted by Thompson et al. [51] in three populations from different continents,

combined with the screening of the null-function allele in more than 1,000 individuals,

revealed an unusual geographical pattern of extreme variation across human populations

and significant correlation with distance from the equator. In particular, the signature of

recent selective sweep associated with the null *CYP3A5*3* allele (rs776746 T>C) was

interpreted as an adaptation underlying salt regulation. A strong sweep would explain why

the derived (i.e. arose by mutation) *CYP3A5*3* is the major allele in most non-African

populations while being nearly fixed in Europe (Figure 1C). Few other human genes show

the same pattern that was previously defined as "west Eurasia sweep" [49]. Interestingly,

the proposed selective mechanism involves the endogenous role of CYP3A5 enzyme, rather

than its detoxification activity. Specifically, in the kidney, CYP3A5 converts cortisol to 6b-

hydroxycortisol that leads to a higher reabsorption of sodium and water retention.

According to the salt-retention hypothesis [52], the allele coding for a full-function protein

confers a selective advantage in equatorial populations where water shortages are common

[53]. Heat dissipation lost its adaptive value in human populations that colonized colder

regions during the out-of-Africa expansion. Moreover, at lower temperatures, salt-retention

alleles may cause salt-sensitive hypertension and, possibly, pre-eclampsia thus having

important fitness reduction effects [54]. This can partially account for the selective sweep

that has increased the frequency of the impaired-function and derived allele outside Africa.

The explanation based on the salt-retention hypothesis is also supported by the correlation

between latitude and specific alleles of genes involved in blood pressure regulation [51,55].

The case of *CYP3A* cluster illustrates how positive selection can give rise to evident global

11

geographic structure of genetic variation in response to environmental variables (Figure 1C). At the other extreme, selection may lead to detectable allele frequency change in a restricted geographic region, as in the case of the Arsenite Methyltransferase *AS3MT* in the populations residing near San Antonio de los Cobres, Argentina. Here, individuals have likely been exposed to poisonous levels of the acutely toxic arsenic present in the water for thousands of years. In a recent association study, Schlebusch et al. [56] identified a potential protective regulatory variant upstream of *AS3MT*, a gene involved in the arsenic methylation pathway. The putatively protective variants are highly frequent in these populations (68.7%), and are embedded in a long haplotype, both of which are signatures of recent positive selection.

(d) Rapid evolution of PGx genes by artificial selection

Examples of evolution by artificial selection reveal how PGx genes evolve under extreme selective pressure. Poisonous substances, such as insecticides and pesticides, increase the fitness of individuals carrying a resistant genotype. The difference in fitness between resistant and sensitive individuals depends on the magnitude of the selection coefficient imposed by the poisonous substance. Larger selection coefficients determine a faster increase in the frequency of the advantageous mutation and, therefore, a more rapid spread of the resistant phenotype in the population [57]. Fungi and insects adapt to pesticides in a matter of years, offering very good examples of acquired resistance via drug metabolizing enzymes' evolution [58,59]. Resistance can result from the rapid increase in frequency of an advantageous mutation conferring less sensitivity to a specific molecule [60], or from an increased production of detoxifying enzyme due to overtranscription or copy number

12

variation [61]. An example of resistance acquired through artificial selection is that of the vitamin K 2,3-epoxide reductase subcomponent 1 (vkorc1) and anticoagulants in rodents. Polymorphism in vkorc1 determines the physiological response of humans and rodents to warfarin [62,63]. In humans, *VKORC1* and *CYP2C9* polymorphisms are included in genotype-guided dosage of this widely prescribed anticoagulant [6], which was also used as rodenticide in Europe since the 1950s. Genetic analyses have shown how vkorc1-mediated resistance convergently evolved in different rodent pest species in few years through selection on adaptive genetic variants [62]. Interestingly, the house mice *Mus musculus domesticus* have acquired adaptive vkorc1 polymorphisms from the Algerian mouse (*M. spretus*) through introgressive hybridization [64]. This example stresses the fact that resistance phenomena can spread rapidly through the gene flow between populations and species.

The fast evolution at the somatic level due to the strong selection coefficient involved in cancer progression and resistance to chemotherapy is an important example of artificial selection in our species [65]. In normal human tissues, the constituent cells are not the unit of selection and their fitness is the same as the fitness of the entire organism. Tissue cells do not compete among each other, and their proliferation and survival is governed entirely by tissue control mechanisms acting though intracellular pathways. Conversely, somatic cells isolated from normal tissue controls can evolve toward cancer and their survival and proliferation is governed by interaction of their own properties with the environment [66]. Therapeutic resistance can then arise from one or several subclones harboring resistance-conferring mutations, which can be either pre-existing (i.e. present at diagnosis) or

13

acquired. Both mechanisms have been observed, for example, in acute myeloid leukemia and urothelial carcinoma [67,68]. Relapse after remission generated from minor genetic subclones already present at diagnosis is particularly interesting from an evolutionary perspective. These subclones, which were carrying resistance-conferring mutation but were outcompeted in the absence of therapy, will eventually be positively selected by treatment and expand [69]. Understanding tumor evolution during therapy is at the same time a unique opportunity to study evolution in action at the cell level, and promising for drug development strategies [70].

4. Rare genetic variants in human PGx genes

The importance of rare genetic variants, whose minor allele has a frequency below 5% in the global human population, has been overlooked for a long time in human genetics. Due to the low frequency, rare variants tend to be geographically restricted, or even private to a population[71], so their predictive role in drug response, for example, is not equally important in different human populations[72]. Additionally, before the advent of Next-Generation Sequencing (NGS) technologies, the lack of cost-effective genotyping approach led to the exclusion of rare genetic variants from the worldwide used SNP reference panels, such as those generated by the HapMap projects[73]. The development of NGS has enabled researchers to obtain information on both common and rare genetic variants, overcoming many intrinsic biases of a SNP typing approach. With a slight delay, this change has also involved pharmacogenetics and pharmacogenomics[10,74] (Supplementary Table 2). In the

14

following, I first discuss important aspects of rare genetic variants that can assist in

interpreting their patterns in human PGx loci; then I provide examples of studies focusing

on rare variants in genes for drug targets, DMEs and transporters.

(a) Important aspects of rare genetic variation in humans

The allele frequency spectrum is the distribution of allele frequencies at a large number of

equivalent loci, and its shape is affected by population size and selection intensity [75]

(Figure 1A). During the last 10,000 years our species has experienced an explosive

population growth, from few million people to roughly 7 billion today [76,77]. Rapid

population growth are characterized by a very large number of recent mutations that skew

patterns of genetic variation increasing the number of rare genetic variants. Most of these

new variants will be either deleterious (or even lethal), slightly deleterious, or neutral (i.e.

that do not affect the fitness), while only very few will be advantageous [78]. The fate of

new mutations is then determined by the interplay of genetic drift and selection, depending

on the population size and on the mutation's effect on fitness (i.e. the selection coefficient,

s) [79]. In sufficiently large populations purifying selection acts against deleterious

variants, which then have lower population frequency and are generally younger (i.e. not

yet purged by selection) than the neutral ones [80] (Figure 1A). Consistent with these

expectations, a study on 6,515 human exomes in populations of European and African

ancestry estimated that most of the rare variants, especially those with deleterious effect,

arose in the past 5,000 years [81]. Conversely, in small populations the strength of selection

is reduced and the fate of each mutations, even if deleterious, and especially if slightly

deleterious, is mainly governed by random genetic drift. It is therefore expected that small

15

population size, bottleneck or founder effect will be characterized by an excess of deleterious variants that can drift to high frequency or even to fixation [78]. In line with this prediction, people of European descent that have experienced the bottleneck associated with the Out-of-Africa dispersal, seem to carry an excess of deleterious variants compared to those of African descent [81,82]. Altogether, in terms of site frequency spectrum, given the equal input of mutations, deleterious variants will be numerous among rare alleles in small as well as in large populations. Conversely, their proportion at intermediate and high frequency depends on the equilibrium between drift and selection [80,82].

(b) Examples of studies on rare genetic PGx variants

In the last few years, several studies focused on the importance of low-frequency genetic variants in PGx genes (Supplementary Table 2). Their results consistently indicate that rare variants in PGx genes are (i) enriched for predicted (in few cases functionally tested) protein-damaging mutations, (ii) abundant and population-specific, and (iii) usually neglected in common pharmacogenetics genotyping panels (and, when included, affected by ascertainment bias[83–85]). Most of these studies analyzed PGx or ADME (absorption, distribution, metabolism, and excretion) loci as a whole (Supplementary Table 2). However, as described in the previous paragraphs, PGx or ADME loci code for proteins with very heterogeneous functions, which means different evolutionary histories and, thus, heterogeneous patterns of rare genetic variation. To the aim of determining the abundance, distribution, and phenotypic effects of rare variants in drug-related phenotypes, studies focused on more homogeneous group of genes may be more informative. For example, in the study by Nelson et al.[86] the observed patterns of genetic variation of 202 drug target

genes in 14,002 ethnically diverse individuals match predictions based on evolutionary

principles. Specifically, the genes in this study show signature of strong purifying selection

(i.e. they are extremely conserved), which is consistent with their role as drug targets and

importance to human health (Figure 1A). The excess of rare genetic variants that

characterizes the large European subset analyzed in the study (12,514 individuals) is

concentrated in gene regions with a more likely functional effect (i.e. untranslated region

(UTR) and non-synonymous variants), which are underrepresented among common-

frequency variants. A significantly different pattern can be observed at putatively neutral

sites (i.e. synonymous and intronic regions). Thus, target genes as a group are under

considerable selective constraint in all populations, even in those with smaller population

size. Purifying selection at these loci effectively prevents new damaging mutations from

becoming common and shared among populations. Patterns of common and rare genetic

variation have also been analyzed in drug metabolizing enzymes. For example, Gordon et

al. [87] sequenced the entire coding region of 12 *CYP* genes in two large population

samples of African American and European origin. The 12 *CYP*-gene panel includes most

of the DMEs for which therapeutic recommendation and guidelines exist (pharmgkb.org,

cpicpgx.org), as these are collectively responsible for nearly 75% of all known drug

oxidizing reactions [88,89]. Non-synonymous variation in the 12-CYP genes is similar to

the rest of the exome, with the interesting exception of the extremely variable *CYP2A6*,

CYP2B6 and *CYP2D6* (see also Zhou et al.[90]). In terms of proportion of variants with

functional effect, the 12 *CYP* genes appear less constrained than drug target genes [86]

(Figure 1A). However, for rare variants, target genes and the 12 CYP genes show similar

17

patterns, and newly discovered variants are exceedingly rare and enriched for predicted damaging mutations in all population samples.

To date, patterns and effects of rare variants in transporter genes have mostly been analyzed in the general context of PGx genes [83–85,90] (Supplementary Table 2). Two notable exceptions concern the *SLCO* gene family. The first is a detailed analysis of rare genetic variation in the gene *SLCO1B1*[91]*,* coding for a hepatic anion transporter of extreme pharmacogenetic interest [4]. By sequencing the gene in a very large group of well-phenotyped individuals, the authors concluded that rare variants significantly contribute to the inter-individual variation in clearance of methotrexate - an antifolate agent used in autoimmune diseases and malignancies. The functional effect of some of these detrimental mutations has been confirmed via *in vitro* analyses [91]. Recently, a comprehensive analysis of *SLCO* variability by Zhang and Lauschke [92] highlighted the considerable impact (although highly gene-dependent) of rare functional variants, which represented the sole source of variation for 7 out of the 11 human *SLCO* genes.

5. Conclusions and practical implications

Evolutionary processes such as mutation, genetic drift, natural selection, migration and non-random mating, are important for understanding the role of genetics in response to drug. The following key points summarize the important aspects discussed in this review and delineate their implications for pharmacogenetics and pharmacogenomics approaches.

1. Comparative studies between different species suggest a direct relationship between the extent of the endogenous role of a PGx protein and its conservation. Proteins mainly interacting with substrates of environmental origin will thus be more polymorphic and more prone to CNV. In the case of unstable genes, genotyping approaches based on few known polymorphic positions may greatly underestimate genetic variation compared to a whole-gene resequencing[28].

2. Factors that influence the evolution of PGx genes in human and non-human species are of demographic or environmental origin. Modern drugs can have an evolutionary effect only if they actually influence the individual's fitness, such as in the case of warfarin for mice and some chemotherapies in human cancer treatment. Given the fast replication of tumor cells and the strong selective coefficient involved in chemotherapy, the application of evolutionary principles in cancer treatment is particularly promising [93]. For example, the application of adaptive evolutionary strategies allowed to control the proliferation of resistant cells in breast cancer by stabilizing the population of therapy-sensitive cells through intratumoral competition [94].

19

3. Local adaptation can generate gene-specific spatial distribution of detrimental alleles [39]. In humans, loci that have undergone directional selective pressure -such as *CYP3A5, VKORC1 or AS3MT*- typically exhibit marked (and locus-specific) differences between populations or geographic regions (Figure 1C). On the other hand, convergent adaptation, or relaxation of natural selection, may be responsible for the presence of slow-function alleles in different geographic regions for genes such as *CYP2D6* and *NAT2* (Figure 1B). From a practical perspective, this means that genetic variation at PGx loci as a whole cannot be partitioned in population groups based, for example, on general continental or ethnic criteria [3].

4. As theoretically postulated, rare damaging PGx genetic variants are abundant, and therefore important, but are also population-specific, and thus difficult to include into standard genetic tests. Given the current human population size and considering a genome-wide average mutation rate of $1.5 *10^{-8}$ per site per generation [95], the number of mutations arising each generation is substantial, at least 100 billion mutations. As sequencing costs continue to decline, analyzing rare and private variants on an individual's genome will become feasible [96]. Until then, a reasonable compromise could be to focus on variants segregating at sufficiently common frequency to be detected without sequencing of the whole genome, and rare enough to be enriched for deleterious variants, as suggested by association studies in disease genetics [97,98].

Finally, it is important and timely to mention the additional level of complexity due to population admixture. Historically admixed populations such as African-Americans,

20

Brazilians and Caribbeans, but also new mixing due to recent and very recent migrations, for example from Asia to Australia, or from Africa and Middle East to Europe, can show particular combinations of genetic variants which are not captured by standard and population-based pharmacogenetics tests [99]. Studies in Brazilians, a classical admixed population, show that for genes highly differentiated between Africans and Europeans such as *ABCB1*, *CYP3A5*, *CYP2C9*, *VKORC1*, the probability of observing a specific genotype may be modelled as a function of individual's continental admixture [100]. In clinical pharmacogenetics, it has been suggested that the inclusion of admixture inferences improve dosage algorithm for warfarin in Puerto Rico [101]. A more in-depth genetic characterization of admixed populations is required to capture the set of genetic variation that is unique to admixed individuals. This approach would provide valuable insights to further develop pharmacogenetic tests specific for individuals with likely multiple ancestry, or to include key SNPs and genes in multi-populations tests.

References

[1]   Weinshilboum R. Inheritance and Drug Response. N Engl J Med 2003;348:529–37.
        doi:10.1056/NEJMra020021.

[2]   Meyer UA, Zanger UM, Schwab M. Omics and Drug Response. Annu Rev
        Pharmacol Toxicol 2013;53:475–502. doi:10.1146/annurev-pharmtox-010510-
        100502.

[3]   Maisano Delser P, Fuselli S. Human loci involved in drug biotransformation:
        Worldwide genetic variation, population structure, and pharmacogenetic
        implications. Hum Genet 2013;132. doi:10.1007/s00439-013-1268-5.

[4]   Giacomini KM, Huang S-M, Tweedie DJ, Benet LZ, Brouwer KLR, Chu X, et al.
        Membrane transporters in drug development. Nat Rev Drug Discov 2010;9:215–36.
        doi:10.1038/nrd3028.

[5]   Santos R, Ursu O, Gaulton A, Bento AP, Donadi RS, Bologa CG, et al. A
        comprehensive map of molecular drug targets. Nat Rev Drug Discov 2017;16:19–34.
        doi:10.1038/nrd.2016.230.

[6]   Van Driest SL, Shi Y, Bowton EA, Schildcrout JS, Peterson JF, Pulley J, et al.
        Clinically actionable genotypes among 10,000 patients with preemptive
        pharmacogenomic testing. Clin Pharmacol Ther 2014;95:423–31.
        doi:10.1038/clpt.2013.229.

[7]   Giacomini KM, Yee SW, Mushiroda T, Weinshilboum RM, Ratain MJ, Kubo M.

23

Genome-wide association studies of drug response and toxicity: An opportunity for genome medicine. Nat Rev Drug Discov 2016;16:1. doi:10.1038/nrd.2016.234.

[8]   Luzum J, Pakyz R, Elsey A, Haidar C, Peterson J, Whirl-Carrillo M, et al. The Pharmacogenomics Research Network Translational Pharmacogenetics Program: Outcomes and Metrics of Pharmacogenetic Implementations Across Diverse Healthcare Systems. Clin Pharmacol Ther 2017;102:502–10. doi:10.1002/cpt.630.

[9]   van der Wouden C, Cambon-Thomsen A, Cecchin E, Cheung K, Dávila-Fajardo C, Deneer V, et al. Implementing Pharmacogenomics in Europe: Design and Implementation Strategy of the Ubiquitous Pharmacogenomics Consortium. Clin Pharmacol Ther 2017;101:341–58. doi:10.1002/cpt.602.

[10]   Drogemoller BI, Wright GEB, Warnich L. Considerations for Rare Variants in Drug Metabolism Genes and the Clinical Implications. Expert Opin Drug Metab Toxicol 2014;10:873–84. doi:10.1517/17425255.2014.903239.

[11]   Fuselli S, de Filippo C, Mona S, Sistonen J, Fariselli P, Destro-Bisol G, et al. Evolution of detoxifying systems: the role of environment and population history in shaping genetic diversity at human CYP2D6 locus. Pharmacogenet Genomics 2010;20:485–99. doi:10.1097/FPC.0b013e32833bba25.

[12]   Podgorná E, Diallo I, Vangenot C, Sanchez-Mazas A, Sabbagh A, Černý V, et al. Variation in NAT2 acetylation phenotypes is associated with differences in food-producing subsistence modes and ecoregions in Africa. BMC Evol Biol 2015;15:1–20. doi:10.1186/s12862-015-0543-6.

[13]   Wells JCK, Nesse RM, Sear R, Johnstone RA, Stearns SC. Evolutionary public

health: introducing the concept. Lancet 2017;390:500–9. doi:10.1016/S0140-

6736(17)30572-X.

[14]   Theys K, Libin P, Pineda-Peña A-C, Nowé A, Vandamme A-M, Abecasis AB. The

impact of HIV-1 within-host evolution on transmission dynamics. Curr Opin Virol

2018;28:92–101. doi:10.1016/j.coviro.2017.12.001.

[15]   Francino MP. Antibiotics and the Human Gut Microbiome: Dysbioses and

Accumulation of Resistances. Front Microbiol 2015;6:1543.

doi:10.3389/fmicb.2015.01543.

[16]   Simons YB, Bullaughey K, Hudson RR, Sella G. A population genetic interpretation

of GWAS findings for human quantitative traits. PLoS Biol 2018;16:e2002985.

doi:10.1371/journal.pbio.2002985.

[17]   Sezutsu H, le Goff G, Feyereisen R. Origins of P450 diversity. Philos Trans R Soc B

Biol Sci 2013;368. doi:10.1098/rstb.2012.0428.

[18]   Moitra K, Dean M. Evolution of ABC transporters by gene duplication and their role

in human disease. Biol Chem 2011;392:29–37. doi:10.1515/BC.2011.006.

[19]   Sabbagh A, Marin J, Veyssière C, Lecompte E, Boukouvala S, Poloni ES, et al.

Rapid birth-and-death evolution of the xenobiotic metabolizing NAT gene family in

vertebrates with evidence of adaptive selection. BMC Evol Biol 2013;13.

doi:10.1186/1471-2148-13-62.

[20]   Nei M, Rooney AP. Concerted and Birth-and-Death Evolution of Multigene

25

Families. Annu Rev Genet 2005;39:121–52.

doi:10.1146/annurev.genet.39.073003.112240.

[21]   Thomas JH. Rapid birth-death evolution specific to xenobiotic cytochrome P450

genes in vertebrates. PLoS Genet 2007;3:e67. doi:10.1371/journal.pgen.0030067.

[22]   Drosophila 12 Genomes Consortium, Clark AG, Eisen MB, Smith DR, Bergman

CM, Oliver B, et al. Evolution of genes and genomes on the Drosophila phylogeny.

Nature 2007;450:203–18. doi:10.1038/nature06341.

[23]   Good RT, Gramzow L, Battlay P, Sztal T, Batterham P, Robin C. The Molecular

Evolution of Cytochrome P450 Genes within and between Drosophila Species.

Genome Biol Evol 2014;6:1118–34. doi:10.1093/gbe/evu083.

[24]   Hein DW. Molecular genetics and function of NAT1 and NAT2: role in aromatic

amine metabolism and carcinogenesis. Mutat Res 2002;506–507:65–77.

[25]   Patin E, Barreiro LB, Sabeti PC, Austerlitz F, Luca F, Sajantila A, et al. Deciphering

the ancient and complex evolutionary history of human arylamine N-

acetyltransferase genes. Am J Hum Genet 2006;78:423–36. doi:10.1086/500614.

[26]   Chatterjee A, Gupta S. The multifaceted role of glutathione S-transferases in cancer.

Cancer Lett 2018;433:33–42. doi:10.1016/j.canlet.2018.06.028.

[27]   Zhu C, Nigam KB, Date RC, Bush KT, Springer SA, Saier MH, et al. Evolutionary

Analysis and Classification of OATs, OCTs, OCTNs, and Other SLC22

Transporters: Structure-Function Implications and Analysis of Sequence Motifs.

PLoS One 2015;10:e0140569. doi:10.1371/journal.pone.0140569.

26

[28] Lauschke VM, Ingelman-Sundberg M. Prediction of drug response and adverse drug reactions: From twin studies to Next Generation Sequencing. Eur J Pharm Sci 2019;130:65–77. doi:10.1016/j.ejps.2019.01.024.

[29] Garrigan D, Hedrick PW. Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. Evolution 2003;57:1707–22.

[30] Yang Z. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol Biol Evol 1998;15:568–73. doi:10.1093/oxfordjournals.molbev.a025957.

[31] Kim S, Cho YS, Kim HM, Chung O, Kim H, Jho S, et al. Comparison of carnivore, omnivore, and herbivore mammalian genomes with a new leopard assembly. Genome Biol 2016;17:1–12. doi:10.1186/s13059-016-1071-4.

[32] Shrestha B, Reed JM, Starks PT, Kaufman GE, Goldstone J V., Roelke ME, et al. Evolution of a major drug metabolizing enzyme defect in the domestic cat and other Felidae: Phylogenetic timing and the role of hypercarnivory. PLoS One 2011;6. doi:10.1371/journal.pone.0018046.

[33] Jukes TH, Kimura M. Evolutionary constraints and the neutral theory. J Mol Evol 1984;21:90–2.

[34] Calderoni L, Rota-Stabelli O, Frigato E, Panziera A, Kirchner S, Foulkes NS, et al. Relaxed selective constraints drove functional modifications in peripheral photoreception of the cavefish P. andruzzii and provide insight into the time of cave colonization. Heredity (Edinb) 2016;117:383–92. doi:10.1038/hdy.2016.59.

27

[35]   Almeida D, Maldonado E, Khan I, Silva L, Gilbert MTP, Zhang G, et al. Whole-genome identification, phylogeny, and evolution of the cytochrome P450 family 2 (CYP2) subfamilies in birds. Genome Biol Evol 2016;8:1115–31. doi:10.1093/gbe/evw041.

[36]   He Y, Hoskins JM, Mcleod HL. Copy Number Variants in pharmacogenetic genes 2011. doi:10.1016/j.molmed.2011.01.007.

[37]   Jirimutu, Wang Z, Ding G, Chen G, Sun Y, Sun Z, et al. Genome sequences of wild and domestic bactrian camels. Nat Commun 2012;3:1202. doi:10.1038/ncomms2192.

[38]   Bellwood PS. First farmers : the origins of agricultural societies. Blackwell Pub; 2005.

[39]   Fan S, Hansen MEB, Lo Y, Tishkoff SA. Going global by adapting local: A review of recent human adaptation. Science 2016;354:54–9. doi:10.1126/science.aaf5098.

[40]   Lopez M, Kousathanas A, Quach H, Harmant C, Mouguiama-Daouda P, Hombert JM, et al. The demographic history and mutational load of African hunter-gatherers and farmers. Nat Ecol Evol 2018;2:721–30. doi:10.1038/s41559-018-0496-4.

[41]   Marciniak S, Perry GH. Harnessing ancient genomes to study the history of human adaptation. Nat Rev Genet 2017;18:659–74. doi:10.1038/nrg.2017.65.

[42]   Sabbagh A, Langaney A, Darlu P, Gérard N, Krishnamoorthy R, Poloni ES. Worldwide distribution of NAT2 diversity: Implications for NAT2 evolutionary history. BMC Genet 2008;9:1–14. doi:10.1186/1471-2156-9-21.

28

[43] Sabbagh A, Darlu P, Vangenot C, Poloni ES. Arylamine *N* -Acetyltransferases in Anthropology. Arylamine N -acetyltransferases Heal. Dis., WORLD SCIENTIFIC; 2018, p. 165–93. doi:10.1142/9789813232013_0007.

[44] Luca F, Bubba G, Basile M, Brdicka R, Michalodimitrakis E, Rickards O, et al. Multiple advantageous amino acid variants in the NAT2 gene in human populations. PLoS One 2008;3:e3136. doi:10.1371/journal.pone.0003136.

[45] Magalon H, Patin E, Austerlitz F, Hegay T, Aldashev A, Quintana-Murci L, et al. Population genetic diversity of the NAT2 gene supports a role of acetylation in human adaptation to farming in Central Asia. Eur J Hum Genet 2008;16:243–51. doi:10.1038/sj.ejhg.5201963.

[46] Mathieson S, Mathieson I. *FADS1* and the Timing of Human Adaptation to Agriculture. Mol Biol Evol 2018;35:2957–70. doi:10.1093/molbev/msy180.

[47] Nielsen R, Hellmann I, Hubisz M, Bustamante C, Clark AG. Recent and ongoing selection in the human genome. Nat Rev Genet 2007;8:857–68. doi:10.1038/nrg2187.

[48] Voight BF, Kudaravalli S, Wen X, Pritchard JK. A Map of Recent Positive Selection in the Human Genome. PLoS Biol 2006;4:e72. doi:10.1371/journal.pbio.0040072.

[49] Coop G, Pickrell JK, Novembre J, Kudaravalli S, Li J, Absher D, et al. The role of geography in human adaptation. PLoS Genet 2009;5. doi:10.1371/journal.pgen.1000500.

[50] Carlson CS, Thomas DJ, Eberle MA, Swanson JE, Livingston RJ, Rieder MJ, et al.

Genomic regions exhibiting positive selection identified from dense genotype data. Genome Res 2005:1553–65. doi:10.1101/gr.4326505.

[51] Thompson EE, Kuttab-Boulos H, Witonsky D, Yang L, Roe B a, Di Rienzo a. CYP3A variation and the evolution of salt-sensitivity variants. Am J Hum Genet 2004;75:1059–69. doi:10.1086/426406.

[52] Gleibermann L. Blood Pressure and Dietary Salt in Human Populations. Ecol Food Nutr 1973;2:143–56. doi:10.1080/03670244.1973.9990329.

[53] Kuehl P, Zhang J, Lin Y, Lamba J, Assem M, Schuetz J, et al. Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. Nat Genet 2001;27:383–91. doi:10.1038/86882.

[54] Sibai B, Dekker G, Kupferminc M. Pre-eclampsia. Lancet (London, England) 2005;365:785–99. doi:10.1016/S0140-6736(05)17987-2.

[55] Young JH, Chang Y-PC, Kim JD-O, Chretien J-P, Klag MJ, Levine M a, et al. Differential susceptibility to hypertension is due to selection during the out-of-Africa expansion. PLoS Genet 2005;1:e82. doi:10.1371/journal.pgen.0010082.

[56] Schlebusch CM, Gattepaille LM, Engström K, Vahter M, Jakobsson M, Broberg K. Human adaptation to arsenic-rich environments. Mol Biol Evol 2015;32:1544–55. doi:10.1093/molbev/msv046.

[57] Nielsen R, Slatkin M. Chapter 7 Selecion I. Sinauer Associates; 2013.

[58] Kelly SL, Kelly DE. Microbial cytochromes P450: biodiversity and biotechnology.

Where do cytochromes P450 come from, what do they do and what can they do for us? Philos Trans R Soc Lond B Biol Sci 2013;368:20120476. doi:10.1098/rstb.2012.0476.

[59] Schmidt JM, Battlay P, Gledhill-Smith RS, Good RT, Lumb C, Fournier-Level A, et al. Insights into DDT Resistance from the Drosophila melanogaster Genetic Reference Panel. Genetics 2017;207:1181–93. doi:10.1534/genetics.117.300310.

[60] Nash D, Nair S, Mayxay M, Newton PN, Guthmann J-P, Nosten F, et al. Selection strength and hitchhiking around two anti-malarial resistance genes. Proceedings Biol Sci 2005;272:1153–61. doi:10.1098/rspb.2004.3026.

[61] Bass C, Field LM. Gene amplification and insecticide resistance. Pest Manag Sci 2011;67:886–90. doi:10.1002/ps.2189.

[62] Pelz H-J, Rost S, Hünerberg M, Fregin A, Heiberg A-C, Baert K, et al. The genetic basis of resistance to anticoagulants in rodents. Genetics 2005;170:1839–47. doi:10.1534/genetics.104.040360.

[63] Pirmohamed M, Kamali F, Daly AK, Wadelius M. Oral anticoagulation: a critique of recent advances and controversies. Trends Pharmacol Sci 2015;36:153–63. doi:10.1016/j.tips.2015.01.003.

[64] Song Y, Endepols S, Klemann N, Richter D, Matuschka FR, Shih CH, et al. Adaptive introgression of anticoagulant rodent poison resistance by hybridization between old world mice. Curr Biol 2011;21:1296–301. doi:10.1016/j.cub.2011.06.043.

31

[65]   Burrell RA, Swanton C. Tumour heterogeneity and the evolution of polyclonal drug resistance. Mol Oncol 2014;8:1095–111. doi:10.1016/j.molonc.2014.06.005.

[66]   Shpak M, Lu J. An Evolutionary Genetic Perspective on Cancer Biology. Annu Rev Ecol Evol Syst 2016;47:25–49. doi:10.1146/annurev-ecolsys-121415-032109.

[67]   Shlush LI, Mitchell A, Heisler L, Abelson S, Ng SWK, Trotman-Grant A, et al. Tracing the origins of relapse in acute myeloid leukaemia to stem cells. Nature 2017;547:104–8. doi:10.1038/nature22993.

[68]   Vlachostergios PJ, Faltas BM. Treatment resistance in urothelial carcinoma: an evolutionary perspective. Nat Rev Clin Oncol 2018;15:495–509. doi:10.1038/s41571-018-0026-y.

[69]   Greaves M, Maley CC. Clonal evolution in cancer. Nature 2012;481:306–13. doi:10.1038/nature10762.

[70]   Sottoriva A, Spiteri I, Piccirillo SGM, Touloumis A, Collins VP, Marioni JC, et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proc Natl Acad Sci 2013;110:4009–14. doi:10.1073/pnas.1219747110.

[71]   Gravel S, Henn BM, Gutenkunst RN, Indap AR, Marth GT, Clark AG, et al. Demographic history and rare allele sharing among human populations. Proc Natl Acad Sci 2011;108:11983–8. doi:10.1073/pnas.1019276108.

[72]   Apellániz-Ruiz M, Inglada-Pérez L, Naranjo MEG, Sánchez L, Mancikova V, Currás-Freixes M, et al. High frequency and founder effect of the CYP3A4*20 loss-of-function allele in the Spanish population classifies CYP3A4 as a polymorphic

enzyme. Pharmacogenomics J 2015;15:288–92. doi:10.1038/tpj.2014.67.

[73]   International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851–61. doi:10.1038/nature06258.

[74]   Ingelman-Sundberg M, Mkrtchian S, Zhou Y, Lauschke VM. Integrating rare genetic variants into pharmacogenetic drug response predictions. Hum Genomics 2018;12:26. doi:10.1186/s40246-018-0157-3.

[75]   Evans SN, Shvets Y, Slatkin M. Non-equilibrium theory of the allele frequency spectrum. Theor Popul Biol 2007;71:109–19. doi:10.1016/J.TPB.2006.06.005.

[76]   Keinan A, Clark AG. Recent explosive human population growth has resulted in an excess of rare genetic variants. Science (80-) 2012;336:740–3. doi:10.1126/science.1217283.

[77]   Roberts L. 9 billion? Science 2011;333:540–3. doi:10.1126/science.333.6042.540.

[78]   Ohta T. The Nearly Neutral Theory of Molecular Evolution. Annu Rev Ecol Syst 1992;23:263–86. doi:10.1146/annurev.es.23.110192.001403.

[79]   Gravel S. When Is Selection Effective? Genetics 2016;203:451–62. doi:10.1534/genetics.115.184630.

[80]   Henn BM, Botigué LR, Bustamante CD, Clark AG, Gravel S. Estimating the mutation load in human genomes. Nat Rev Genet 2015;16:333–43. doi:10.1038/nrg3931.

33

[81]   Fu W, O'Connor TD, Jun G, Kang HM, Abecasis G, Leal SM, et al. Analysis of

       6,515 exomes reveals the recent origin of most human protein-coding variants.

       Nature 2013;493:216–20. doi:10.1038/nature11690.

[82]   Lohmueller KE, Indap AR, Schmidt S, Boyko AR, Hernandez RD, Hubisz MJ, et al.

       Proportionally more deleterious genetic variation in European than in African

       populations. Nature 2008;451:994–7. doi:10.1038/nature06611.

[83]   Hovelson DH, Xue Z, Zawistowski M, Ehm MG, Harris EC, Stocker SL, et al.

       Characterization of ADME gene variation in 21 populations by exome sequencing.

       Pharmacogenet Genomics 2017;27:89–100. doi:10.1097/FPC.0000000000000260.

[84]   Han SM, Park J, Lee JH, Lee SS, Kim H, Han H, et al. Targeted Next-Generation

       Sequencing for Comprehensive Genetic Profiling of Pharmacogenes. Clin Pharmacol

       Ther 2017;101:396–405. doi:10.1002/cpt.532.

[85]   Bush WS, Crosslin DR, Owusu-Obeng A, Wallace J, Almoguera B, Basford MA, et

       al. Genetic variation among 82 pharmacogenes: The PGRNseq data from the

       eMERGE network. Clin Pharmacol Ther 2016;100:160–9. doi:10.1002/cpt.350.

[86]   Nelson MR, Wegmann D, Ehm MG, Kessner D, St. Jean P, Verzilli C, et al. An

       abundance of rare functional variants in 202 drug target genes sequenced in 14,002

       people. Science (80-) 2012;336:100–4. doi:10.1126/science.1217876.

[87]   Gordon AS, Tabor HK, Johnson AD, Snively BM, Assimes TL, Auer PL, et al.

       Quantifying rare, deleterious variation in 12 human cytochrome P450 drug-

       metabolism genes in a large-scale exome dataset. Hum Mol Genet 2014;23:1957–63.

doi:10.1093/hmg/ddt588.

[88] Evans WE, Relling M V. Pharmacogenomics: translating functional genomics into rational therapeutics. Science 1999;286:487–91.

[89] Sim SC, Kacevska M, Ingelman-Sundberg M. Pharmacogenomics of drug-metabolizing enzymes: a recent update on clinical implications and endogenous effects. Pharmacogenomics J 2013;13:1–11. doi:10.1038/tpj.2012.45.

[90] Zhou Y, Ingelman-Sundberg M, Lauschke VM. Worldwide Distribution of Cytochrome P450 Alleles: A Meta-analysis of Population-scale Sequencing Projects. Clin Pharmacol Ther 2017;102:688–700. doi:10.1002/cpt.690.

[91] Ramsey LB, Bruun GH, Yang W, Treviño LR, Vattathil S, Scheet P, et al. Rare versus common variants in pharmacogenetics: SLCO1B1 variation and methotrexate disposition. Genome Res 2012;22:1–8. doi:10.1101/gr.129668.111.

[92] Zhang B, Lauschke VM. Genetic variability and population diversity of the human SLCO (OATP) transporter family. Pharmacol Res 2019;139:550–9. doi:10.1016/J.PHRS.2018.10.017.

[93] Gatenby R, Brown J. The Evolution and Ecology of Resistance in Cancer Therapy. Cold Spring Harb Perspect Med 2018;8:a033415. doi:10.1101/cshperspect.a033415.

[94] Silva AS, Kam Y, Khin ZP, Minton SE, Gillies RJ, Gatenby RA. Evolutionary approaches to prolong progression-free survival in breast cancer. Cancer Res 2012;72:6362–70. doi:10.1158/0008-5472.CAN-12-2235.

35

[95]　Narasimhan VM, Rahbari R, Scally A, Wuster A, Mason D, Xue Y, et al. Estimating the human mutation rate from autozygous segments reveals population differences in human mutational processes. Nat Commun 2017;8:303. doi:10.1038/s41467-017-00323-y.

[96]　Wetterstrand K. DNA Sequencing Costs: Data - National Human Genome Research Institute (NHGRI) 2018. https://www.genome.gov/sequencingcostsdata/ (accessed August 30, 2018).

[97]　Price AL, Kryukov G V., de Bakker PIW, Purcell SM, Staples J, Wei L-J, et al. Pooled Association Tests for Rare Variants in Exon-Resequencing Studies. Am J Hum Genet 2010;86:832–8. doi:10.1016/j.ajhg.2010.04.005.

[98]　Whiffin N, Minikel E, Walsh R, O'Donnell-Luria AH, Karczewski K, Ing AY, et al. Using high-resolution variant frequencies to empower clinical genome interpretation. Genet Med 2017;19:1151–8. doi:10.1038/gim.2017.26.

[99]　Soares-Souza G, Borda V, Kehdy F, Tarazona-Santos E. Admixture, Genetics and Complex Diseases in Latin Americans and US Hispanics. Curr Genet Med Rep 2018;6:208–23. doi:10.1007/s40142-018-0151-z.

[100]　Suarez-Kurtz G, Struchiner CJ. Differential effects of predictors of warfarin dose according to race/color categories in the admixed Brazilian population. Pharmacogenet Genomics 2017;27:210–1. doi:10.1097/FPC.0000000000000273.

[101]　Duconge J, Ramos AS, Claudio-Campos K, Rivera-Miranda G, Bermúdez-Bosch L, Renta JY, et al. A Novel Admixture-Based Pharmacogenetic Approach to Refine

Warfarin Dosing in Caribbean Hispanics. PLoS One 2016;11:e0145480.

doi:10.1371/journal.pone.0145480.

[102] Marcus JH, Novembre J. Visualizing the geography of genetic variants.

Bioinformatics 2016;33:btw643. doi:10.1093/bioinformatics/btw643.

37

Fig. 1. Evolutionary scenarios affecting drug response in humans.

A. Allelic frequency spectrum (AFS) expected under the *exogenous-endogenous* model of evolution. Molecules mainly involved in endogenous pathways (left, e.g. drug targets) are most probably under purifying selection (red) with a skew of the AFS towards rare polymorphic sites compared to neutrality (white). Selective constraint on molecules with exogenous substrates (right, e.g. DMEs) are moderate, thus the AFS is expected to follow a neutral distribution (grey). Alternatively, positive selection may favor diversification and may increase common variants (green). In all cases, a demographic model of a population of large and constant size with no population subdivision is assumed. The effect of demography on AFS is discussed in the text.
B. Influence of human lifestyle on PGx genetic variation. For genes such as *NAT2* and *CYP2D6,* hunter-gatherers (brown) are predicted to be faster metabolizers compared to food producers (green)[11,12].
C. Positive selection due to landscape variables according to the sodium retention hypothesis[51]. *CYP3A5* rs776746 (*3*): the ancestral nucleotide C is the minor allele outside Africa, with the exception of admixed American populations. The derived T allele (null-function) is almost fixed in Europe. Allelic frequency from the 1000genomes project; Distribution of Plot generated with Geography of Genetic Variants Browser [102].

39