PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Review

Cite this article: Fuselli S. 2019 Beyond drugs: the evolution of genes involved in human response to medications. *Proc. R. Soc. B* **286**: 20191716.
http://dx.doi.org/10.1098/rspb.2019.1716

Received: 22 July 2019
Accepted: 30 September 2019

Subject Category:
Genetics and genomics

Subject Areas:
evolution, genetics, health and disease and epidemiology

Keywords:
evolution, pharmacogenomics, drug-metabolizing enzymes, rare genetic variants, natural selection

Author for correspondence:
Silvia Fuselli
e-mail: fss@unife.it

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.4690451.

THE ROYAL SOCIETY PUBLISHING

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

SF, 0000-0002-1442-6171

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 the 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*), thiopurines (*TPMT*) and tacrolimus (*CYP3A5*) (pharmagkb.org, [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 programmes 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 range of variation in DME 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: (i) the knowledge on HIV evolution is essential for the design of preventive and therapeutic measures [14]; (ii) 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; and (iii) 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 show extreme levels of single nucleotide polymorphism (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 50–80 CYP450 genes in 10 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 (GST) 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, because 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. Predictions based on this model may be useful for pharmacogenetic applications (figure 1*a*). For example, patterns of genetic variation of 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 [28]. 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 [18,29,30].

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 defence, such as those coding for the major histocompatibility complex of vertebrates, which typically evolve under long-term balancing selection, or the lysozyme of primates, evolving under positive selection in specific phylogenetic lineages [31,32]. 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 [33]. 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* is more common in groups consuming a hypercarnivore diet (i.e. a diet comprising of more than 70% meat) [34]. Because the rate of nucleotide substitutions in genes is inversely related to the stringency of the functional constraints acting upon them [35], genes that code for enzymes which no longer have a metabolic role are expected to accumulate mutations and become pseudogenes [36].

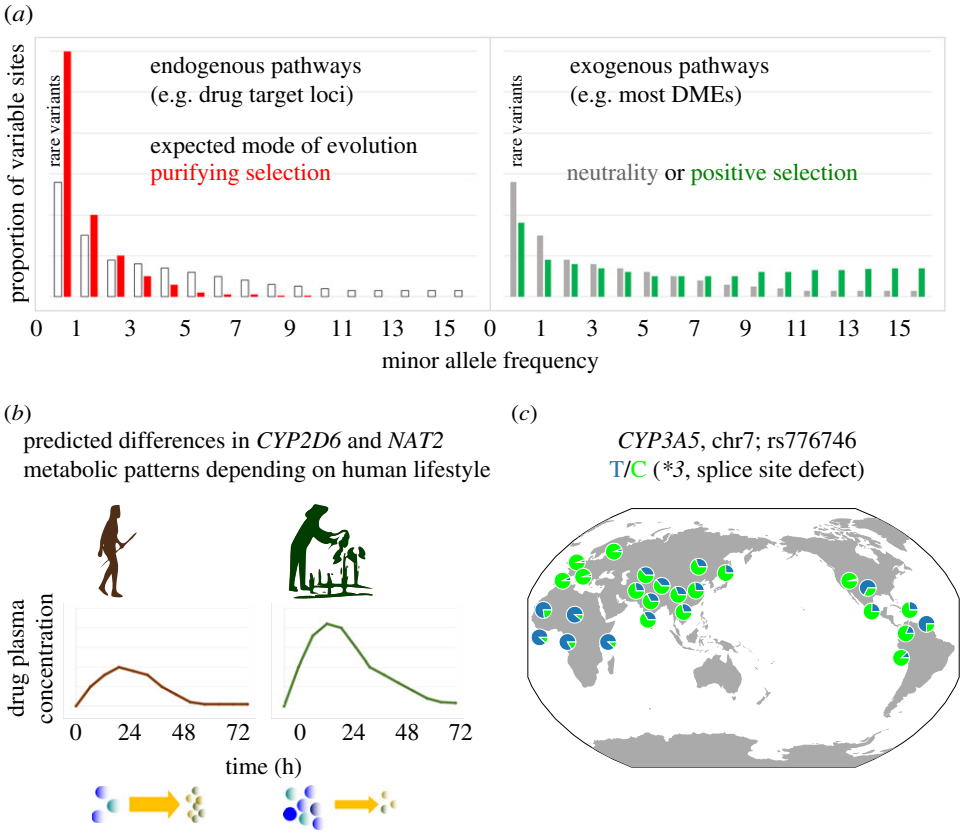

Figure 1. Evolutionary scenarios affecting drug response in humans. (*a*) Allelic frequency spectrum (AFS) expected under the *exogenous–endogenous* model of evolution. Molecules mainly involved in endogenous pathways (left), e.g. drug targets are most probably under purifying selection (red) with a 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 favour 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 owing to landscape variables according to the sodium retention hypothesis [26]. *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 1000 genomes project; distribution of plot generated with geography of genetic variants browser [27]. (Online version in colour.)

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 [34].

In birds, whole-genome comparisons of 48 species with different feeding habits, geographical distribution and migratory behaviour revealed the influence of these variables on the CYP2 enzyme family evolution [37]. 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 frequency. 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, probably leading to an increased or a decreased metabolic effect [38]. 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 *CYP2 J* families are expanded. Because 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 [39].

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

Human populations subsisted by hunting wild animals and gathering wild plants for about 90% of their history. Between 10 000 and 4000 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 [40]. All these changes left a footprint in our genome [41–43]. 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 (electronic supplementary material, table S1; figure 1*b*). Most of these studies focused on the gene *NAT2* coding for the Phase II DME NAT 2. Detailed information on the *NAT2* genetic variation is available for all geographical regions across the world [44]. 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 [45]. 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 [45]. 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, the 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 [46,47] invoke a non-neutral explanation for a reduction in the NAT2 enzymatic activity associated with 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 CYP450 2D6. By controlling for environment and demography, a comparison of 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 1*b*). 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 [43]. Ancient DNA allows measurement of 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 [48].

(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 [49]. Genome scans in humans have occasionally identified PGx loci as the target of recent positive selection, including signals at the *SLC24A5* locus in non-African populations, possibly owing to its association with skin pigmentation [50,51], or at the *CYP3A4* and *CYP3A5* loci [52]. 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 commonly used drugs (pharmgkb.org, cpicpgx.org). A whole-gene resequencing study conducted by Thompson *et al*. [26] in three populations from different continents, combined with the screening of the null-function allele in more than 1000 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 (rs776746T>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 1*c*). Few other human genes show the same pattern that was previously defined as the 'west Eurasia sweep' [51]. Interestingly, the proposed selective mechanism involves the endogenous role of the 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 [53], the allele coding for a full-function protein confers a selective advantage in equatorial populations where water shortages are common [54]. 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 [55]. 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 [26,56]. The case of the *CYP3A* cluster illustrates how positive selection can give rise to an evident global geographical structure of genetic variation in response to environmental variables (figure 1*c*). At the other extreme, selection may lead to detectable allele frequency change in a restricted geographical region, as in the case of the Arsenite Methyltransferase *AS3MT* in the populations residing near San Antonio de los Cobres, Argentina. Here, individuals have probably 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*. [57] 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 [58]. Fungi and insects adapt to pesticides in a matter of years, offering very good examples of acquired resistance via DMEs' evolution [59,60]. Resistance can result from the rapid increase in the frequency of an advantageous mutation conferring less sensitivity to a specific molecule [61], or from an increased production of the detoxifying enzyme owing to overtranscription or CNV [62]. 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 [63,64]. In humans, *VKORC1* and

CYP2C9 polymorphisms are included in a genotype-guided dosage of this widely prescribed anticoagulant [65], which was also used as a rodenticide in Europe since the 1950s. Genetic analyses have shown how vkorc1-mediated resistance convergently evolved in different rodent pest species in a few years through selection on adaptive genetic variants [63]. Interestingly, the house mice *Mus musculus domesticus* have acquired adaptive vkorc1 polymorphisms from the Algerian mouse (*Mus spretus*) through introgressive hybridization [66]. 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 owing to the strong selection coefficient involved in cancer progression and resistance to chemotherapy is an important example of artificial selection in our species [67]. 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 are governed entirely by tissue control mechanisms acting through intracellular pathways. Conversely, somatic cells isolated from normal tissue controls can evolve towards cancer and their survival and proliferation are governed by the interaction of their own properties with the environment [68]. Therapeutic resistance can then arise from one or several subclones harbouring resistance-conferring mutations, which can be either pre-existing (i.e. present at diagnosis) or acquired. Both mechanisms have been observed, for example, in acute myeloid leukaemia and urothelial carcinoma [69,70]. Relapse after remission generated from minor genetic subclones already present at diagnosis is particularly interesting from an evolutionary perspective. These subclones, which were carrying the resistance-conferring mutation but were outcompeted in the absence of therapy, will eventually be positively selected by treatment and expand [71]. Understanding tumour 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 [72].

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. Owing to the low frequency, rare variants tend to be geographically restricted or even private to a population [73], so their predictive role in drug response, for example, is not equally important in different human populations [74]. Additionally, before the advent of next-generation sequencing (NGS) technologies, the lack of a 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 [75]. The development of NGS has enabled researchers to obtain information on both common and rare genetic variants, overcoming many intrinsic biases of an SNP typing approach. With a slight delay, this change has also involved pharmacogenetics and pharmacogenomics [10,76] (electronic supplementary material, table S2). In the 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 [77] (figure 1*a*). During the last 10 000 years, our species has experienced an explosive population growth, from a few million people to roughly 7 billion today [78,79]. Rapid population growth is 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 [80]. 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) [81]. 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 [82] (figure 1*a*). Consistent with these expectations, a study on 6515 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 5000 years [83]. Conversely, in small populations, the strength of selection is reduced and the fate of each mutation, even if deleterious, and especially if slightly deleterious, is mainly governed by random genetic drift. It is therefore expected that small 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 [80]. 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 [83,84]. Altogether, in terms of the 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 [82,84].

(b) Examples of studies on rare genetic PGx variants

In the last few years, several studies have focused on the importance of low-frequency genetic variants in PGx genes (electronic supplementary material, table S2). 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 [85–88]). Most of these studies analysed PGx or ADME (absorption, distribution, metabolism and excretion) loci as a whole (electronic supplementary material, table S2). 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. With the aim of determining the abundance, distribution and phenotypic effects of rare variants in drug-related phenotypes, studies focused on a more homogeneous group of genes may be more informative. For example, in the study by Nelson *et al*. [89], 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 a 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 1*a*). The excess of rare genetic variants that characterizes the large European subset analysed in the study (12 514 individuals) is concentrated in gene regions with a more likely functional effect (i.e. untranslated regions and non-synonymous variants), which are under-represented 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 analysed in DMEs. For example, Gordon *et al*. [90] 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 [91,92]. 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*. [93]). In terms of the proportion of variants with functional effect, the 12 *CYP* genes appear less constrained than drug target genes [89] (figure 1*a*). However, for rare variants, target genes and the 12 *CYP* genes show similar 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 analysed in the general context of PGx genes [85–87,93] (electronic supplementary material, table S2). Two notable exceptions concern the *SLCO* gene family. The first is a detailed analysis of rare genetic variation in the gene *SLCO1B1* [94], 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 the 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 [94]. Recently, a comprehensive analysis of *SLCO* variability by Zhang & Lauschke [95] 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.

Besides rare variants of known or easy-to-predict phenotypic effect, exome and genome sequencing generate lists of rare variants that harbour one or more potentially deleterious properties (e.g. change the amino acid and have never been observed in other species), but whose functional consequences remain unclear. These variants are therefore considered of uncertain significance (VUS) and challenge the clinical implementation of pharmacogenomics [88]. Principles of population genetics and molecular evolution described at the beginning of this fourth paragraph, together with comparative genomics approaches, can guide the categorization of VUS into subgroups with more or less probable functional effect [96].

5. Conclusion 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 the drug. The following key points summarize the important aspects discussed in this review and delineate their implications for pharmacogenetics and pharmacogenomics approaches.

(i) 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 [30].

(ii) 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 tumour cells and the strong selective coefficient involved in chemotherapy, the application of evolutionary principles in cancer treatment is particularly promising [97]. 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 intratumoural competition [98].

(iii) Local adaptation can generate a gene-specific spatial distribution of detrimental alleles [41]. 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 geographical regions (figure 1*c*). On the other hand, convergent adaptation, or relaxation of natural selection, may be responsible for the presence of slow-function alleles in different geographical regions for genes such as *CYP2D6* and *NAT2* (figure 1*b*). 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].

(iv) As theoretically postulated, rare damaging PGx genetic variants, as well as newly discovered VUS, 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 [99], the number of mutations arising each

generation is substantial, at least 100 billion mutations. As sequencing costs continue to decline, analysing rare and private variants on an individual's genome will become feasible [100]. Until then, a reasonable compromise could be to focus on variants segregating at a 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 [101,102].

Finally, it is important and timely to mention the additional level of complexity owing to population admixture. Historically admixed populations such as African-Americans, Brazilians and Caribbeans, but also new mixing because of 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 [103]. 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 [104]. In clinical pharmacogenetics, it has been suggested that the inclusion of admixture inferences improve dosage algorithm for warfarin in Puerto Rico [105]. 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.

Data accessibility. This article has no additional data.

Competing interests. The author declares that she has no competing interests.

Funding. The author received no funding for this study.

Acknowledgements. The author thanks Giorgio Bertorelle, Eduardo Tarazona-Santos, Estella Poloni, Pierpaolo Maisano Delser and Francesca Tassi for helpful discussion and comments on the manuscript. She also thanks the two anonymous reviewers for their useful comments.

References

1. Weinshilboum R. 2003 Inheritance and drug response. *N. Engl. J. Med.* **348**, 529–537. (doi:10.1056/NEJMra020021)
2. Meyer UA, Zanger UM, Schwab M. 2013 Omics and drug response. *Annu. Rev. Pharmacol. Toxicol.* **53**, 475–502. (doi:10.1146/annurev-pharmtox-010510-100502)
3. Maisano Delser P, Fuselli S. 2013 Human loci involved in drug biotransformation: worldwide genetic variation, population structure, and pharmacogenetic implications. *Hum. Genet.* **132**, 563–577. (doi:10.1007/s00439-013-1268-5)
4. Giacomini KM *et al.* 2010 Membrane transporters in drug development. *Nat. Rev. Drug Discov.* **9**, 215–236. (doi:10.1038/nrd3028)
5. Santos R *et al.* 2017 A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34. (doi:10.1038/nrd.2016.230)
6. Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, Altman RB, Klein TE. 2012 Pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Ther.* **92**, 414–417. (doi:10.1038/clpt.2012.96)
7. Giacomini KM, Yee SW, Mushiroda T, Weinshilboum RM, Ratain MJ, Kubo M. 2016 Genome-wide association studies of drug response and toxicity: an opportunity for genome medicine. *Nat. Rev. Drug Discov.* **16**, 1. (doi:10.1038/nrd.2016.234)
8. Luzum J *et al.* 2017 The pharmacogenomics research network translational pharmacogenetics program: outcomes and metrics of pharmacogenetic implementations across diverse healthcare systems. *Clin. Pharmacol. Ther.* **102**, 502–510. (doi:10.1002/cpt.630)
9. Cambon-Thomsen A *et al.* 2017 Implementing pharmacogenomics in Europe: design and implementation strategy of the ubiquitous pharmacogenomics consortium. *Clin. Pharmacol. Ther.* **101**, 341–358. (doi:10.1002/cpt.602)
10. Drogemoller BI, Wright GEB, Warnich L. 2014 Considerations for rare variants in drug metabolism genes and the clinical implications. *Expert Opin. Drug Metab. Toxicol.* **10**, 873–884. (doi:10.1517/17425255.2014.903239)
11. Fuselli S, de Filippo C, Mona S, Sistonen J, Fariselli P, Destro-Bisol G, Barbujani G, Bertorelle G, Sajantila A. 2010 Evolution of detoxifying systems: the role of environment and population history in shaping genetic diversity at human *CYP2D6* locus. *Pharmacogenet. Genomics* **20**, 485–499. (doi:10.1097/FPC.0b013e32833bba25)
12. Podgorná E, Diallo I, Vangenot C, Sanchez-Mazas A, Sabbagh A, Černý V, Poloni ES. 2015 Variation in NAT2 acetylation phenotypes is associated with differences in food-producing subsistence modes and ecoregions in Africa. *BMC Evol. Biol.* **15**, 1–20. (doi:10.1186/s12862-015-0543-6)
13. Wells JCK, Nesse RM, Sear R, Johnstone RA, Stearns SC. 2017 Evolutionary public health: introducing the concept. *Lancet* **390**, 500–509. (doi:10.1016/S0140-6736(17)30572-X)
14. Theys K, Libin P, Pineda-Peña A-C, Nowé A, Vandamme A-M, Abecasis AB. 2018 The impact of HIV-1 within-host evolution on transmission dynamics. *Curr. Opin. Virol.* **28**, 92–101. (doi:10.1016/j.coviro.2017.12.001)
15. Francino MP. 2015 Antibiotics and the human gut microbiome: dysbioses and accumulation of resistances. *Front. Microbiol.* **6**, 1543. (doi:10.3389/fmicb.2015.01543)
16. Simons YB, Bullaughey K, Hudson RR, Sella G. 2018 A population genetic interpretation of GWAS findings for human quantitative traits. *PLoS Biol.* **16**, e2002985. (doi:10.1371/journal.pbio.2002985)
17. Sezutsu H, le Goff G, Feyereisen R. 2013 Origins of P450 diversity. *Phil. Trans. R. Soc. B* **368**, 20120428. (doi:10.1098/rstb.2012.0428)
18. Moitra K, Dean M. 2011 Evolution of ABC transporters by gene duplication and their role in human disease. *Biol. Chem.* **392**, 29–37. (doi:10.1515/bc.2011.006)
19. Sabbagh A, Marin J, Veyssière C, Lecompte E, Boukouvala S, Poloni ES, Darlu P, Crouau-Roy B. 2013 Rapid birth-and-death evolution of the xenobiotic metabolizing NAT gene family in vertebrates with evidence of adaptive selection. *BMC Evol. Biol.* **13**, 62. (doi:10.1186/1471-2148-13-62)
20. Nei M, Rooney AP. 2005 Concerted and birth-and-death evolution of multigene families. *Annu. Rev. Genet.* **39**, 121–152. (doi:10.1146/annurev.genet.39.073003.112240)
21. Thomas JH. 2007 Rapid birth-death evolution specific to xenobiotic cytochrome P450 genes in vertebrates. *PLoS Genet.* **3**, e67. (doi:10.1371/journal.pgen.0030067)
22. Drosophila 12 Genomes Consortium *et al.* 2007 Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* **450**, 203–218. (doi:10.1038/nature06341)
23. Good RT, Gramzow L, Battlay P, Sztal T, Batterham P, Robin C. 2014 The molecular evolution of cytochrome P450 genes within and between *Drosophila* species. *Genome Biol. Evol.* **6**, 1118–1134. (doi:10.1093/gbe/evu083)
24. Hein DW. 2002 Molecular genetics and function of NAT1 and NAT2: role in aromatic amine metabolism and carcinogenesis. *Mutat. Res.* **506–507**, 65–77. (doi:10.1016/S0027-5107(02)00153-7)

25. Patin E *et al.* 2006 Deciphering the ancient and complex evolutionary history of human arylamine N-acetyltransferase genes. *Am. J. Hum. Genet.* **78**, 423–436. (doi:10.1086/500614)
26. Thompson EE, Kuttab-Boulos H, Witonsky D, Yang L, Roe BA, Di Rienzo A. 2004 CYP3A variation and the evolution of salt-sensitivity variants. *Am. J. Hum. Genet.* **75**, 1059–1069. (doi:10.1086/426406)
27. Marcus JH, Novembre J. 2016 Visualizing the geography of genetic variants. *Bioinformatics* **33**, btw643. (doi:10.1093/bioinformatics/btw643)
28. Chatterjee A, Gupta S. 2018 The multifaceted role of glutathione S-transferases in cancer. *Cancer Lett.* **433**, 33–42. (doi:10.1016/j.canlet.2018.06.028)
29. Zhu C, Nigam KB, Date RC, Bush KT, Springer SA, Saier MH, Wu W, Nigam SK. 2015 Evolutionary analysis and classification of OATs, OCTs, OCTNs, and other SLC22 transporters: structure-function implications and analysis of sequence motifs. *PLoS ONE* **10**, e0140569. (doi:10.1371/journal.pone.0140569)
30. Lauschke VM, Ingelman-Sundberg M. 2019 Prediction of drug response and adverse drug reactions: from twin studies to next generation sequencing. *Eur. J. Pharm. Sci.* **130**, 65–77. (doi:10.1016/j.ejps.2019.01.024)
31. Garrigan D, Hedrick PW. 2003 Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution* **57**, 1707–1722. (doi:10.1111/j.0014-3820.2003.tb00580.x)
32. Yang Z. 1998 Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol. Biol. Evol.* **15**, 568–573. (doi:10.1093/oxfordjournals.molbev.a025957)
33. Kim S *et al.* 2016 Comparison of carnivore, omnivore, and herbivore mammalian genomes with a new leopard assembly. *Genome Biol.* **17**, 1–12. (doi:10.1186/s13059-015-0866-z)
34. Shrestha B, Reed JM, Starks PT, Kaufman GE, Goldstone JV, Roelke ME, O'Brien SJ, Koepfli KP, Frank LG. 2011 Evolution of a major drug metabolizing enzyme defect in the domestic cat and other Felidae: phylogenetic timing and the role of hypercarnivory. *PLoS ONE* **6**, e18046. (doi:10.1371/journal.pone.0018046)
35. Jukes TH, Kimura M. 1984 Evolutionary constraints and the neutral theory. *J. Mol. Evol.* **21**, 90–92. (doi:10.1007/BF02100633)
36. Calderoni L, Rota-Stabelli O, Frigato E, Panziera A, Kirchner S, Foulkes NS, Kruckenhauser L, Bertolucci C, Fuselli S. 2016 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.*) **117**, 383–392. (doi:10.1038/hdy.2016.59)
37. Almeida D *et al.* 2016 Whole-genome identification, phylogeny, and evolution of the cytochrome P450 family 2 (CYP2) subfamilies in birds. *Genome. Biol. Evol.* **8**, 1115–1131. (doi:10.1093/gbe/evw041)
38. He Y, Hoskins JM, Mcleod HL. 2011 Copy number variants in pharmacogenetic genes. *Trends Mol. Med.* **17**, 244–251. (doi:10.1016/j.molmed.2011.01.007)
39. Sequencing TB *et al.* 2012 Genome sequences of wild and domestic bactrian camels. *Nat. Commun.* **3**, 1202. (doi:10.1038/ncomms2192)
40. Bellwood PS. 2005 *First farmers: the origins of agricultural societies*. Oxford, UK: Blackwell Publications.
41. Fan S, Hansen MEB, Lo Y, Tishkoff SA. 2016 Going global by adapting local: a review of recent human adaptation. *Science* **354**, 54–59. (doi:10.1126/science.aaf5098)
42. Lopez M *et al.* 2018 The demographic history and mutational load of African hunter-gatherers and farmers. *Nat. Ecol. Evol.* **2**, 721–730. (doi:10.1038/s41559-018-0496-4)
43. Marciniak S, Perry GH. 2017 Harnessing ancient genomes to study the history of human adaptation. *Nat. Rev. Genet.* **18**, 659–674. (doi:10.1038/nrg.2017.65)
44. Sabbagh A, Langaney A, Darlu P, Gérard N, Krishnamoorthy R, Poloni ES. 2008 Worldwide distribution of NAT2 diversity: implications for NAT2 evolutionary history. *BMC Genet.* **9**, 1–14. (doi:10.1186/1471-2156-9-21)
45. Sabbagh A, Darlu P, Vangenot C, Poloni ES. 2018 Arylamine *N*-acetyltransferases in anthropology. In *Arylamine N-acetyltransferases in health and disease*. From pharmacogenetics to drug discovery (eds N Laurieri, E Sim), pp. 165–193. Singapore: World Scientific.
46. Luca F *et al.* 2008 Multiple advantageous amino acid variants in the NAT2 gene in human populations. *PLoS ONE* **3**, e3136. (doi:10.1371/journal.pone.0003136)
47. Magalon H, Patin E, Austerlitz F, Hegay T, Aldashev A, Quintana-Murci L, Heyer E. 2008 Population genetic diversity of the NAT2 gene supports a role of acetylation in human adaptation to farming in Central Asia. *Eur. J. Hum. Genet.* **16**, 243–251. (doi:10.1038/sj.ejhg.5201963)
48. Mathieson S, Mathieson I. 2018 *FADS1* and the timing of human adaptation to agriculture. *Mol. Biol. Evol.* **35**, 2957–2970. (doi:10.1093/molbev/msy180)
49. Nielsen R, Hellmann I, Hubisz M, Bustamante C, Clark AG. 2007 Recent and ongoing selection in the human genome. *Nat. Rev. Genet.* **8**, 857–868. (doi:10.1038/nrg2187)
50. Voight BF, Kudaravalli S, Wen X, Pritchard JK. 2006 A map of recent positive selection in the human genome. *PLoS Biol.* **4**, e72. (doi:10.1371/journal.pbio.0040072)
51. Coop G *et al.* 2009 The role of geography in human adaptation. *PLoS Genet.* **5**, e1000500. (doi:10.1371/journal.pgen.1000500)
52. Carlson CS, Thomas DJ, Eberle MA, Swanson JE, Livingston RJ, Rieder MJ, Nickerson DA. 2005 Genomic regions exhibiting positive selection identified from dense genotype data. *Genome Res.* **1**, 1553–1565. (doi:10.1101/gr.4326505)
53. Gleibermann L. 1973 Blood pressure and dietary salt in human populations. *Ecol. Food Nutr.* **2**, 143–156. (doi:10.1080/03670244.1973.9990329)
54. Kuehl P *et al.* 2001 Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. *Nat. Genet.* **27**, 383–391. (doi:10.1038/86882)
55. Sibai B, Dekker G, Kupferminc M. 2005 Pre-eclampsia. *Lancet* (*London*) *England* **365**, 785–799. (doi:10.1016/S0140-6736(05)17987-2)
56. Young JH, Chang Y-PC, Kim JD-O, Chretien J-P, Klag MJ, Levine MA, Ruff CB, Wang NY, Chakravarti A. 2005 Differential susceptibility to hypertension is due to selection during the out-of-Africa expansion. *PLoS Genet.* **1**, e82. (doi:10.1371/journal.pgen.0010082)
57. Schlebusch CM, Gattepaille LM, Engström K, Vahter M, Jakobsson M, Broberg K. 2015 Human adaptation to arsenic-rich environments. *Mol. Biol. Evol.* **32**, 1544–1555. (doi:10.1093/molbev/msv046)
58. Nielsen R, Slatkin M. 2013 *Chapter 7 selection I.* In *An introduction to population genetics*, pp. 129–151. Sunderland, MA: Sinauer Associates.
59. Kelly SL, Kelly DE. 2013 Microbial cytochromes P450: biodiversity and biotechnology. Where do cytochromes P450 come from, what do they do and what can they do for us? *Phil. Trans. R. Soc. B* **368**, 20120476. (doi:10.1098/rstb.2012.0476)
60. Schmidt JM, Battlay P, Gledhill-Smith RS, Good RT, Lumb C, Fournier-Level A, Robin C. 2017 Insights into DDT resistance from the *Drosophila melanogaster* genetic reference panel. *Genetics* **207**, 1181–1193. (doi:10.1534/genetics.117.300310)
61. Nash D, Nair S, Mayxay M, Newton PN, Guthmann J-P, Nosten F, Anderson TJ. 2005 Selection strength and hitchhiking around two anti-malarial resistance genes. *Proc. R. Soc. B* **272**, 1153–1161. (doi:10.1098/rspb.2004.3026)
62. Bass C, Field LM. 2011 Gene amplification and insecticide resistance. *Pest Manag. Sci.* **67**, 886–890. (doi:10.1002/ps.2189)
63. Pelz H-J *et al.* 2005 The genetic basis of resistance to anticoagulants in rodents. *Genetics* **170**, 1839–1847. (doi:10.1534/genetics.104.040360)
64. Pirmohamed M, Kamali F, Daly AK, Wadelius M. 2015 Oral anticoagulation: a critique of recent advances and controversies. *Trends Pharmacol. Sci.* **36**, 153–163. (doi:10.1016/j.tips.2015.01.003)
65. Van Driest SL, Shi Y, Bowton EA, Schildcrout JS, Peterson JF, Pulley J, Denny JC, Roden DM. 2014 Clinically actionable genotypes among 10,000 patients with preemptive pharmacogenomic testing. *Clin. Pharmacol. Ther.* **95**, 423–431. (doi:10.1038/clpt.2013.229)
66. Song Y, Endepols S, Klemann N, Richter D, Matuschka FR, Shih CH, Nachman MW, Kohn MH. 2011 Adaptive introgression of anticoagulant rodent poison resistance by hybridization between old world mice. *Curr. Biol.* **21**, 1296–1301. (doi:10.1016/j.cub.2011.06.043)
67. Burrell RA, Swanton C. 2014 Tumour heterogeneity and the evolution of polyclonal drug resistance. *Mol. Oncol.* **8**, 1095–1111. (doi:10.1016/j.molonc.2014.06.005)

68. Shpak M, Lu J. 2016 An evolutionary genetic perspective on cancer biology. *Annu. Rev. Ecol. Evol. Syst.* **47**, 25–49. (doi:10.1146/annurev-ecolsys-121415-032109)
69. Shlush LI *et al.* 2017 Tracing the origins of relapse in acute myeloid leukaemia to stem cells. *Nature* **547**, 104–108. (doi:10.1038/nature22993)
70. Vlachostergios PJ, Faltas BM. 2018 Treatment resistance in urothelial carcinoma: an evolutionary perspective. *Nat. Rev. Clin. Oncol.* **15**, 495–509. (doi:10.1038/s41571-018-0026-y)
71. Greaves M, Maley CC. 2012 Clonal evolution in cancer. *Nature* **481**, 306–313. (doi:10.1038/nature10762)
72. Sottoriva A, Spiteri I, Piccirillo SGM, Touloumis A, Collins VP, Marioni JC, Curtis C, Watts C, Tavaré S. 2013 Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc. Natl Acad. Sci. USA* **110**, 4009–4014. (doi:10.1073/pnas.1219747110)
73. Gravel S, Henn BM, Gutenkunst RN, Indap AR, Marth GT, Clark AG, Yu F, Gibbs RA, Bustamante CD. 2011 Demographic history and rare allele sharing among human populations. *Proc. Natl Acad. Sci. USA* **108**, 11 983–11 988. (doi:10.1073/pnas.1019276108)
74. Apellániz-Ruiz M *et al.* 2015 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.* **15**, 288–292. (doi:10.1038/tpj.2014.67)
75. International HapMap Consortium *et al.* 2007 A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861. (doi:10.1038/nature06258)
76. Ingelman-Sundberg M, Mkrtchian S, Zhou Y, Lauschke VM. 2018 Integrating rare genetic variants into pharmacogenetic drug response predictions. *Hum. Genomics* **12**, 26. (doi:10.1186/s40246-018-0157-3)
77. Evans SN, Shvets Y, Slatkin M. 2007 Non-equilibrium theory of the allele frequency spectrum. *Theor. Popul. Biol.* **71**, 109–119. (doi:10.1016/J.TPB.2006.06.005)
78. Keinan A, Clark AG. 2012 Recent explosive human population growth has resulted in an excess of rare genetic variants. *Science* **336**, 740–743. (doi:10.1126/science.1217283)
79. Roberts L. 2011 9 billion? *Science* **333**, 540–543. (doi:10.1126/science.333.6042.540)
80. Ohta T. 1992 The nearly neutral theory of molecular evolution. *Annu. Rev. Ecol. Syst.* **23**, 263–286. (doi:10.1146/annurev.es.23.110192.001403)
81. Gravel S. 2016 When is selection effective? *Genetics* **203**, 451–462. (doi:10.1534/genetics.115.184630)
82. Henn BM, Botigué LR, Bustamante CD, Clark AG, Gravel S. 2015 Estimating the mutation load in human genomes. *Nat. Rev. Genet.* **16**, 333–343. (doi:10.1038/nrg3931)
83. Fu W *et al.* 2013 Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature* **493**, 216–220. (doi:10.1038/nature11690)
84. Lohmueller KE *et al.* 2008 Proportionally more deleterious genetic variation in European than in African populations. *Nature* **451**, 994–997. (doi:10.1038/nature06611)
85. Hovelson DH *et al.* 2017 Characterization of ADME gene variation in 21 populations by exome sequencing. *Pharmacogenet. Genomics* **27**, 89–100. (doi:10.1097/FPC.0000000000000260)
86. Han SM *et al.* 2017 Targeted next-generation sequencing for comprehensive genetic profiling of pharmacogenes. *Clin. Pharmacol. Ther.* **101**, 396–405. (doi:10.1002/cpt.532)
87. Bush WS *et al.* 2016 Genetic variation among 82 pharmacogenes: the PGRNseq data from the eMERGE network. *Clin. Pharmacol. Ther.* **100**, 160–169. (doi:10.1002/cpt.350)
88. Daly A, Rettie A, Fowler D, Miners J. 2017 Pharmacogenomics of CYP2C9: functional and clinical considerations. *J. Pers. Med.* **8**, 1. (doi:10.3390/jpm8010001)
89. Nelson MR *et al.* 2012 An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. *Science* **336**, 100–104. (doi:10.1126/science.1217876)
90. Gordon AS *et al.* 2014 Quantifying rare, deleterious variation in 12 human cytochrome P450 drug-metabolism genes in a large-scale exome dataset. *Hum. Mol. Genet.* **23**, 1957–1963. (doi:10.1093/hmg/ddt588)
91. Evans WE, Relling MV. 1999 Pharmacogenomics: translating functional genomics into rational therapeutics. *Science* **286**, 487–491. (doi:10.1126/science.286.5439.487)
92. Sim SC, Kacevska M, Ingelman-Sundberg M. 2013 Pharmacogenomics of drug-metabolizing enzymes: a recent update on clinical implications and endogenous effects. *Pharmacogenomics J.* **13**, 1–11. (doi:10.1038/tpj.2012.45)
93. Zhou Y, Ingelman-Sundberg M, Lauschke VM. 2017 Worldwide distribution of cytochrome P450 alleles: a meta-analysis of population-scale sequencing projects. *Clin. Pharmacol. Ther.* **102**, 688–700. (doi:10.1002/cpt.690)
94. Ramsey LB *et al.* 2012 Rare versus common variants in pharmacogenetics: *SLCO1B1* variation and methotrexate disposition. *Genome Res.* **22**, 1–8. (doi:10.1101/gr.129668.111)
95. Zhang B, Lauschke VM. 2019 Genetic variability and population diversity of the human SLCO (OATP) transporter family. *Pharmacol. Res.* **139**, 550–559. (doi:10.1016/J.PHRS.2018.10.017)
96. Cooper GM. 2015 Parlez-vous VUS? *Genome Res.* **25**, 1423–1426. (doi:10.1101/gr.190116.115)
97. Gatenby R, Brown J. 2018 The evolution and ecology of resistance in cancer therapy. *Cold Spring Harb. Perspect. Med.* **8**, a033415. (doi:10.1101/cshperspect.a033415)
98. Silva AS, Kam Y, Khin ZP, Minton SE, Gillies RJ, Gatenby RA. 2012 Evolutionary approaches to prolong progression-free survival in breast cancer. *Cancer Res.* **72**, 6362–6370. (doi:10.1158/0008-5472.CAN-12-2235)
99. Narasimhan VM *et al.* 2017 Estimating the human mutation rate from autozygous segments reveals population differences in human mutational processes. *Nat. Commun.* **8**, 303. (doi:10.1038/s41467-017-00323-y)
100. Wetterstrand K. 2018 DNA sequencing costs: data – National Human Genome Research Institute (NHGRI) 2018. See https://www.genome.gov/sequencingcostsdata/ (accessed 30 August, 2018).
101. Price AL, Kryukov GV, de Bakker PIW, Purcell SM, Staples J, Wei L-J, Sunyaev SR. 2010 Pooled association tests for rare variants in exon-resequencing studies. *Am. J. Hum. Genet.* **86**, 832–838. (doi:10.1016/j.ajhg.2010.04.005)
102. Whiffin N *et al.* 2017 Using high-resolution variant frequencies to empower clinical genome interpretation. *Genet. Med.* **19**, 1151–1158. (doi:10.1038/gim.2017.26)
103. Soares-Souza G, Borda V, Kehdy F, Tarazona-Santos E. 2018 Admixture, genetics and complex diseases in Latin Americans and US Hispanics. *Curr. Genet. Med. Rep.* **6**, 208–223. (doi:10.1007/s40142-018-0151-z)
104. Suarez-Kurtz G, Struchiner CJ. 2017 Differential effects of predictors of warfarin dose according to race/color categories in the admixed Brazilian population. *Pharmacogenet. Genomics* **27**, 210–211. (doi:10.1097/FPC.0000000000000273)
105. Duconge J *et al.* 2016 A novel admixture-based pharmacogenetic approach to refine warfarin dosing in Caribbean Hispanics. *PLoS ONE* **11**, e0145480. (doi:10.1371/journal.pone.0145480)