REVIEWS

PHARMACOGENETICS GOES GENOMIC

David B. Goldstein*, Sarah K. Tate* and Sanjay M. Sisodiya‡

Most people in the developed world will sooner or later be given prescription drugs to treat common diseases or to reduce the risk of getting them. Almost everyone who takes medicines will, at some stage, encounter those that do not work as well as they do in other people or even that cause an adverse reaction. Pharmacogenetics seeks to reduce the variation in how people respond to medicines by tailoring therapy to individual genetic make-up. It seems increasingly likely that investment in this field might be the most effective strategy for rapidly delivering the public health benefits that are promised by the Human Genome Project and related endeavours.

*Department of Biology (Galton Laboratory), University College London, The Darwin Building, Gower Street, London WC1E 6BT, UK. ‡Department of Clinical and Experimental Epilepsy, Institute of Neurology, University College London, London WC1N 3BG, UK. Correspondence to D.B.G. e-mail: d.goldstein@ucl.ac.uk

Pharmacogenetics is an old discipline, with antecedents that stretch back to the beginning of the twentieth century[1]. Like many other branches of the biomedical sciences it has been invigorated by recent advances in genomics, which have led to expectations that the safety and efficacy of medicines will soon be notably improved by genetic personalization. The excitement has also generated yet another compound name: pharmacogenomics. Although often used interchangeably, pharmacogenetics and pharmacogenomics have different connotations and a range of alternative definitions have been offered. Some have suggested that the difference is just in scale and that pharmacogenetics implies the study of a single gene whereas pharmacogenomics implies the study of many genes or entire genomes. Others have suggested that pharmacogenomics covers levels above that of DNA, such as mRNA or proteins, or that it relates more to drug development than does pharmacogenetics. These distinctions are variable and fuzzy, and are probably not worth formalizing. Here, we use the term pharmacogenetics, which has historical priority, according to its broadest meaning — relating heritable variation to inter-individual variation in drug response.

Whatever the name, efforts to personalize medicines have entered into a new era. For ~40 years, pharmacogenetics focused on describing variation, either genetically or biochemically, in a handful of proteins and genes.

Today, it is possible to assess entire pathways that might be relevant to disease or to drug responses at the DNA, mRNA and protein levels. Eventually, the entire genome, transcriptome and proteome will be within reach. As a consequence, pharmacogenetics and disease genetics are undergoing similar transitions, with a shift in focus from Mendelian examples to more complex modes of genetic causation.

Here, we assess the state of pharmacogenetic research and use this assessment to make recommendations for areas that require increased attention and investment. We also review the ethical and economic implications of pharmacogenetic research. Our overall conclusion is that most of the supposed obstacles to the clinical application of pharmacogenetics can be overcome, but that some refocusing and expansion of research effort will be required to realize the potential of pharmacogenetics in a reasonable timeframe. In particular, there is an urgent need to foster integrated pharmacogenetics research programmes, bringing together clinicians and basic scientists, and to greatly expand the coverage of pharmacogenetics in medical curricula.

Pre- and post-genomic pharmacogenetics

The intellectual foundations of pharmacogenetics were first articulated by Arno Motulsky in 1957 (REF. 2). Referring to examples such as sensitivity to the anti-malarial drug primaquine and the muscle relaxant

succinylcholine, Motulsky argued that "otherwise innocuous genetic traits" might underlie variation among individuals in drug response. The challenge now is to find these gene variants, to understand how they interact with one another and the environment, and to adjust therapy accordingly.

The early years. Early pharmacogenetic studies focused on variants with Mendelian effects on response. One important early example is the classical DEBRISOQUINE poor-metabolizer phenotype. In the 1970s, Mahgoub began work on debrisoquine metabolism at St Mary's Hospital Medical School in London. As was, and still is, often the case, members of the laboratory participated in the study. In May 1977, five colleagues took the medicine,

including Robert Smith, who was head of the laboratory. Within hours, Smith felt dizzy and disorientated, and he soon collapsed with severe hypotension. In most people, debrisoquine is relatively quickly broken down to inactive metabolites and eliminated through the urine. However, Smith's urine contained no metabolites of debrisoquine[3]. The study was extended to his family members, some of whom suffered similar outcomes. Subsequently, the investigations were extended to an unrelated population that consisted of medical students. This work ultimately led to the molecular cloning of the gene that is primarily responsible for the metabolism of debrisoquine, cytochrome P450 subfamily IID polypeptide 6 (CYP2D6), and to the characterization of polymorphisms that eliminate CYP2D6 activity.

DEBRISOQUINE
A drug (now superseded) that was used to treat high blood pressure and is metabolized by the enzyme CYP2D6.

PRO-DRUG
A biologically inactive compound that is changed in the body into a biologically active form, for example, by the action of one or more drug-metabolizing enzymes.
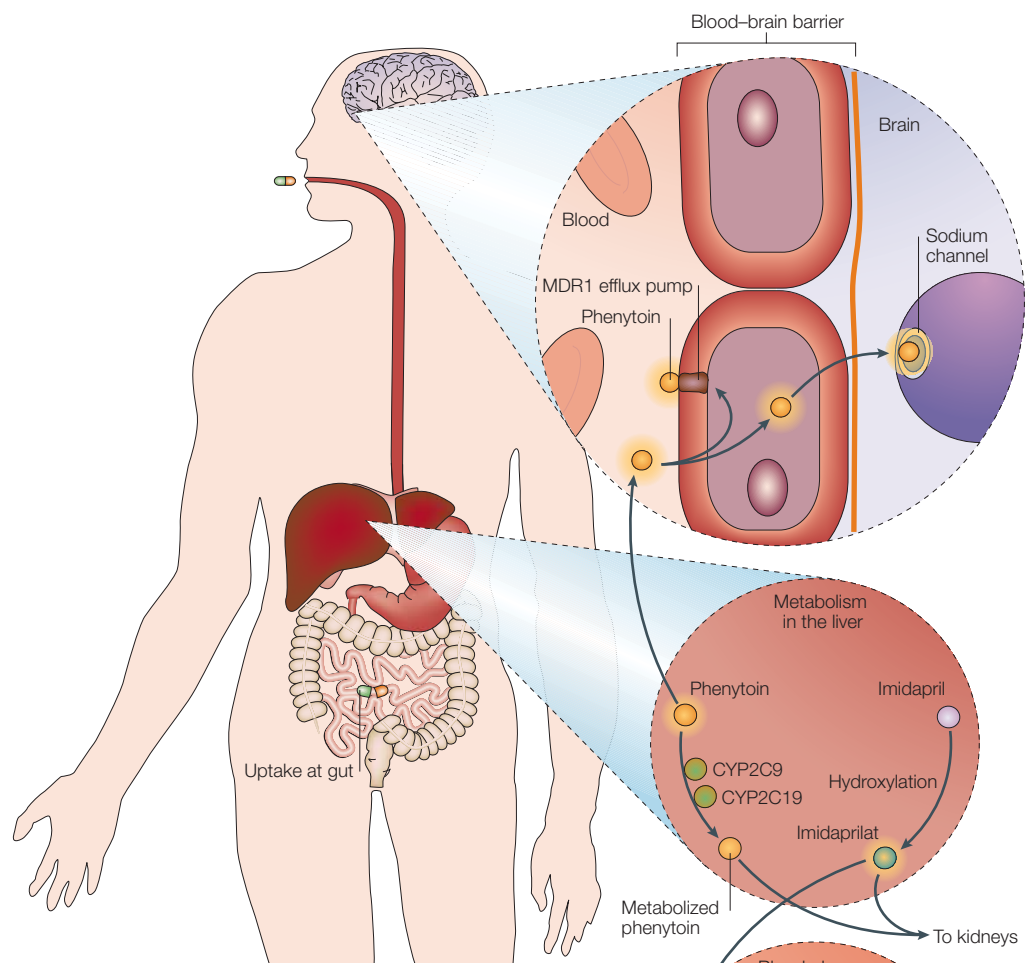
Figure 1 | **The path of phenytoin and imidapril after ingestion.**
This figure shows the paths that are taken by the anti-epileptic drug phenytoin and the angiotensin-converting enzyme (ACE) inhibitor imidapril in the human body. Phenytoin is absorbed into the bloodstream at the gut and circulated through the liver to the brain. It crosses the blood–brain barrier where it binds and inhibits its target, neuronal sodium channels. It is pumped back out across the blood–brain barrier into the bloodstream by multidrug resistance protein 1 (MDR1, also known as ABCB1) efflux pumps. Note that MDR1 efflux pumps are also active in the gut, where they promote drug excretion (not shown). At the liver, phenytoin is metabolized by the cytochrome P450 enzymes CYP2C9 and CYP2C19, and it is eliminated through the kidneys. Imidapril is a PRO-DRUG. After its absorption from the gut into the bloodstream it is hydroxylated in the liver to the active metabolite imidaprilat. Imidaprilat binds and inhibits ACE in the plasma. Imidaprilat is also eliminated through the kidneys.

Drug response: a simpler complex trait? As well as debrisoquine, there were many other early examples of variable responses to drugs owing to variable metabolism, such as the peripheral neuropathy that is experienced by some patients in response to the antimycobacterial drug isoniazid. These examples identified as CANDIDATE GENES for variable drug response those genes that encoded enzymes that metabolize drugs and their products. The relevant enzymes are now called drug-metabolizing enzymes (DMEs) and have evolved to neutralize XENOTOXINS and/or to control concentrations of signalling molecules in endogenous pathways[4]. The principal site of drug metabolism is in the liver, although many DMEs are also active in other tissues; for example, some DMEs form a component of the blood–brain barrier[5].

Although the best studied by far, the genes that encode DMEs are only one of several obvious groups of genes that might have variation that is relevant to drug response. Drug transport is another process that is of obvious importance. Drugs are actively moved between or out of body compartments by a range of specialized transporters. Finally, drugs must physically bind to their targets, such as receptors or enzymes, to modulate their behaviours, so genes that encode proteins that are involved in these processes might also be involved in variation in drug response.

The anti-epileptic drug phenytoin and the angiotensin-converting enzyme (ACE) inhibitor imidapril provide two good examples of the ways in which the body acts on, and is acted on by, drugs (FIG. 1). Both parent compounds are metabolized in the liver. In the case of phenytoin, the active drug is metabolized to forms with insignificant anti-epileptic activity, whereas the pro-drug imidapril is transformed into the active form imidaprilat, which then passes to the plasma where it binds to and inhibits ACE. The drug transporter ABCB1 at the blood–brain barrier is also likely to affect the amount of phenytoin that eventually reaches the brain-expressed sodium channels that are its target[6]. Other transporters might also influence phenytoin distribution[7].

These examples illustrate clearly defined groups of candidate genes that could harbour variation relevant to variable drug response: the first contains genes such as those that encode DMEs and drug transporters, which control the PHARMACOKINETIC properties (including DISPOSITION) of the drug; and the second group contains genes that encode drug targets (plus elements of the associated pathways) that influence drug PHARMACODYNAMICS.

The second category includes not only the specific target of the drug, but also the broader pathway in which the target acts. To this we also add molecules that are similar to the target, which the drug might also modulate, such as the effect that many non-cardiovascular drugs can have on potassium channels and thereby on QT-INTERVAL duration[8].

One notable feature of pharmacogenetics is how often the obvious candidate genes carry variants that seem to influence drug response. Most DME-encoding genes have polymorphisms that have been shown to influence enzymatic activity[9], either through *in vitro* studies or through direct measurement of drug or metabolite levels

in vivo. The effects of most of these variants on responses to relevant drugs have not yet been characterized. The importance of metabolism in adverse reactions in particular was highlighted recently by Phillips and colleagues[10], who studied the metabolism of 27 drugs that are relatively prone to produce adverse reactions. Of these, 59% are metabolized by at least one enzyme that is known to have low-activity forms, whereas less than 22% of randomly selected drugs satisfy this criterion.

Drug transporters also show considerable genetic variation, including many potentially functional polymorphisms[11], although to our knowledge only one variant has been associated with drug response in more than one study, in the ATP-binding-cassette sub-family B member 1 (*ABCB1*) gene[7,12]. The *ABCB1* gene encodes a transmembrane efflux pump that is known to transport a broad range of drugs and is expressed in the gut, liver and other tissues (FIG. 1). Variation in genes that encode drug targets has also been implicated in several drug responses. The β-2 adrenergic receptor (*ADRB2*) has polymorphisms that have been associated with response to both beta-agonists in the treatment of asthma[13–15] and beta-blockers as part of the management of congestive heart failure[16]. Similarly, the arachidonate 5-lipoxygenase (*ALOX5*) gene encodes an enzyme that is involved in the production of LEUKOTRIENES, which promotes bronchial constriction. A promoter-region VARIABLE NUMBER OF TANDEM REPEATS (VNTR) that affects *ALOX5* expression has been associated with the efficacy of both inhibitors of *ALOX5* and antagonists of its receptor[17].

These examples must be considered in light of the fact that until recently, pharmacogenetic studies were limited in scope and considered only a small subset of the variation in a few genes. Given these constraints, the above examples show considerable scope for the identification of gene variants that influence drug response through the systematic study of candidate genes. This is not to say that variants that influence drug response will not be found in unexpected genes. However, it does seem reasonable to conclude that some of the important pharmacogenetic variation resides in a relatively small number of candidate genes for any given drug. It would be much harder to make such a case for disease predisposition, which indicates that pharmacogenetic research might progress more rapidly during the candidate-gene phase of study, before the tools become available for systematic genome-wide scans (see below).

Where are we now? To assess the state of pharmacogenetics research we have identified from the literature a set of variants that have been significantly associated with drug response in at least two studies (TABLE 1a,b; for further information, see ONLINE TABLE 1). This includes examples of direct replication, in which the same drug response was studied in the first and at least one subsequent study (26; 62%) and examples in which a similar drug response (that is, the same class of drug) was studied (8; 19%). Examples are also included of functional replication; that is, where the polymorphism was associated with a different drug response but there was apparently a common underlying physiological cause for the

Table 1a | **Pharmacogenetic variants that have been significantly associated with drug response in at least two studies**

Gene	Name	Allele‡	Associated phenotype
Drug target/pathway protein			
ACE	Angiotensin-I converting enzyme	Ins/del	Del/Del associated with decreased proteinurea in response to ACE inhibitor treatment for renal disease
ADRB1	β-1 adrenergic receptor	Arg389Gly	Arg/Arg homozygotes have greater response to β-adrenergic receptor antagonists
ADRB2	β-2 adrenergic receptor	Arg16Gly	Gly allele associated with decreased response to albuterol and salbutamol
AGT	Angiotensinogen	Met235Thr	Thr allele associated with greater reduction of blood pressure and greater decrease in left ventricular mass with anti-hypertensive treatment
AGTR1	Angiotensin-II receptor type 1	A1166C	C allele associated with greater response to angiotensin-II receptor antagonists
ALOX5	Arachidonate 5-lipoxygenase	Promoter VNTR (wt = 5; non-wt = 3,4 or 6)	Non-wt homozygotes have decreased response to 5-lipoxygenase inhibitor and leukotriene receptor antagonist
BDKRB2	Bradykinin receptor B2	C-58T	T allele associated with ACE inhibitor-related cough
CETP	Cholesteryl ester transfer protein	TaqIB polymorphism B1/B2	B1/B1 associated with increased response to pravastatin and atorvastatin
DRD2	Dopamine D2 receptor gene	3′ UTR Taq1A A1/A2	A1 associated with greater positive response to the antipsychotics haloperidol and nemonapride
DRD3	Dopamine D3 receptor gene	Ser9Gly	Gly associated with response to clozapine; Ser/Ser with non-response
DRD4	Dopamine D4 receptor gene	Exon 3 VNTR (twofold to sevenfold repeat)	7 allele significantly less frequent than 4 allele in patients responding to typical neuroleptics (compared with patients responding to clozapine); 4/4 homozygotes have a higher rate of good response to neuroleptics.
GNB3	Guanine nucleotide-binding protein (G protein), β-polypeptide 3	Exon 10 C825T	CC associated with response to antidepressants
GRIN2B	Glutamate receptor, ionotropic, N-methyl D-aspartate 2B	C2664T	CC associated with higher mean clozapine dosage for schizophrenia
HTR2A	5-hydroxytryptamine (serotonin) receptor 2A	T102C	CC or C associated with susceptibility to tardive dyskinesia in response to antipsychotic drugs
HTR2A	5-hydroxytryptamine (serotonin) receptor 2A	His452Tyr	Tyr associated with reduced response to the antipsychotic clozapine
LIPC	Hepatic lipase	C-514T	CC genotype associated with increased response to statins
MTHFR	5,10-methylene-tetrahyrdrofolate reductase	C677T	TT patients have increased toxicity for methotrexate therapy
SLC6A3	Dopamine transporter	3′ VNTR 10-repeat allele	10/10 genotype associated with poor response to methylphenidate for treatment of attention deficit/hyperactivity disorder
SLC6A4	Serotonin transporter (5-HTT)	Promoter ins/del (long/short)	long/long better response to fluoxetine or paroxetine than short/short
TPH1	Tryptophan hydroxylase1	A218C	The A218C polymorphism has been associated with response to anti-depressants
TYMS	Thymidylate synthase	TSER*2/*3	*3 reduced response to 5-flurouracil treatment compared with *2; *3/*3 requires a higher dose
Drug transporter			
ABCB1	MDR1, P-glycoprotein 1	C3435T	TT patients are less likely to have drug-resistance epilepsy than CC and have increased immune recovery after the initiation of antiretroviral treatment

‡As described in online table 1, summaries of the types of mutation that are responsible for the alleles are as follows: 16 missense, 8 in the promoter region, 4 synonymous, 4 intronic, 3 in the 3′ untranslated region, 2 gene deletions, 2 nonsense, 1 exonic VNTR, 1 splice variant and one HLA allelic variant. Amino acids are represented by their standard three letter abbreviations, whereas nucleotides are represented by their standard single letter abbreviations. Del, deletion; Ins, insertion; UTR, untranslated region; VNTR, variable number of tandem repeats; wt, wild type.

associations (8; 19%). For example, the exon 26 polymorphism of the *ABCB1* gene has been associated with the efficacy of anti-epileptic drugs and response to anti-retroviral drugs. In both cases, low- and high-activity *ABCB1* genotypes apparently influence the bioavailability of the drugs, although this is presumably owing to ABCB1 activity in different tissues. The associations were compiled from review papers and from PubMed literature searches using the following keywords

and phrases: pharmacogenetics OR pharmacogenomics, association study AND drug response, polymorphism AND drug response.

This list omits many polymorphisms and probably includes some false positives (compare with REFS 18,19). Many of the associations that are reported in TABLE 1a,b, however, have been replicated more than twice and many have well-studied functional effects that are consistent with the observed drug response (for example,

Table 1b | **Pharmacogenetic variants that have been significantly associated with drug response in at least two studies**

Gene	Name	Allele‡	Associated phenotype
Metabolism			
BCHE	Butyrylcholinesterase	Several mutations including Asp70Gly (dibucaine or atypical variant) and Ala539Thr (K allele)	Variants associated with adverse effects in response to succinylcholine
COMT	Catechol O-methyltransferase	Val108/135Met	Met associated with higher daily neuroleptic dosage and poor response
CYP2C19	Cytochrome p450 2C19	*2,*3	Extensive metabolizers have decreased response to omeprazole for *Helicobacter pylori* infection
CYP2C9	Cytochrome p450 2C9	*2,*3	Non-wt alleles associated with reduced warfarin daily dose requirement
CYP2D6	Cytochrome p450 2D6	Many inactive alleles, such as *3, *4 and *5	Non-wt alleles associated with susceptibility to tardive dyskinesia in response to antipsychotics
DPYD	Dihydropyrimidine dehydrogenase (DPD)	DPYD*2A (IVS14+1G>A)	*2A associated with severe toxicity and fatal outcomes for 5-fluorouracil treatment
GSTM1	Glutathione S-transferase M1	GSTM1-null	Null carriers have increased survival time and a progression-free interval following paclitaxel and cisplatinum treatment for ovarian cancer; decreased risk of relapse for cytotoxic therapy for leukaemia
GSTM3	Glutathione S-transferase M3 (brain)	GSTM3*A/GSTM3*B	GSTM1*0/GSTM3*A haplotype less likely to show a beneficial response to D-penicillaminein rheumatoid arthritis; *3A has increased risk of cisplatin ototoxicity
GSTP1	Glutathione S-transferase-π	Ile105Val	Val associated with increased survival for 5-fluorouracil and oxaloplatin therapy for colorectal cancer, and following therapy for multiple myeloma
GSTT1	Glutathione S-transferase-η1	GSTT1-null (homozygote frequent; deletion	GSTT1 associated with susceptibility to tacrine hepatotoxicity and troglitazone hepatotoxicity in combination with GSTM1-null allele
NAT2	N-acetyltransferase 2	Slow acetylator alleles include NAT2*5B, NAT2*6A, NAT2*7A or B, NAT2*14A or B	Slow-acetylator status of NAT2 increased risk of antituberculosis drug-induced hepatotoxicity
TPMT	Thiopurine methyltransferase	TPMT*2, TPMT*3A, TPMT*3C	Homozygotes for non-wt alleles at high risk of severe haematopoietic toxicity after thiopurine treatment; heterozygotes intermediate risk of dose-limiting toxicity
UGT1A1	UDP-glucurono-syltransferase 1A1	UGT1A1*28	*28 associated with increased chance of developing diarrhoea and leukopaenia during irinotecan therapy
Other			
ADD1	Adducin 1 (α)	Gly460Trp	Trp associated with increased response to diuretics in hypertensives (hydrochlorothiazide)
APOE	Apolipoprotein E	E4	E4 allele associated with lesser response to statins
FCGR3A	FcγRIIIa receptor protein	Phe158Val	Val/Val patients greater response to rituximab for non-Hodgkin lymphoma and higher complete remission rate with immunosupressive medication for idiopathic thrombocytopenic purpura
HLA-B	Major histocompatibility complex class I-B	HLA-B*5701	Associated with hypersensitivity to abacavir
IL10	Interleukin 10	A-1082G	GG has better response to prednisone in leukaemia patients and sustained response in antiviral therapy for chronic hepatitis C infection (as part of haplotype (108 base pairs)–(-2575T)–(-2763C)–(-1082A)–(-819T)–(-592A))
TNF	Tumour-necrosis factor (TNF superfamily, member 2)	G-308A	A associated with good response to immunosuppressive therapy in patients with aplastic anaemia and carbamazepine hypersensitivity
XRCC1	DNA-repair protein XRCC1	Arg399Gln	Gln associated with resistance to 5-fluoruracil and oxaliplatin chemotherapy, Gln/Gln individuals less likely to develop therapy related acute myeloblastic leukaemia

‡As described in online table 1, summaries of the types of mutation that are responsible for the alleles are as follows: 16 missense, 8 in the promoter region, 4 synonymous, 4 intronic, 3 in the 3′ untranslated region, 2 gene deletions, 2 nonsense, 1 exonic VNTR, 1 splice variant and one HLA allelic variant. Amino acids are represented by their standard three letter abbreviations, whereas nucleotides are represented by their standard single letter abbreviations. wt, wild type

thiopurine S-methyltransferase (*TPMT*), *ADRB2* and *ABCB1*). Most of the polymorphisms are either in the drug target or in a protein that is in the pathway in which the target acts (55%), whereas nearly one-third are in DMEs (TABLE 1a,b). These examples confirm an important role for the two categories of candidate genes listed above: genes that influence pharmacokinetics, for example, DMEs and transporters; and genes that influence pharmacodynamics, for example, targets and other elements of the pathway in which the target acts. However, there is an obvious ascertainment bias as studies have concentrated mainly on genes that are expected to influence drug response. So, for example, drug transporters have only one example (in the *ABCB1* gene). It is only recently that the role of drug transporters in multidrug resistance in anti-tumour therapy has sparked interest in transporter pharmacogenetics[20]. The entries in TABLE 1a,b also highlight another category of candidate

genes for drug response: genes that are involved in the underlying disease condition or intermediate phenotype. For example, *APOE*4* has been associated with response to STATINS, yet statins inhibit an enzyme that is involved in cholesterol synthesis to which *APOE* has no direct connection (although it is involved in lipid transport). In some cases, the variant in question might both influence the disease and interact directly with the drug. For example, it is thought that the *ALOX5* regulatory variant determines whether leukotrienes have a role in asthma in a particular patient, and only if they do will inhibitors of *ALOX5* be effective[17].

The future of pharmacogenetics

Although it clearly illustrates the scope for genetic influence on drug response, our survey also shows that most pharmacogenetic studies so far have fallen well short of the ideal approach to studying the genetics of drug responses and indeed of what is now possible. Moreover, the studies seem small and unplanned in comparison with many studies of disease predisposition.

Sample sizes. To detect even relatively strong associations between genetic variants at a specific locus and variation in drug response, many cases and controls are needed. For example, to detect the effect of a gene variant that explains 5% of the total phenotypic variation in a quantitative response to a drug by typing 100 independent single nucleotide polymorphisms (SNPs) would require 500 patients to provide an 80% chance of detection, assuming an experiment-wide false-positive rate (type I error threshold) of 5% (using the power calculation of Sham and colleagues[21]). Most of the sample sizes for the studies that are reported in TABLE 1a,b are much smaller than required for this relatively modest aim. All but 3 of the 84 first and second studies have fewer than 500 patients, whereas 35 have fewer than 100 subjects. This means that these studies are generally underpowered, even for detecting variants of relatively large effect (compare with REF. 18).

Moreover, complications such as STRATIFICATION[22,23] have been ignored in the pharmacogenetic literature. Stratification occurs when a patient population is genetically structured, and can create spurious associations between gene variants and drug responses. As far as we are aware, only one study in pharmacogenetics has used genetic methods to rule out stratification as an explanation of the association[6].

Of equal importance is that all of the studies so far have been CANDIDATE-POLYMPORPHISM STUDIES; that is, they have been focused on specific polymorphisms without systematic representation of the variation in the genes that are under study. Furthermore, none of the studies systematically tested for interactions between polymorphisms in different genes in determining drug response. These are both important omissions. Our knowledge of functional variation is incomplete, even in well-studied genes. Furthermore, although pharmacogenetics might be more tractable than disease genetics, the behaviour of most drugs will be influenced by a wide range of gene

products (DMEs, transporters, targets and others), and in many cases the importance of polymorphisms in one of the relevant genes might depend on polymorphisms in other genes. As a simple example, *CYP1A2* and *N*-acetyltransferase 2 (*NAT2*) act in different stages in the pathway that metabolizes compounds in burnt meat, and variants might interact to influence the risk of colorectal cancer[24].

One promising development is the availability of large patient cohorts with detailed clinical records. For example, the deCode genetics project, which has enrolled many of the 280,000 inhabitants of Iceland in a database that is available for genetic research, has found an association between schizophrenia and variation in neuregulin 1 (REF. 25). This finding has implications for pharmacogenetics as well as disease predisposition, as neuregulin HYPOMORPH mice have fewer *N*-methyl-d-aspartate (NMDA) receptors than wild-type mice and their behavioural abnormalities are partially reversed by the atypical antipyschotic drug clozapine[25].

Population-genetic study design. Two approaches that could be used to systematically represent variation in candidate genes (or ultimately the whole genome) have been termed 'MAP BASED' and 'SEQUENCE BASED' by Peltonen and McKusick[26]. In the former, genomic patterns of LINKAGE DISEQUILIBRIUM (LD) are used to select a set of markers that are statistically associated with many other variants in the genome. This 'map' of markers, often called tagging SNPs (tSNPs; FIG. 2), is then typed to economically represent genomic variation. In the sequence-based strategy, SNPs that are potentially functional are identified and assayed directly; for example, SNPs that are in and near exons. Botstein and Risch[27] have recently emphasized the importance of a sequence-based strategy in the study of common disease, arguing that the example of Mendelian disease indicates that causal variants are most likely to reside in and near exons. They further argue that the variants are likely to be rare and therefore hard to represent with tSNPs.

We agree that sequence-based strategies are an important complement to LD mapping, especially given that it remains unclear how well tSNPs will be able to represent variants with lower minor-allele frequency. Also, sequence-based strategies can represent an important shortcut when a set of possibly functional variants is already known. But this approach is likely to miss at least some regulatory variation and so should not replace map-based strategies, even for genes that have been subjected to extensive re-sequencing of exons. Our survey of pharmacogenetic studies, for example, identified eight promoter-region polymorphisms that were associated with drug response, as, for example, in the case of *ALOX5* discussed above. Coupled with evidence that drug–drug interactions influence variable responses to medicines, these polymorphisms indicate that regulatory variants have a far more important role in variable drug response than they do in Mendelian diseases[27]. Moreover, most of the variants that are listed in TABLE 1a,b are common and, presumably, would be easily represented by tagging strategies.

STATINS
A class of cholesterol-lowering drugs that inhibit a key enzyme in the synthesis of cholesterol.

STRATIFICATION
If a genetic-association study is done in a population that is genetically structured, associations might be observed between polymorphisms that reflect allele-frequency differences between the unknown subgroups in the test populations. There are methods available to both detect and correct for such spurious association.

CANDIDATE-POLYMORPHISM STUDY
A previously known polymorphism, usually thought to be functional, which is tested for its effect on a drug response. This should not be confused with a candidate-gene study, which would seek to exhaustively represent variation in the gene.

HYPOMORPH
Low activity of forms of a gene.

MAP BASED
An approach to genetic-association studies that is focused on putatively functional SNPs, for example, identified by re-sequencing exons and other functional regions in relatively large samples, or directly in patients. This approach is also sometimes called direct.

SEQUENCE BASED
An approach to genetic-association studies that is focused on a set of genetic markers, often now called tagging SNPs, which are statistically associated with whichever variants influence the phenotype.

LINKAGE DISEQUILIBRIUM
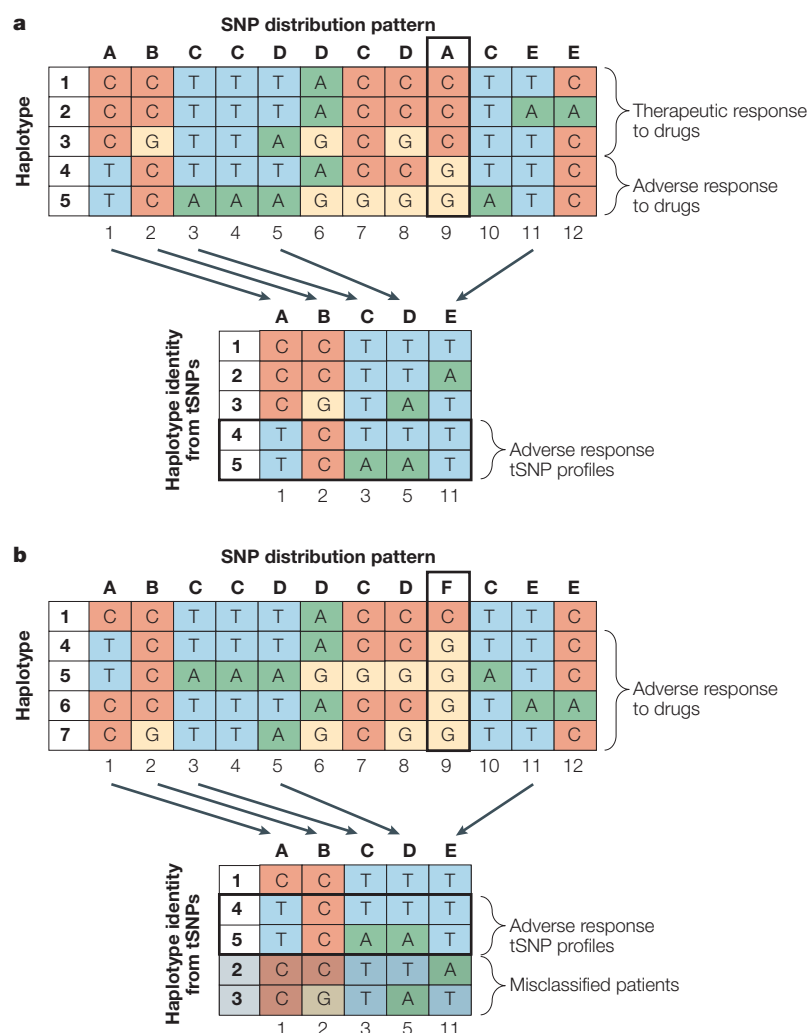The non-random association of alleles at different polymorphic sites in the gene.

Figure 2 | **Illustration of tagging SNPs. a** | The diagram shows five haplotypes. Twelve single nucleotide polymorphisms (SNPs) are localized in order along the chromosome. The letters on the top indicate groups of SNPs that have perfect pairwise linkage disequilibrium (LD) with one another, and the numbers on the bottom indicate each of the 12 SNPs. SNP 9 is the causal variant, which in this simple example determines drug response: allele C results in a therapeutic response, whereas allele G results in an adverse reaction. In this example, the selection of just one SNP from each of the groups A–E would be sufficient to fully represent all of the haplotype diversity. Each haplotype can be identified by just five tagging SNPs (tSNPs), and the causal variant would be tagged even if it were not itself typed (in fact, multi-marker approaches to tSNP selection would reduce the set of tags to fewer than five, but this is ignored for simplicity). So, tSNP profiles that are highlighted predict an adverse reaction to the medicine. Normally, LD patterns are not so clear-cut and statistical methods are required to select appropriate sets of tSNPs. **b** | The diagram depicts the same 12 SNPs, but with different associations among them, as might happen in a different population group. Because patterns of LD are different, some patients would be misclassified if the same five tSNPs were used and interpreted in the same way; that is, using the same SNP profiles as defined in population A, haplotype profiles 1, 2 and 3 are predicted to have allele C at the causal SNP 9 (a therapeutic response), whereas haplotype profiles 4 and 5 are predicted to have an adverse response. However, because the pattern of association has changed, the new haplotypes 6 and 7 are misclassified as haplotype patterns 6 and 7 in population B.

The approach that we advocate is illustrated with a simple example. Phenytoin (FIG. 1) is metabolized by cytochrome P450 system enzymes, is a possible substrate for the drug transporters ABCB1, multidrug resistance-associated protein 1 (MRP1) and MRP2, and acts on voltage-gated sodium channels (the brain-expressed genes that encode the channel are *SCN1A*, *SCN2A*,

SCN3A, *SCN8A*, *SCN1B* and *SCN2B*). These 'front-line' candidate genes for variable response to phenytoin represent a total of ~1 Mb of discontinuous genomic sequence. Although the optimal approach for identifying and using tSNPs remains controversial, recent work indicates that fewer tSNPs are required to represent common variation than was previously assumed[28–31]. Taking the required number of tSNPs as 1 in every 10 kb[31], a full study of front-line candidate genes for phenytoin would require the typing of only ~100 tSNPs, once the haplotype structure had been determined. Typing this number of tSNPs in patients with different response profiles to phenytoin would allow the identification of any common polymorphism — regulatory or structural — that influences response.

The obvious importance of drug targets and the associated pathways that are represented in TABLE 1a,b makes it clear that it is an urgent priority for pharmacogenetics research to identify appropriate tSNPs for not only DMEs and transporters, but also for the important targets and associated pathways. Biochemical considerations and the empirical evidence (for example, TABLE 1a,b) indicate that genes that are relevant to disease will also influence drug response. In general, disease genes will probably be harder to itemize than the drug-specific candidate genes that are listed above and will eventually motivate genome-wide scans in the study of some drug responses.

We emphasize that in the study of candidate genes for drug response, it will be important to consider groups of potentially interacting genes as a set, such as those that act in common pathways, to identify interactions between polymorphisms in different genes. For example, genes that encode enzymes that act at different points in the metabolism of a drug should be considered collectively, as should all of the genes that encode a receptor complex (for example, the set of genes that encode the subunits of ion channels or the set of genes that encode receptors that form multimers). On the basis of the number of variants that have already been identified using candidate-polymorphism studies, it seems likely that the next few years will see a notable increase in our knowledge of how genetic variation influences drug response.

Getting to the clinic. Most pharmacogenetic studies are retrospective; that is, the genetic work has been carried out after the responses have already been observed. Few gene variants have been studied prospectively to assess how knowledge of the relevant genotype might improve clinical outcomes, for example, by adjusting doses or selecting the most appropriate drug as a function of the genotype. There might be some examples in which the effects of certain variants are so extreme that knowledge of them could inform drug choice without explicit prospective evaluation. A theoretical example of this would be *CYP2D6*-nulls and codeine efficacy. As codeine is a pro-drug that must be transformed into morphine to have an analgesic effect, knowledge of *CYP2D6* genotype alone could identify patients who would fail to respond to codeine.

In general, however, translating pharmacogenetic research into improved therapies will require a drastic expansion of prospective studies of how variants influence responses to drugs. To assess the usefulness of a pharmacogenetic test in general, it is important to have a reliable estimate of positive and negative predictive values (PPV and NPV, respectively[32]). Following Holtzman, PPV for drug response X is defined as the proportion of people with a positive test result who have response X, and NPV is the proportion of people with a negative test who do not have drug response X. Accurate estimates of PPV and NPV are rarely available in retrospective-study designs[32].

More fundamentally, the required thresholds for negative and positive predictive value will clearly depend on the precise clinical situation, which will usually need explicit evaluation. For example, if the condition being predicted is a severe adverse drug reaction (ADR), and there are alternative, but less efficacious, medicines available, a test with a high NPV but low PPV might be clinically useful. Also, specific clinical interventions that are indicated by identified genetic associations might not be straightforward, for example, such interventions might be complicated by new drug interactions. For these reasons, it is likely that the clinical application of pharmacogenetic research will considerably lag behind the first identified associations between genes and drug responses. This also explains why so few tests are in use today (for an exception, see BOX 1).

Follow-up of associations. The association of tSNPs with a drug response is only the first step. As we have discussed above, determining whether and how to use a pharmacogenetic association clinically will generally require prospective evaluation. This evaluation will be greatly facilitated by the identification of the causal variants that underlie the genotype–phenotype correlation.

So, after associations are first detected, it is of high priority to assess patterns of LD that surround associated variants to identify the ASSOCIATED INTERVAL and the set of candidate causal variants within it[33]. These polymorphisms will then need to be prioritized for functional evaluation.

There are two fundamental reasons why causal variants should be identified to make clinical use of associations. First, clinical complications can be better predicted and accommodated if the biological cause of the association is understood, as in, for example, the receptor regulation model that has been used to explain differences in short- and long-term response to beta-agonists as a function of the *ADRB2* genotype[34].

Second, we noted that the decision whether to use a pharmacogenetic test will depend on its positive and negative predictive value[32]. If a marker, or set of markers, is used instead of the causal variant as a diagnostic indicator, the positive and negative predictive value will depend on the degree of association between the markers and the underlying causal variants. As these associations are known to vary among population groups, it is unlikely that tests could be developed that would have stable predictive value without identification of the underlying causal variants. So, unless the test was useful over a broad range of predictive values, this would seem to prohibit widespread use of the test. The consequences of applying such a test in a different population from that which it was developed in could be severe (see, for example, FIG. 2). For this reason, it is unlikely that genome-wide SNP profiles will provide the basis for the genetic personalization of medicines in the near future. Rather, these approaches will provide pointers to the underlying causal variants that will need to be identified and exploited for clinical application.

Even if the pharmacogenetic results are not used directly in the clinic, but are used to improve the design of new medicines, it will still be crucial to identify causal variants. Overcoming the variable or limited efficacy of medicines with new compounds will be made much easier if the biological cause of the variable reaction is understood in as much detail as possible.

Pharmacogenetics in society

Economic and ethical considerations will influence both the directions that are taken in pharmacogenetic research and its clinical application. Overlapping but distinguishable concerns will arise for health-care providers, patients and the pharmaceutical industry. Noah[35] recently argued that the clinical use of pharmacogenetics runs counter to trends among payers of health care to adopt 'economizing mechanisms, such as restricted formularies and therapeutic substitution'. To the extent that gene variants could offer significant diagnostic value about how patients respond to medicine, it seems likely that health-care providers will benefit economically by paying for pharmacogenetic diagnostics, as well as providing better clinical care for their patients both from savings on expensive medicines that do not work in many patients and by savings on costs that are associated with ADRs.

Box 1 | **TPMT pharmacogenetics and clinical practice**

The thiopurine S-methyltransferase (*TPMT*) gene is one of the few examples in which pharmacogenetic testing has been effectively intergrated into clinical practice. In this case, testing for TPMT genotypes is used to modify doses of thiopurines such as 6-mercaptopurine and azathioprine that are used to treat acute lymphoblastic leukaemia and inflammatory bowel disease[47,48]. The *TPMT* polymorphism is of great clinical importance because individuals with deficient or intermediate TPMT activity risk toxicity, including fatal MYELOSUPPRESSION, at standard thiopurine doses. Testing for low-activity variants for TPMT has been in clinical practice for more than 10 years in the United States and has been shown to be cost-effective in certain health-care settings[49]. Even in this case, which might be the best studied so far, there are important gaps in knowledge that interfere with the broad application of pharmacogenetic testing. For example, present practice relies on a set of functional alleles described primarily in Caucasian populations. These variants have different frequencies in different population groups, and some populations are likely to carry functional alleles that are absent or rare in Caucasians[50]. So, in future, if the clinical application of pharmacogenetic testing is to become widespread, identification of the causal pharmacogenetic variants that underlie the response and explicit evaluation of the distribution of functional alleles across different populations groups is likely to be important[51]. Furthermore, there is great variability in TPMT activity in both the homozygous wild type and heterozygous groups, so it is likely that other genetic and environmental factors have a role in determining TPMT activity.

Box 2 | Adverse reactions to felbamate

The anti-epileptic drug felbamate is a good example of a drug the use of which could be made cost-effective if we were able to identify patients that were at risk of having an adverse response. Although many new anti-epileptic drugs tend to be expensive, they can still represent good value for society if patients are rendered seizure free, avoid repeated hospitalization and become net financial contributors to society. Felbamate is structurally unrelated to other anti-epileptic drugs (AEDs), but is related to the anxiolytic drug meprobamate. The discovery that felbamate is a potent AED in standard models and was able to control seizures in trials in some patients with otherwise refractory epilepsy led to its eventual licensing in 1993 in the United States. Within a year, probably more than 100,000 patients had been exposed to felbamate. Clinical trials, which were necessarily undertaken in selected and comparatively small populations (<2,000 individuals), showed no evidence of clinically significant blood or liver disorder. However, in the first year after licensing, 34 cases of aplastic anaemia were reported (with 13 fatalities) and 23 cases of hepatic failure (with 5 deaths)[52]. Although many patients had co-morbidities and were on other AEDs, these serious events led to severely restricted usage or withdrawal of its licence, which represented a huge loss of investment for the manufacturers, and also of time and resources for administrative and clinical bodies, and, most importantly, the loss of an important anti-epileptic drug for most patients with refractory epilepsy for whom it would probably have been safe to use. Certain individuals seemed more at risk: those with previous sensitivity to AEDs, immune disorders and reduced blood-cell counts in particular[52].

Detailed studies of the metabolism of felbamate have been undertaken to explore these potentially fatal idiosyncratic reactions. Felbamate is metabolized in the liver to highly reactive toxic intermediates (for example, atropaldehyde and 2-phenylpropenal)[53,54], which are usually rapidly detoxified (for example, by CONJUGATION with glutathione). Overproduction of these (or other) reactive metabolites, or, as indicated by more recent research, inadequate conjugation, might lead to an excess of such metabolites. This, in turn, might be the cause of idiosyncratic reactions by a range of mechanisms, such as immune responses. The genes that encode enzymes in metabolic pathways are, in general, variable among people, as discussed in the main text. So, pharmacogenetic factors might contribute to the idiosyncratic reactions that have curtailed the use of felbamate. The ability to reliably predict which patients might be at risk of such reactions to felbamate might allow greater confidence in its use, and again make available a potentially valuable AED for patients with epilepsy that can itself be life-threatening. It might now be too late to carry out population-pharmacogenetic studies for felbamate, but predictive tests for felbamate-naive individuals in whom felbamate might usefully be considered are in development and the general principles might productively be applied in the future development of other AEDs.

The economic calculations for industry are more complicated, because of the concern about market segmentation[36]. But there are potential benefits. Some drugs that work well generally are rejected, withdrawn or limited in use because of rare but serious ADRs, for example the anti-epileptic felbamate, the atypical antipsychotic clozapine and most drug withdrawals owing to QT-interval-associated arrhythmias (for example, terfenadine). If pharmacogenetic predictors of adverse events could prevent the exposure of genetically vulnerable patients and so preserve even a single drug, the costs of any large-scale research effort in pharmacogenetics could be fully recovered. Clinical care could also be improved by allowing the treatment of most patients that do not experience adverse events (BOX 2). In this regard, one of the highest priorities is to understand the factors that influence predisposition to QT-interval prolongation (which is considered to be a surrogate for risk of life-threatening ventricular arrhythmias), which has been responsible for more drug withdrawals than any other category of adverse event in recent times, with eight non-cardiac drugs withdrawn between June 1990 and March 2001 (REF. 8).

Overall, economic considerations will not impede pharmacogenetics in general, but could drive a differentiation between the research focus in the industrial and academic sectors. This makes clear the need to establish appropriate structures that would facilitate the transfer of research projects and information between the industrial and academic research sectors.

In considering ethical implications, it is useful to distinguish pharmacogenetic research that is applied during the development of new drugs and that which is applied to existing licensed medicines in routine clinical use. One motivation for pharmacogenetic testing during drug development is that the stratification of patients by genotype might allow the identification of responses that would have been missed in an unselected cohort. This could allow efficacy for subgroups to be shown in drugs that might not have been considered effective in general populations, which could potentially improve the success rate of compounds. Because of the huge costs per patient of clinical trials, it has also been suggested that such streamlining might reduce the average cost of developing new compounds[37]. Such drugs would then be indicated for a subset of patients, as HERCEPTIN is for the subpopulation of breast cancer patients who overexpress *HER2*, and Gleevec is for those patients with chronic myeloid leukaemia that results from the PHILADELPHIA CHROMOSOME.

However, a concern with this use of pharmacogenetics is the possibility of focusing on 'easier to treat' subsets of the population and excluding from trials those with unfavourable, or simply unusual, genetic constitutions, even though a proportion of these might actually respond to the drug or to a modest alteration of it. This possibility is of particular concern given the extensive geographic variation in genes that are relevant to drug response[38] (for further information, see ONLINE TABLE 1).

A practical and ethical concern is the transferability of diagnostic tests across ethnic groups, particularly in the prediction of ADRs in which mistakes might have severe consequences. There are two distinct reasons

why diagnostic tests might not transfer well across ethnic groups. One is that the underlying genetic causes of the response might be different in the different groups. Most polymorphisms that have been associated with drug response show significant differences in allele frequency across populations and these differences are often substantial (ONLINE TABLE 1). So, in considering the effect of a given polymorphism, the relevant genetic background that might influence the effect of the polymorphism could differ considerably among populations. The second reason is the variable pattern of LD across populations when markers, instead of causal variants, are used diagnostically. It is well known that patterns of LD vary sharply between populations. A marker that has been associated with a phenotype in a given population, but that is not itself causal, is likely to have less or even no diagnostic value in other ethnic groups (FIG. 2). The importance of tracking down causal variants therefore becomes apparent. Both of these issues would need to be addressed in the application of pharmacogenetics on a global scale.

Finally, the widespread adoption of pharmacogenetic testing would raise serious issues about privacy. Variants that predict drug response might also influence disease predisposition. At least 23 of the variants that are listed in TABLE 1a,b have been associated with common diseases (in three or more association studies for each[18]). In many instances, they are associated with a common disease other than that which is being treated. For example, the apolipoprotein E4 allele (*APOE4*) is associated with a lesser response to statin treatment for the lowering of cholesterol[39,40] but also with an increased risk of Alzheimer disease[41,42].

Conclusions

Much of the enthusiasm surrounding recent advances in genomics has focused on its application to understanding common diseases. This is justified because understanding the causes of disease can point to new directions for drug development and for environmental interventions.

Pharmacogenetics, however, could offer more immediate clinical returns. This is in part because many drug responses seem to be genetically and physiologically simpler traits than are common diseases. From an evolutionary perspective this is not surprising. Gene variants that have an important effect on disease predisposition would be selected against in most environments, whereas there has been no such selection on how variants affect drug response. So, without selection working to reduce their frequency, there are even examples of common pharmacogenetic variants that have essentially Mendelian effects on drug response, such as the effect of *CYP2D6*-null alleles on codeine efficacy.

The second reason that pharmacogenetics might move more quickly than the genetics of common disease is that an association between genotype and drug response might be of direct diagnostic use; for example, in using genetic predictors to avoid rare ADRs[43] or to select which of several alternative drugs has the highest

efficacy[44]. Identification of a new variant that predisposes to disease might indicate a new therapeutic target or a new pathway with numerous candidate targets. However, it takes considerable time to develop new medicines to hit these targets and so finding them is less immediately useful than identifying pharmacogenetic variants, which through diagnostic testing can rapidly increase the efficacy of existing therapies[44–46].

It should be noted, however, that pharmacogenetics will also have indirect benefits for future drug development by helping to subgroup diseases genetically and providing pointers towards the genetic and physiological causes of variable and adverse reactions.

Realization of the health-care benefits of pharmacogenetics will require a close collaboration between clinicians and scientists. As we have argued, pharmacogenetic associations cannot guide therapy without explicit clinical evaluation. Similarly, clinical assessment must be fully informed by what is known about the pharmacogenetic variants that are under study and how they influence drug response. To achieve sufficient interaction between the research and clinical arms, environments need to be established that encompass basic and clinical pharmacogenetics, which house appropriate research tools and expertise, and involve real patients. This approach would both increase the access of researchers to the most relevant clinical questions and expertise, and would also expedite the application of research, which might otherwise lag years behind the initial discoveries. Progress will also require considerable improvements in medical informatics to make the relevant clinical data available for research.

At a practical level, some simple steps might facilitate greater interaction between clinical and laboratory aspects of pharmacogenetics. Clinicians should be involved in framing research questions in pharmacogenetics and be integrated in research projects. Geneticists need to explain the principles and details of research to clinical colleagues, as in, for example, the recent efforts of some journals to bring pharmacogenetics to a broader non-genetic audience (for example, see REF. 43). Clinicians will need to be informed through clinical routes about relevant pharmacogenetic findings, how a given test can actually be ordered for their patients and the actual pharmacogenetics contribution to variable drug response in specific diseases. Trials that recruit patients should include pharmacogenetic components. Pharmacogenetics itself needs to be brought to the lay public and patients, perhaps through patient-support organizations.

These will not be trivial undertakings, in part because many clinicians are not persuaded that genetic information has much relevance to health care. Given the state of knowledge this is probably a broadly reasonable judgement, but as we argue here, gaining the necessary knowledge will require increased enthusiasm in the clinical community. This makes the inclusion of expanded genetics coverage in medical courses an urgent priority and argues for mechanisms of promoting greater interaction between clinicians and the genetics research community.

1. Weber, W. W. *Pharmacogenetics* (Oxford Univ. Press, New York and Oxford, 1997).
2. Motulsky, A. G. Drug reactions, enzymes, and biochemical genetics. *J. Am. Med. Assoc.* **165**, 835–837 (1957). **This paper was the first to make explicit the hypothesis that variation among individuals in drug response might be owing to subtle genetic differences with little or no obvious phenotypes except in response to the relevant drug. This perspective contributed directly to the development of pharmacogenetics as a distinct discipline.**
3. Mahgoub, A., Idle, J. R., Dring, L. G., Lancaster, R. & Smith, R. L. Polymorphic hydroxylation of debrisoquine in man. *Lancet* **2**, 584–586 (1977).
4. Nebert, D. W. & Dieter, M. Z. The evolution of drug metabolism. *Pharmacology* **61**, 124–135 (2000).
5. El-Bacha, R. S. & Minn, A. Drug metabolizing enzymes in cerebrovascular endothelial cells afford a metabolic protection to the brain. *Cell. Mol. Biol.* **45**, 15–23 (1999).
6. Siddiqui, A. *et al.* Association of multidrug resistance in epilepsy with a polymorphism in the drug-transporter gene *ABCB1*. *N. Engl. J. Med.* **348**, 1442–1448 (2003). **This paper reports the first polymorphism that was associated with the efficacy of anti-epileptic drugs and was one of the first pharmacogenetic studies to assess LD patterns around the associated variant and rule out population stratification.**
7. Sisodiya, S. M. Mechanisms of antiepileptic drug resistance. *Curr. Opin. Neurol.* **16**, 197–201 (2003).
8. Shah, R. R. The significance of QT interval in drug development. *Br. J. Clin. Pharmacol.* **54**, 188–202 (2002).
9. Evans, W. E. & Relling, M. V. Pharmacogenomics: translating functional genomics into rational therapeutics. *Science* **286**, 487–491 (1999).
10. Phillips, K. A., Veenstra, D. L., Oren, E., Lee, J. K. & Sadee, W. Potential role of pharmacogenomics in reducing adverse drug reactions: a systematic review. *JAMA* **286**, 2270–2279 (2001).
11. Leabman, M. K. *et al.* Natural variation in human membrane transporter genes reveals evolutionary and functional constraints. *Proc. Natl Acad. Sci. USA* **100**, 5896–5901 (2003).
12. Fellay, J. *et al.* Response to antiretroviral treatment in HIV-1-infected individuals with allelic variants of the multidrug resistance transporter 1: a pharmacogenetics study. *Lancet* **359**, 30–36 (2002).
13. Gratze, G. *et al.* β-2 adrenergic receptor variants affect resting blood pressure and agonist-induced vasodilation in young adult Caucasians. *Hypertension* **33**, 1425–1430 (1999).
14. Dishy, V. *et al.* The effect of common polymorphisms of the β2-adrenergic receptor on agonist-mediated vascular desensitization. *N. Engl. J. Med.* **345**, 1030–1035 (2001).
15. Drysdale, C. M. *et al.* Complex promoter and coding region β2-adrenergic receptor haplotypes alter receptor expression and predict *in vivo* responsiveness. *Proc. Natl Acad. Sci. USA* **97**, 10483–10488 (2000). **This paper relates *ADRB2* haplotypes to beta-agonist response in asthmatics and shows that haplotypes are a better predictor of response than individual SNPs.**
16. Kaye, D. M. *et al.* β-adrenoceptor genotype influences the response to carvedilol in patients with congestive heart failure. *Pharmacogenetics* **13**, 379–382 (2003).
17. Palmer, L. J., Silverman, E. S., Weiss, S. T. & Drazen, J. M. Pharmacogenetics of asthma. *Am. J. Respir. Crit. Care Med.* **165**, 861–866 (2002).
18. Hirschhorn, J. N., Lohmueller, K., Byrne, E. & Hirschhorn, K. A comprehensive review of genetic association studies. *Genet. Med.* **4**, 45–61 (2002).
19. Lohmueller, K. E., Pearce, C. L., Pike, M., Lander, E. S. & Hirschhorn, J. N. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nature Genet.* **33**, 177–182 (2003). **This paper analyses 301 published studies covering 25 different reported associations for common disease. The analyses show that false-positive results cannot explain all of the associations.**

20. Gottesman, M. M., Fojo, T. & Bates, S. E. Multidrug resistance in cancer: role of ATP-dependent transporters. *Nature Rev. Cancer* **2**, 48–58 (2002).
21. Sham, P. C., Cherny, S. S., Purcell, S. & Hewitt, J. K. Power of linkage versus association analysis of quantitative traits, by use of variance-components models, for sibship data. *Am. J. Hum. Genet.* **66**, 1616–1630 (2000).
22. Pritchard, J. K. & Rosenberg, N. A. Use of unlinked genetic markers to detect population stratification in association studies. *Am. J. Hum. Genet.* **65**, 220–228 (1999).
23. Reich, D. E. & Goldstein, D. B. Detecting association in a case-control study while correcting for population stratification. *Genet. Epidemiol.* **20**, 4–16 (2001).
24. Kohlmeier, L., DeMarini, D. & Piegorsch, W. in *Design Concepts in Nutritional Epidemiology* (eds Margetts, B. & Nelson, M.) 312–337 (Oxford Univ. Press, Oxford, 1997).
25. Stefansson, H. *et al.* Neuregulin 1 and susceptibility to schizophrenia. *Am. J. Hum. Genet.* **71**, 877–892 (2002).
26. Peltonen, L. & McKusick, V. A. Genomics and medicine. Dissecting human disease in the postgenomic era. *Science* **291**, 1224–1229 (2001).
27. Botstein, D. & Risch, N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nature Genet.* **33** (Suppl.), 228–237 (2003).
28. Daly, M. J., Rioux, J. D., Schaffner, S. F., Hudson, T. J. & Lander, E. S. High-resolution haplotype structure in the human genome. *Nature Genet.* **29**, 229–232 (2001). **This paper introduces the idea of the human genome being partitioned into discrete haplotype blocks with limited diversity, which provided motivation for the HapMap project.**
29. Johnson, G. C. *et al.* Haplotype tagging for the identification of common disease genes. *Nature Genet.* **29**, 233–237 (2001).
30. Gabriel, S. B. The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229 (2002).
31. Goldstein, D. B., Ahmadi, K. R., Weale, M. E. & Wood, N. W. Genome scans and candidate gene approaches in the study of common diseases and variable drug responses. *Trends Genet.* **19**, 615–622 (2003).
32. Holtzman, M. J. Drug development for asthma. *Am. J. Respir. Cell Mol. Biol.* **29**, 163–171 (2003).
33. Goldstein, D. B. Pharmacogenetics in the laboratory and the clinic. *N. Engl. J. Med.* **348**, 553–556 (2003).
34. Liggett, S. B. Polymorphisms of the β2-adrenergic receptor and asthma. *Am. J. Respir. Crit. Care Med.* **156**, 156–162 (1997). **This paper introduces a model of receptor transport to explain responses to beta-agonists and shows the importance of combining functional and association data.**
35. Noah, L. Author reply. *N. Engl. J. Med.* **348**, 2041–2043 (2003).
36. Shah, J. Economic and regulatory considerations in pharmacogenomics for drug licensing and healthcare. *Nature Biotechnol.* **217**, 747–753 (2003).
37. Roses, A. D. Pharmacogenetics and future drug development and delivery. *Lancet* **355**, 1358–1361 (2000).
38. Xie, H. G., Kim, R. B., Wood, A. J. & Stein, C. M. Molecular basis of ethnic differences in drug disposition and response. *Annu. Rev. Pharmacol. Toxicol.* **41**, 815–850 (2001).
39. Carmena, R., Roederer, G., Mailloux, H., Lussier-Cacan, S. & Davignon, J. The response to lovastatin treatment in patients with heterozygous familial hypercholesterolemia is modulated by apolipoprotein E polymorphism. *Metabolism* **42**, 895–901 (1993).
40. Ojala, J. P. *et al.* Effect of apolipoprotein E polymorphism and XbaI polymorphism of apolipoprotein B on response to lovastatin treatment in familial and non-familial hypercholesterolaemia. *J. Intern. Med.* **230**, 397–405 (1991).
41. Corder, E. H. *et al.* Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* **261**, 921–923 (1993).

42. Strittmatter, W. J. *et al.* Apolipoprotein E: high-avidity binding to β-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc. Natl Acad. Sci. USA* **90**, 1977–1981 (1993).
43. Roses, A. D. Genome-based pharmacogenetics and the pharmaceutical industry. *Nature Rev. Drug Discov.* **1**, 541–549 (2002). **This paper describes the use of a candidate-gene approach to identify variants for rare adverse drug reactions (abacavir hypersensitivity).**
44. DiMasi, J. A., Hansen, R. W., Grabowski, H. G. & Lasagna, L. Cost of innovation in the pharmaceutical industry. *J. Health Econ.* **10**, 107–142 (1991).
45. DiMasi, J. A., Hansen, R. W., Grabowski, H. G. & Lasagna, L. Research and development costs for new drugs by therapeutic category. A study of the US pharmaceutical industry. *Pharmacoeconomics* **7**, 152–169 (1995).
46. DiMasi, J. A., Hansen, R. W. & Grabowski, H. G. The price of innovation: new estimates of drug development costs. *J. Health Econ.* **22**, 151–185 (2003).
47. Mancinelli, L., Cronin, M. & Sadee, W. Pharmacogenomics: the promise of personalized medicine. *AAPS PharmSci.* **2**, 4 (2002).
48. Seidman, E. G. Clinical use and practical application of TPMT enzyme and 6-mercaptopurine metabolite monitoring in IBD. *Rev. Gastroenterol. Disord.* **3** (Suppl.), 30–38 (2003).
49. Marra, C. A., Esdaile, J. M. & Anis, A. H. Practical pharmacogenetics: the cost effectiveness of screening for thiopurine s-methyltransferase polymorphisms in patients with rheumatological conditions treated with azathioprine. *J. Rheumatol.* **29**, 2507–2512 (2002).
50. McLeod, H. L. & Siva, C. The thiopurine S-methyltransferase gene locus — implications for clinical pharmacogenomics. *Pharmacogenomics* **3**, 89–98 (2002).
51. van Aken, J., Schmedders, M., Feuerstein, G. & Kollek, R. Prospects and limits of pharmacogenetics: the thiopurine methyl transferase (TPMT) experience. *Am. J. Pharmacogenomics* **3**, 149–155 (2003).
52. Pellock, J. M. Felbamate. *Epilepsia* **40** (Suppl.), 57–62 (1999).
53. Kapetanovic, I. M. *et al.* Reactivity of atropaldehyde, a felbamate metabolite in human liver tissue *in vitro*. *Chem. Biol. Interact.* **142**, 119–134 (2002).
54. Roller, S. G. *et al.* Interaction between human serum albumin and the felbamate metabolites 4-hydroxy-5-phenyl-[1,3]oxazinan-2-one and 2-phenylpropenal. *Chem. Res. Toxicol.* **15**, 815–824 (2002).

◉ Online links

DATABASES
The following terms in this article are linked online to:
LocusLink: http://www.ncbi.nlm.nih.gov/LocusLink
ABCB1 | *ADRB2* | *ALOX5* | *APOE* | *CYP1A2* | *CYP2D6* | *NAT2* | *SCN1A* | *SCN1B* | *SCN2A* | *SCN2B* | *SCN3A* | *SCN8A* | *TPMT*
OMIM: http://www.ncbi.nlm.nih.gov/omim
acute lymphoblastic leukaemia | Alzheimer disease | chronic myeloid leukaemia | schizophrenia
Swiss-Prot: http://us.expasy.org/sprot
CYP2C9 | CYP2C19 | CYP2D6 | MDR1 | MRP1 | MRP2

FURTHER INFORMATION
deCode genetics: http://www.decodegenetics.com
PubMed: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi
Tag selection software (Goldstein laboratory web site):
http://popgen.biol.ucl.ac.uk
Access to this interactive links box is free online.