INTRODUCTION

As soon as the two independent sequence drafts of the human genome were surfaced (1–2), it was clearly established that the two ‘genomes’ that each of us carry, inherited from our parents, most often differ from each other, and from the genomes of other humans (3). These genetic differences interact with the environmental factors in producing varied phenotypes, inducing differential human susceptibility to diseases and leading to differential response to pharmacological agents (4–5). Thereafter, the scientific literature got flooded with the reports of role of genetic variation in disease susceptibility, the age of onset, its severity and finally, its treatment (6–9).

The findings were more important in the light of the fact that the response of patients to drug therapy varies widely from one individual to another worldwide, both in terms of the beneficial effects of the drug as in the occurrence of serious and often unpredictable side-effects or adverse drug reactions. The clinical consequences of adverse drug reactions range from patient discomfort to serious illness, which requires hospitalization and may even result into
death. Adverse drug reactions accounted for 2.2 million hospitalizations and over 100,000 deaths in the USA in 1994, making adverse drug reactions the 4th-6th leading cause of death in the USA. In the UK, 1 in 5 hospitalizations can be attributed to adverse drug reactions (8). None of the factors including age, gender, body weight, patient health, disease status, diet, smoking, alcohol, exercise and drug interactions provided any guarantee that a given treatment will be effective or well tolerated in a given patient. However, the projects of human genome sequencing and human genome diversity and the subsequent reports compelled the gentists and pharmacologists all over the world to believe that the major cause for variability in drug responses lies in a patient’s genetic make-up (8–13).

The findings lead to the emergence of the filed of ‘Personalized Medicine’; it simply means the prescription of specific treatments and therapeutics best suited for an individual and avoids the trial and error approach of conventional medicines. Pharmacogenomics, Pharmacokinetics and Pharmacoproteomics are the basic foundation of personalized medicine while molecular diagnostics being the major tool.

1.1 Pharmacogenetics and pharmacogenomics

The study of genetic variants or polymorphisms that influence drug responses has given rise to a principal challenge in the analysis of these data mainly, the difficulty in linking information about the variation in human genes to the variation in drug response (pharmacogenetics) and to understand how interacting genes determine individual drug responses (pharmacogenomics) (7, 13). The joint disciplines of pharmacogenetics and pharmacogenomics are interdisciplinary and collaborative fields requiring the cooperative efforts of research and clinical scientists (Fig. 1).

Pharmacogenetics is defined as the study of variability in drug responses due to heredity. Pharmacogenomics examines the role of the entire genome in both disease susceptibility and drug response, in an attempt to identify specific genes that are associated with specific diseases and that may be the targets for new drugs (10). Pharmacogenetic variations can be classified into three main types according to their mechanism of action. The first is pharmacokinetic, in which genetic variants are associated with drug transporters and metabolizing enzymes, and lead to alterations in the uptake, distribution and elimination of drugs. The second is pharmacodynamic, in which genetic variation occurs in the drug target or a component of the target pathway leading to altered drug efficacy. Pharmacodynamic targets include receptors, ion channels, enzymes, transducer and regulatory proteins, and immune molecules. A third mechanism is idiosyncratic, in which genetic variants result in unintended actions of a drug outside its therapeutic indication (11–13).

The role of genetic variation in drug response or adverse reaction defines the science of Pharmacogenetics (13), and it actually started in 1950s with the emergence of human biochemical genetics. The best cited example is occurrence of hemolytic anemia due to G6PD deficiency. Way back in 1957, Motulsky published a report about
drug reactions, enzymes and biochemical genetics. The term ‘pharmacogenetics’ was basically coined by Friedrich Vogel of Heidelberg, Germany in 1959 (14). In late 1960s, Vesell showed remarkable similarity of disposal for several drugs in monozygotic twins who share 100% of their genes as contrasted to dizygotic twins who share only 50% (15). The term ‘pharmacogenomics’ was re-introduced in 1990s with emergence of the Human Genome Project and the development of the human genome sciences.

Latest high throughput technology has added newer dimension to the search of multiple genes and their expression affecting drug responses. Recent developments in technology have revealed large number of new mutations even for diseases like hemolytic anemia. A 24 bp deletion of nucleotide 953–976 in the exon 9 of the G6PD gene causes the G6PD deficiency. The parents were found to be heterozygous for this mutation and appropriately advised on the condition and the importance of taking folic acid regularly (16). Search for characteristic cellular DNA abnormalities in disease is now beginning to guide
construction of therapeutic drugs acting on
disease specific DNA mutations (17). A
somatic mutation in chronic myelocytic
leukemia responds to the drug Gleevec in
almost 100% of cases (18).

Similarly, Long-QT syndrome is a
clinically and genetically heterogeneous
syndrome characterized by lengthening of
the QT interval and increased dispersion of
the ventricular repolarization on surface
electrocardiogram and a propensity to
malignant ventricular arrhythmias, torsade
de pointes and ventricular fibrillation, which
may lead to sudden cardiac death. Long-QT
syndrome mostly affects adolescents and
young adults with structurally and
functionally normal hearts and is caused by
aberrations in potassium and sodium ion
channels (19). Standard therapies for long-
QT syndrome include correction of the
underlying cause, alleviation of the
precipitating factors, magnesium sulfate,
isoproterenol, antiadrenergic therapy (beta-
adrenergic receptor blockers, left
cervicothoracic sympathectomy), cardiac
pacing, and implantable cardioverter
defibrillator. The potential therapies include
sodium channel blockers (mexiletine,
flecainide, lidocaine, pentisomide, phenytoin),
potassium, potassium channel activators
(nicorandil, pinacidil, cromakalim), alpha-
adrenergic receptor blockers, calcium
channel blockers, atropine, and protein
kinase inhibitors. There is individual
variation and different therapies are being
tried in response to the personal genetic
variation (19).

1.2 How genetic variation affects susceptibility
to a disease or drug action

Humans are considerably more similar
to each other than other species as any two
randomly chosen humans differ at ~1 in 1000
nucleotide pairs, whereas two random
chimpanzees differ in ~1 in 500 nucleotide
pairs (20). Nevertheless, there are ~3 billion
nucleotides on haploid human genome,
therefore on an average two humans differ
at ~3 million nucleotides which is ~0.1–0.2%
of the haploid human genome (21). Most of
these variants are neutrals but still common
variants are present in coding and regulatory
regions of genes that alter the amino acid
sequences and gene expression respectively
(22).

Majority of the polymorphisms observed
are single nucleotide polymorphisms (SNPs),
short Tandem repeat polymorphisms (STRPs)
or insertion-deletion polymorphisms (indels).
If these polymorphisms are present in the
coding or the regulatory region of various
genes whose products participates in the
pathology, physiology or treatment of a
particular disease, then their presence might
lead to individual variation both for the
occurrence of the disease or the action of
drug against it.

An ideal example of such genetic control
is evidenced by the identification of
polymorphisms in the gene regulatory region
of cytokines, chemokines and growth factors
that correlates with intra individual
variations in actual cytokine or chemokine
production (23–25). As these polymorphisms
segregate independently, each individual is
a mosaic of high, moderate and low cytokine
producing phenotype. Presence of such
high, moderate and low producing SNPs
in the regulatory domain of cytokines leads
to their differential, responses against
immunosupressors in case of organ and bone
marrow transplantations and against anti-inflammatory drugs in various diseases like asthma.

A vast majority of such polymorphisms are found in the regulatory promoter regions, while various others are found in the intronic, exonic and untranslated regions. The promoter gene polymorphisms may disrupt or abolish transcription regulatory elements such as those involved for NFKB and STATS (signal transducers and activators of transcription) (26–27). All these regulatory elements regulate RNA polymerase binding, influencing the rate at which gene is transcribed into mRNA. Intronic variation may affect enhancer/silencer sequences and certain polymorphisms may alter architectural transcription factor binding elements (27).

Furthermore, the differential distribution of the genetic variation in different populations owing to various factors like mutations, selection, and genetic drift and to some extent migrations and mating patterns also leads to the differential susceptibility to diseases and differential responses to drugs (5–6). Natural selection is the result of population variation among individual genotypes in their probabilities of survival and/or reproduction. Earliest evidences include selection of Heterozygotes of hemoglobin A/S polymorphism for having greater resistance to malaria. Similarly, FOXP2 gene has shown a two amino acid difference in human and chimpanzees suggesting the role of this gene in evolution of speech and language in modern humans (28). Strong molecular evidences are available for the selection of activity of G6PD locus to confer resistance to malaria (29) and TNFSF5 in response to infectious agents (30). Other potent examples of genes having a signature of selection are Dufry antigen (31), drug metabolism- CYP1A2 (32) and alcohol metabolism- ADH1B and ALDH2 (33).

Random genetic drift occurs due to a finite number of individuals participating in the formation of the next generation. This process is responsible behind the genetic differences between African and non-African populations (3). The impact of drift on differential susceptibility of diseases could be seen from the fact that sickle-cell anemia is found in wide range of people including Hispanics and inhabitants of north-western India (34), but the blacks of South Africa do not carry sickle cell traits (34–35). Similarly, high frequency of null allele of CYP2D6 makes Arabs more capable of transforming codeine into the active form morphine.

1.3 Approaches to pharmacogenomics and pharmacogenetics studies

An ideal pharmacogenomics or pharmacogenetics study involves various crucial steps.

(a) Identification of appropriate candidate genes whose expression may influence drug action or disease pathogenesis.

(b) Identifying all observed polymorphisms with complete sequence.

(c) Carry out case-control studies in various populations to identify the association of a genetic variant with disease or drug response, and

(d) Assess the effect of the genetic variant on the expression profile of the gene.
These four steps together with the elements of pharmacokinetics and pharmacodynamics, i.e., the variability in pharmacokinetics of the concerned drug and variability in association with drug and/or genetic variant on the phenotype which completes the overall requirement for an ideal study of personalized medicines. Such study require a database that can model key elements of the data, acquire data efficiently, provide query tools for analysis and deliver the resulting system to the scientific community (7). Pharmacokinetics includes absorption, distribution, metabolism and elimination of a drug. Pharmacodynamics includes pharmacological effects and clinical response leading to toxicity and efficacy for a particular drug or metabolite (7).

(A) Candidate gene selection

The candidate gene approach for discovering genetic markers use experimentally derived a priori knowledge about a disease or a drug. Scientists do background research, employing both public and proprietary databases, to identify appropriate candidate genes whose expression may impact drug action or disease pathogenesis. Candidate genes are commonly selected based on metabolic pathways, molecular targets, biological response pathways and/or disease risk. Generally, the genes are ranked based upon their perceived likelihood of being involved in the drug response. The stronger candidates can then be tested first. If these top candidates fail to explain sufficiently the variation in drug response, additional candidate genes can then be tested.

A classic example of this approach is the variation in individual responses to the anti-leukemia drug, 6-mercaptopurine (38–39). Thiopurine drugs are metabolized, in part, by S-methylation catalyzed by thiopurine S-methyltransferase (TPMT). Most people metabolize the drug quickly. Some individuals, with a genetic variation for the enzyme TPMT, do not. Consequently, they need lower doses of 6-mercaptopurine for effective treatment as normal doses can be lethal. Patients with very low or undetectable TPMT activity are at high risk of severe, potentially fatal hematopoietic toxicity when they are treated with standard doses of thiopurines (38). As human TPMT activity is controlled by a common genetic polymorphism, it is an excellent candidate for the clinical application of pharmacogenetics. The point mutations in the TPMT gene that cause the loss of TPMT activity can be detected by a fluorescently labeled amplified DNA which is hybridized with oligonucleotide DNA probes immobilized in gel pads on a biochip. The specially designed TPMT biochip can recognize six point mutations in the TPMT gene and eight corresponding alleles associated with TPMT deficiency: TPMT(*2); TPMT(*3A), TPMT(*3B), TPMT(*3C), TPMT(*3D), TPMT(*7), and TPMT(*8). In such instance identification of the candidate gene TPMT was the most decisive step and information about its action on thiopurine drugs helped in its recognition as a potent gene responsible for varied pharmacological action (39).

(B) Identification of genetic variants

The pharmacogenomics studies require a detailed model of genomic sequence, in order to represent accurately DNA sequence
data, gene structure and polymorphisms in sequence, much more than simply storing the DNA sequence. Recent advents of automated DNA sequencing (40), denaturing HPLC (41), mass spectrophotometry (42), array based re-sequencing (43), automated fragment size analysis and SnapShot PCR have not only increased our repertoire of mutation but have also provided a high throughput techniques for quick and reliable genotyping. However, still the choice of technique varies from RFLP, ARMS, gene-scanning and re-sequencing to dHPLC depending upon the type of marker being studied.

Single nucleotide polymorphisms, SNPs, are the most abundant and the simplest form of DNA variation. A SNP originates with a mistake in copying a single nucleotide letter in a DNA sequence. The mistake is simply that one letter gets replaced with another. There are different effects that this change may have. Their effects vary from being silent to lethal (44). There are 1.42 million known SNPs found at a density of one SNP per 1.91 kb. This means that more than 90% of any stretch of sequence 20 kb long will contain one or more SNPs. The density is even higher in regions containing genes. The International SNP Map Working Group estimates that they have identified 60,000 SNPs within genes (‘coding’ SNPs), or one coding SNP per 1.08 kb of gene sequence. Moreover, 93% of genes contain a SNP, and 98% are within 5 kb of a SNP (44).

The main use of the human SNP scoring is to dissect the contributions of individual genes to various diseases that have a complex and polygenic basis. Variation in genome sequences underlies differences in our susceptibility to, or protection from, all kinds of diseases; in the age of onset and severity of illness; and in the way our bodies respond to treatment. These gene variants lead to tissue and organ incompatibility, affecting the success of transplants (22).

Large number of genes and their variants are implicated in the success of renal transplantation; as a result many impending and potent methods have also become available for genotyping of SNPs and STR loci in last few decades. These methods vary from sequence specific primer (SSP) based typing to automated DNA sequencing of the whole fragment where the SNP is located. Other flourishing methods include restriction fragment length polymorphism and amplification refractory mutational system (ARMS). However, Studies covering underlying genetic component of a diverse number of conditions require systems for genotyping large numbers of SNPs. In this regard, a mini-sequencing method known as Snapshot and Taqman probes based real time PCR are ideal techniques, where former is effective for small-scale investigation while later is used for high throughput analysis (22).

(C) Case-control based association studies

Association studies offer a potentially powerful approach to identify genetic variants that influence susceptibility to common disease as well as response of a particular drug (45–47). However, such studies are often plagued by non-reproducible results (48). In principle, the inconsistency may be due to false positive studies, false negative studies or true variability in association among different populations.
It has been widely accepted that undetected population stratification in case-control studies is a major reason behind the false positive associations. Association studies can yield large numbers of spurious associations if population subgroups are unequally represented among cases and controls. The problem of population stratification increases when cases or controls are derived from a metropolitan city where people of numerous ethnic backgrounds live together or there has been a genetic mixing of two or more groups.

Knowler et al have studied the Gm polymorphism of human immunoglobulin IgG gene and type II Diabetes among 4290 Native American individuals belonging to Pima and Papago Indians. They found a highly significant negative association. However, once they included the criteria of genetic ancestry then they found that frequency of Gm polymorphism is much higher and occurrence of type II diabetes is very low in people of no American Indian ancestry than those of American Indian ancestry.

With recent advancement in human genetic variation studies, there has been a broad consensus on the fact that genetic knowledge of population sub-structuring and stratification is an essential requirement for proper selection of controls and for identifying disease pre-disposing alleles that may differ across ethnic groups. An interesting example in this regard would be of Indian populations. Various reports cite the possible Caucasian genetic ancestry of north Indian populations reflected by the high frequency of western Eurasian and central Asian NRY haplogroup-R1a Rib, R2 etc. However, both FVL mutation (A2086G) and prothrombin gene mutation (C10965T) that have been reported to be associated with CAD and thrombotic events among Europeans were not found in any of the samples in our study on normal as well as CAD and RSA patients from north India. Similarly, the ACE DD genotype has been associated with hypertension and CAD but no such results are found among north Indians. However, same ACE DD genotype has been found highly associated (P=0.0001; OR 25.7) with patients of end stage renal disease. Such studies ultimately help in deciding the doses and type of anti-hypertensive therapies.

Similarly, studies based on successful association of MTHFR (677C/T) SNP with increase in the homocysteine levels due to production of thermpliable MTHFR enzyme leads to deciding a quantity and dosages of folic acid supplementation. Another interesting examples is strong association Factor V-leiden SNP (1691G/A), and Prothrombin gene mutation (20210G/A) with hyper coagulable state that helps in to fix a proper anti-coagulatory therapy. Furthermore, immunosuppressor drugs like cyclosporin and tacrolimus carry a narrow therapeutic range and a wide inter individual variation in its pharmacokinetics. The most important factor that affects their inter-patient variability is p-glycoprotein (p-GP), a product of multidrug resistant gene-1 (MDR-1) that uses these immuno-suppressors as its substrate. Four important SNPs have been studied in MDR-1 gene namely 129T/C in exon 1b, 1236C/T in exon 12, 2677G/T in exon 21 and 3435C/T in exon 26, where later three SNPs are more frequently found in general
population (69–71). Exon 21 SNP was found highly associated with tacrolimus dose as TT homozygous showed 40% higher dose requirement of teclarimus but not for cyclosporin doses (70). Exon 26 SNP-3435C/T is reported to be associated with increased expression of MDR1 gene in CD56+ NK cells (70–71). However, the report regarding its role in drug kinetics is contradictory. More focused and targeted studies on such genes will lead to individualization of drug treatment.

(D) Expression analysis

Expression analysis both at mRNA and protein level is an ardent need to understand the contribution of genetic variant in stimulating or suppressing the expression and thereby assessing the exact role on the drug kinetics. This provides the precise mechanism involved in particular individual behind the success or failure of a drug. In fact, this particular step is the most essential step before declaring a particular SNP to be involved in the variability of drug responses (22). Various recent studies have done such analysis where the information generated by the genotypic profile of the patients was substantiated by the mRNA expression and ELISA based assays (23). Such an exercise can clear the controversy about the role of the SNPs in the pleotropic molecules like cytokine IL6 that acts as both pro and anti-inflammatory.

(E) Modeling of phenotypic data and its correlation with the genotype

Finally, it is mandatory to have phenotype data revealing molecular, cellular and clinical profile to correlate with the presence of a genetic variants and its effect on the expression of the related gene. Molecular and cellular phenotype data include enzyme kinetic measurements, such as binding, catalysis and inhibition constants for particular drugs, cellular drug processing rates, homology modeling of three-dimensional structures and pharmacodynamic assays (7). Clinical phenotype is perhaps the most difficult data to model and link with genomic and molecular/cellular phenotypic data. Clinical phenotype data include basic pharmacokinetic measurements (such as drug absorption, distribution, elimination and metabolism) as well as pharmacodynamic profiles, which currently include pulmonary, cardiac and psychological function tests, and cancer chemotherapeutic side effects.

Apart from the most successful candidate gene approaches, whole genome analysis and use of various statistical parameters is also often used for pharmacogenetic studies.

Whole genome analysis

A whole genome analysis is effectively the opposite of a candidate gene based study. Rather than focusing on a set of genes that are already expected to be involved in how patients respond to treatment with a drug, one attempts to test the entire genome. The obvious benefit of this approach is that genetic loci that no one ever expected to be involved in the response may be revealed, potentially adding greatly to the understanding of the drug, the disease or general biology. Another reason to take this approach is if a good candidate gene list cannot be assembled based on prior knowledge.
There are reasons why most Pharmacogenomic studies are not done this way. First, the expense is likely to be prohibitive, although some researchers have pooled their samples, which raises several issues. Second, as a compromise to reduce the expense, each gene generally has to be measured with lower resolution. For example, by choosing only one SNP to represent a gene, the power to detect the true genetic factors involved in drug response is reduced. Third, given the vast number of genetic loci that have to be tested to cover the entire genome (several thousand to several hundred thousand), the true findings are likely to be swamped by a sea of false positives. It is generally a bad experimental design to have many more tests than subjects and most clinical trials are limited to a few thousand subjects. (72). This final problem will remain regardless of how significantly genotyping costs may decrease as new technologies are developed.

One of disease tried widely is hypertension. Increasingly, detailed characterization of human molecular genetic variation will facilitate the use of genetic information in preventing, diagnosing, and treating common diseases (73). One promising application is the identification of genetic variants influencing responses to drugs used to lower blood pressure (BP) and prevent target-organ complications of hypertension. This update on gene markers to guide antihypertensive therapy highlights polymorphisms recently reported to predict inter individual differences in response to antihypertensive medications. However, single-site variation in most genes makes only a small contribution to differences in BP response, and, after all known genetic and environmental predictors have been considered, most variation in responses still remains unexplained. Advancing beyond our current "trial-and-error" approach in selecting drug therapy in individual patients will undoubtedly require whole-genome approaches to discover additional, novel genetic pathways influencing drug response. In addition, larger samples will be required to more fully characterize genetic variation within candidate genes and to consider the joint effects of gene-gene and gene-environment interactions. Eventually, knowledge of genetic variants that influence BP responses may allow more individualized tailoring of therapy to optimally reduce BP and target-organ damage (73).

**Statistical approach**

After all the hard clinical and laboratory work, the data produced must be interpreted to determine which genetic variants affect which clinical variables and under what conditions (74–76). This part of the process is called statistical analysis. Regardless of the experimental design (candidate gene or whole genome, prospective or retrospective), statistical analysis of the resulting data is a tremendous challenge. While it is important to have a statistical analysis plan in place before launching into a Pharmacogenomic project (74), the nuances of how to interpret the results may not be apparent until the data are produced. Here we present a few of the common challenges with which one must deal when working with genetic data, which extend beyond the complexities inherent in any clinical study.

1. Genetic variables are not simple categorical or quantitative traits like gender or age. The variation found in a single gene is instead quite complex. Often, there may be a thousand of
different ways to break the variation into smaller pieces, or “markers” that range in scale from a single SNP variant to a whole-gene haplotype. Because the true causative aspect of a gene variation could occur at any level, we would like to test all of these possibilities. Many of these alternative markers within a gene, however, are correlated with each other to some degree but differ nevertheless. Determining the most plausible association may be more involved than just finding the best “P-value”.

2. For each genetic variable, there may be several ways, or “genetic models”, for how the variable relates to a clinical trait. Is the genetic marker dominant or recessive? and does this increases or reduces the efficacy or side effects of the drug? These genetic models are all possible. Examining the variation found in the clinical trait may suggest which models are likely. For instance, if there are two distinct responses to the drug, dominant and recessive models would be reasonable models to consider. However, in most cases, the pattern of the variation in treatment response does not allow a clear determination of which models to consider, so multiple models should be tested.

3. As many genes are found on the same chromosome, sometimes very close together, a finding for one gene may really reflect the effect of a neighboring gene. All positive findings must be checked for this possibility. If a neighboring gene is plausibly related to the clinical trait (e.g., the response to a drug), secondary analyses, perhaps requiring additional laboratory work, may be required to identify the gene that is really involved in the response.

4. Because biology is rarely as simple as one might hope, variation in multiple genes is likely to be involved in determining each aspect of drug response. Once the most promising individual genes are identified then how these components interact or combine to affect the clinical trait should be determined.

5. In a “multiple comparison testing”, a hundred of genes with a thousand markers, each tested in two models for five clinical traits, requires 1,000,000 tests. A large fraction of these tests will be found to be nominally significant in the statistical sense. Correcting for this over-abundance of positive findings require a clear understanding and determination of the complex correlations among markers, among genes, among models and among clinical traits.

6. Just as many biological characteristics, such as disease risks, differ between ethnic groups, so do the frequencies of many genetic variants. This fact raises the concern that when a genetic marker is found that predicts drug response, the marker might really just be tracking ethnicity, as would countless other genetic variants that have nothing to do with the clinical trait in which we are interested. While it is possible to control this problem by including factors, such as ethnic self-identity of patients in the analyses, the level of control provided might not be adequate and sometimes is not available or the information is
unreliable or ambiguous. One analytical tool that is now used to test for this confounding effect is known as “genomic control”. Genomic control is simply a fancy way of saying that scientists can measure a set of genetic loci that are independent of one another and unlikely to be related to the clinical traits and then use the unrelated loci to assess whether there are unsuspected ethno geographic differences between the patients with different clinical traits, e.g. those who respond well to the drug versus those who do not.

1.4 Recent diagnostic tools for studies on personalized medicine

**Ramification amplification method**: It is an isothermal process and sensitive and specific for amplifying nucleic acids with flexibility to analyze proteins and other small molecules on the same analytical system. This is a high throughput technique where in one go many SNPs can be amplified (77–78).

**Invader assays**: This technique is quite robust where perfect match enzyme substrate reactions using propriety cleavage enzymes, recognizes and cut only the specific structure formed during the invader process. It can be used directly to recognize the SNPs (79–80).

**Molecular beacons**: These are folded probes which gives no fluorescent signals in the folded position due to quenching of the label but upon hybridization of the molecular beacons to the target sequence the probe unfolds and the fluorescent labels emits light. Molecular beacons are able to discriminate alleles in real time PCR assays of genomic DNA. Molecular beacons are ideal tools for genetic screening and diagnosis (81).

**Matrix-assisted laser desorption mean spectrophotometry**: This is also known as MALDI - TOF - MS. This determines the mass of the variant DNA sequences. As large number of sequences can be detected in a single reaction so it is a cost effective technique for genotyping (82).

**Biochip microarray**: This is a technique which can be adapted for genomic DNA, RNA expression analysis and now even the protein chips have been developed. It has powerful data processing software, automated flow-through system and label free detection of the hybridization signals (83–84).

**Nanochip technologies**: This is an accurate method that allows multiplex assays and on chip amplification of DNA material directly on the Nanochip cartridge, which eliminates a time-consuming preparatory step and involves only a single step (85).

1.5 Pharmacogenomics: success story till date

As it has been clearly figured out that the evolutionary processes of mutation, migration, random genetic drift and selection resulted into differential distribution of normal genetic variation and also that of genetic variation affecting diseases and drug responses (4). The genetic variations studies based on the role of differential genomic profiles on disease susceptibility and drug response have contributed tremendously in two major areas (i) molecular sub-classification of the diseases based on the genetic profile (ii) genetic contribution to differential drug response.
Molecular sub-classification of the diseases based on the genetic profile

The impact of evolutionary forces on the differential distribution of disease genes is currently better understood in the context of the worldwide distribution of the monogenic traits (86). The distinctive examples include the parallel presence of high frequency of hemoglobin HbS allele, variant of G6PD and sickle cell anemia among sub-Saharan Africans (87–88) and Mediterranean populations or that of C28Y-HFE allele and hematochromatosis in Europeans. Each of these polymorphisms underlies a monogenic trait and hence inference of genetic ancestry can be unnecessary. However, a closer look suggest that cystic fibrosis is also found among groups of Arab and African ancestry (86, 89), similarly, sickle-cell anemia is found in wide range of people including Hispanics and inhabitants of north-western India (35). Furthermore, the blacks of South Africa do not carry sickle cell traits (35) but its occurrence in central Greece has two-fold increase than that of African Americans (35, 90). Therefore, labeling the disease only on the basis of ethnic affiliation or phenotypic occurrence can be wrong interpretation and could possess serious health consequences. The concept of genetic ancestry is a much better indicator than race or ethnicity to determine that whether one carries the marker of a genetic disease. The neutral genetic variation studies have revealed that the proportion of European ancestry in African-Americans averages ~21% (44) while that among South Indian populations is 16% compared to 84% East Asian ancestry (91).

The geographical distribution of genes associated with common diseases is more complicated as they result from complex interactions between genes and environment. There are two schools of thoughts; one view put forward the 'common disease- common variant' (CD/CV) hypothesis, which states that the common genetic diseases are affected by common disease susceptibility alleles (or variants) at a few loci that exist at high frequency across ethnically diverse populations (44, 92). These alleles probably arose before population differentiation and are common across populations. Supporters of CD/CV hypothesis cite that geographic or ethnic clustering variants are mainly a phenomenon of the monogenic disorders. Furthermore, additional support to CD/CV hypothesis is provided by examples of 235T variant of angiotensin (AGT) gene that codes a key component of rennin-angiotensin, blood pressure regulatory pathway. This variant is present across all human groups, as high as 90% among Africans and as low as 30% among Europeans (93). The allele is associated with 20–30% increase risk of developing hypertension (94). Similarly, frequency of the null allele of CYP2D6 gene varies from 6% in Asians to 7% in Africans and up to 30% in Europeans (95). CYP2D6 encodes a member of cytochrome P450 family involved in metabolism of important drugs (13), and its null allele renders the gene product inactive to an extent that homozygous null allele individuals experience little or no analgesic effect. Therefore it has been quoted that although substantial genetic variation is there in etiology of common diseases but it is present in all populations.

The other perspective view is that of multiple rare variant (MRV) hypotheses...
according to which the diseases are associated with substantial proportions of genetic polymorphisms and will probably be specific to groups that experience similar evolutionary forces of selection or drift (96–97). Recently, Bamshad et al, 2004 (98) has shown that a sequencing based analysis of 63,724 SNPs in the coding and regulatory regions of 3931 human genes reveal that large number of private alleles are present in different population groups. This clearly indicates that CD/CV hypothesis was supported because only commonly occurring SNPs have been studied and even if CD/CV hypothesis is correct then also differential effects of risk allele in people with different genetic ancestry have been reported as homozygous APOE4 Asian individuals have ~5 fold higher risk of developing Alzheimer’s disease even when this allele of APOE is frequent in Alzheimer patients of Africa, Asia and Europe. Several polymorphisms in the 5’ cis-regulatory region of CCR5 influence the progression of AIDS and even death in HIV patients (99–100). However, one CCR5 haplotype (HHE) is associated with delayed progression of AIDS in European-Americans but with faster progression African-Americans (101). Similarly, three important variants of CARD 15 or NOD2-R702W, G908R and 1007fs have been associated with an inflammatory bowel disorder- Crohn’s disease in European-Americans (102–103) but not in Europeans or Asians (104). Therefore even if the same risk allele for a complex trait is present in different group, it might be associated with different outcomes. Overall, if more is learned about the genetic bases of the complex diseases and if it is supplemented with the information about the genetic ancestry of the populations then such diseases can be divided into distinct subclasses with similar phenotypes but different underlying genetic bases.

Genetic contribution to differential drug response

The genetic differences among ethnic groups often lead to differences in drug responses. Tate and Goldstein, 2004 (6) have reviewed various genetic variants associated with drug responses and have stated that out of 42 associated genetic variants, more than 30 variants significantly differ between people of Europeans and African ancestry (Table I). An important example is that of receptor polymorphisms i.e. the β2-adrenoceptor which has been extensively studied as a prototype of G-protein coupled receptors, and is the target for bronchodilators drugs used in the treatment of asthma. The β2-receptor gene is known to have 9 single nucleotide polymorphisms (or SNPs) in the coding region; five are degenerate and are unlikely be functional, but four result in the amino acid substitutions within the protein at positions 16, 27, 34 and 164 (105). The functional relevance of these polymorphisms has been investigated by mimicking the polymorphisms by site-directed mutagenesis, expressing the variant receptor in host cells that lack β-receptor expression, and assessing the pharmacologic properties of these cells. Using this approach several polymorphisms have been shown alter β2-receptor function by decreasing receptor-G protein coupling (Thr to He, position 164), increasing receptor desensitization (Arg to Gly, position 16) or decreasing receptor desensitization (Gln to Glu, position 27) (105–106). Although changes in receptor function occur in vitro, there has not been a
<table>
<thead>
<tr>
<th>Category of drugs</th>
<th>Implication</th>
<th>Reasons</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Drugs with an indication of genetic causation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>I-Beta-adrenoceptor blockers</strong></td>
<td>More effective in EA than in AA in treatment of hypertension</td>
<td>Hypertensive and healthy controls carrying two copies of the R389G variant in beta-1-adrenergic receptor have more response to Beta-adrenoceptor blockers</td>
<td>Soffowora et al (105)</td>
</tr>
<tr>
<td>Propranolol</td>
<td>More effective in EA than in AA for systemic hypertension</td>
<td>The R389G variant is more frequent in Europeans than in African-Americans</td>
<td>Johnson et al (106)</td>
</tr>
<tr>
<td>Nadolol</td>
<td>Mean blood pressure reduction less for African-Americans than for European ancestry</td>
<td></td>
<td>Cubettu et al (114)</td>
</tr>
<tr>
<td>Oxprenolol</td>
<td>Mean blood pressure reduction less for African-Americans than for European ancestry</td>
<td></td>
<td>Friedman et al (115)</td>
</tr>
<tr>
<td>Bucindulol</td>
<td>Survival benefit only in non-African Americans</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Drugs whose association has an essential physiological basis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>I-ACE Inhibitors</strong></td>
<td>Less response in AA than in EUR with congestive heart failure</td>
<td>Probably related to lower bio-activity of endogenous nitric oxide</td>
<td>Kalinowski et al (116)</td>
</tr>
<tr>
<td>Enalapril</td>
<td>Response in EUR but not in AA ancestry</td>
<td></td>
<td>Exner et al (117)</td>
</tr>
<tr>
<td>Lisinopril</td>
<td>AA with hypertension required 2-4 times higher dose to obtain similar lowering of blood pressure to EUR ancestry</td>
<td></td>
<td>Weir et al (118)</td>
</tr>
<tr>
<td>Trandolapril</td>
<td>AA with hypertension required 2-4 times higher dose to obtain similar lowering of blood pressure to EUR ancestry</td>
<td></td>
<td>Weir et al (119)</td>
</tr>
<tr>
<td><strong>II-Combination of two vasodilators</strong></td>
<td>Greater efficacy in African Americans than in European ancestry with congestive heart failure</td>
<td></td>
<td>Carson et al (120)</td>
</tr>
<tr>
<td><strong>III-Alpha adrenoreceptor blockers</strong></td>
<td>More effective in EUR ancestry than in AA ancestry for hypertension</td>
<td>Non-adrenergic mechanism contribute to blood pressure maintenance in AA</td>
<td>Cushman et al (121)</td>
</tr>
<tr>
<td>Prazosin</td>
<td>More effective in EUR ancestry than in AA ancestry for hypertension</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IV-Thiazide (Diuretics)</strong></td>
<td>Greater cystolic and diastolic blood pressure responses in AA than in EUR</td>
<td>Probably related to lower bio-activity of endogenous nitric oxide</td>
<td>Kalinowski et al (116)</td>
</tr>
<tr>
<td>Hydrochlorothiazide</td>
<td></td>
<td>Chapman et al (122)</td>
<td></td>
</tr>
<tr>
<td><strong>V-Calcium Channel blocker</strong></td>
<td>More effective in AA than in EUR ancestry for hypertension</td>
<td></td>
<td>Cushman et al (121)</td>
</tr>
<tr>
<td>Dilpiazem</td>
<td></td>
<td>Avic (123)</td>
<td></td>
</tr>
<tr>
<td><strong>VI-Hepatitis anti-viral treatment</strong></td>
<td>AA have lower rate of response to treatment than EUR ancestry</td>
<td>Due to differing immune ability as AA produces more cytokines than EUR</td>
<td>Maier et al (124)</td>
</tr>
<tr>
<td>Ribavirin</td>
<td></td>
<td>Kimball et al (125)</td>
<td></td>
</tr>
<tr>
<td>(Interferon α)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Drugs in which difference is replicated in numerous studies but no inference about physiological basis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>I-Vasodilator antihypertensive</strong></td>
<td>Attenuated response to multiple vasodilators in AA than EUR ancestry</td>
<td>Mechanism not fully under stood</td>
<td>Stein et al (126)</td>
</tr>
<tr>
<td>Sodium nitropruside</td>
<td></td>
<td>Rosenbaum et al (127)</td>
<td></td>
</tr>
<tr>
<td><strong>II-Glucocorticoids</strong></td>
<td>Adverse effect more common in AA than in EUR ancestry</td>
<td>Also altered pharmacokinetics between AA and EU</td>
<td>Tornatoro et al (128)</td>
</tr>
<tr>
<td>Methyl prednisolone</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>III-Anti-Diabetic</strong></td>
<td>Insulin sensitivity significantly lower in Hispanics and African AA than in EUR</td>
<td>Differences remain after adjusting for body fat results well replicated</td>
<td>Goran et al (129)</td>
</tr>
<tr>
<td>Insulin</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

EUR=Europeans; AA=African-America
consistent association of individual SNPs with bronchodilator responsiveness. Rather, complex promoter and coding region haplotypes containing multiple SNPs alter receptor expression and predict \textit{in vivo} responsiveness. The \( \beta\)-1-adrenoreceptor variant (Arg 389) has been reported as associated with increased response to beta-blockers (105–106). Frequency of this variant varies significantly between European Americans –0.723 and African Americans –0.575.

As mentioned earlier, null allele of CYP2D6, a drug-metabolizing enzyme (DME) is reported in a frequency of 10% among north European ancestry and therefore they do not experience an analgesic effect from the prodrug codeine (95). On the contrary, about 98% Arabs are able to transform codeine into the active form morphine (107). CYP2D6 enzyme is a member of the hepatic cytochrome P450 family, and metabolizes 25–30% of all clinically used medications, including antidepressants, antipsychotics, \( \beta\)-blockers, antiarrhythmics and opioid analgesics. The CYP2D6 gene is the most variable of the P450 family, with over 75 different alleles. Variants arise from point mutations, single base-pair deletions and additions, deletion of the entire gene, and gene duplication resulting in two or more copies of the gene (95). The functional consequences include an increase, decrease or loss of enzyme activity that can be correlated with a change in \textit{in vivo} function. Thus, extensive metabolizers (75–85% of the population) are homozygous or heterozygous for the wild-type enzyme or for variants with normal enzyme activity; intermediate metabolizers (10-15%) or poor metabolizers (5-10%) are carriers of two decreased-activities or loss-of-function alleles; and ultra-rapid metabolizers (1–10%) are carriers of duplicate or multiple active genes (95, 107).

Another important issue of using the information of genetic variation is to assess the adverse drug reactions. Use of an antiretroviral drug ‘\textit{Abacavir}’ to treat HIV infections develops a hypersensitivity reaction in about 5% of the people. It has been reported that people carrying HLA-B*5701 allele are associated with hypersensitivity to ‘\textit{Abacavir}’ among people of European American ancestry (108) and not among African Americans (109).

Importance of using the genetic information instead of proxy ethnic or racial ancestry was well documented in the trial of ‘\textit{BilDil}’ drug among African-Americans. ‘\textit{BilDil}’ is a combination drug that combines isosorbide dinorate (a nitric oxide donor) and hydralazine (a vasodilator agent) designed to restore low or depleted nitric oxide levels in the blood to treat or prevent cases of congenital heart failure (110). The trial was conducted due to inefficacy of BilDil in treating congestive heart failure among African-Americans in two ethnically mixed clinical trials (111–112). The latest trial has been recently halted because ‘\textit{BilDil}’ was found highly effective in treating all blacks (113).

Conclusion

Categorical reviewing of the studies carried out on different genetic variants laid emphasis on the fact that the genetic variation between different populations and individuals is a key factor behind pathogenesis of diseases as well as responses and side effect of drugs. However, this
consensus view appears too narrow because of the heterogeneity of results of pharmacognomics studies carried out at different centres on patients of different ethnic descent. This makes it difficult to identify the most potent genetic variant that can be used as a predictive marker for the success of a drug with some notable exception.

Despite the challenges of Pharmacogenomic research, the potential impact of revealing the genetic underpinnings of variable drug response is too significant to ignore and significant strides have already been made in using this new science. DNA diagnostic tests will become available to define a population of patients that are more likely to respond to a drug and be at less risk of a side effect or adverse reaction. Implementing the results of Pharmacogenomic research will revolutionize medicine. Such studies open a plethora of options that can bring into practice like timing and doses of drug therapies like that of immunosuppressive regimens in case of renal or bone marrow graft rejections or the supplementation of folic acids in neural tube defects or type of anti-hypertensive and anti-coagulatory therapy to overcome the cases of coronary heart diseases and thrombofilia. Incorporation of such study will allow an advance anticipation of clinical outcome and drug response and will cause a shift from ‘One treatment fits all’ approach.

ACKNOWLEDGEMENTS

Authors are thankful to Department of Biotechnology (DBT) New Delhi and Sanjay Gandhi Post Graduate Institute of Medical Sciences Lucknow.

REFERENCES


34. Rotimi CN. Are medical and nonmedical uses of large-scale genomic markers conflating genetics and ‘race’? *Nat Genet* 2004; 36: S43–S47.


82. Wyatt MF, Stein BK, Brenton AG. Investigation into accurate mass capability of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, with respect to radical ion species. *J Am Soc Mass Spectrom* 2006; 17: 672–675.


89. Rotimi CN. Are medical and nonmedical uses of large-scale genomic markers conflating genetics and ‘race’? Nat Genet 2004; 36: S43–S47.


111. Cohn JN, Archibald DG, Ziesche S et al. Effect of vasodilator therapy on mortality in chronic congestive heart failure. Results of a Veterans


