Pharmacogenetic applications of the Human Genome project

Although we have known for many years that genetic variations encode characteristics such as hair and eye color, we have only recently come to realize that genetic variations affect almost all aspects of biological function. The sequence and analysis of the human genome provide a map for gaining a better understanding of the genetic basis of human diversity. Variations that occur in greater than 1% of the population are termed polymorphisms. Although deletions and insertions of DNA sequence are infrequent, substitutions of single nucleotides (termed single nucleotide polymorphisms, or SNPs) are the most common type of polymorphism. SNPs can occur in the coding regions of genes (cSNPs), or in non-coding sequence. Due to the degeneracy of the genetic code, some nucleotide substitutions do not represent a change in the encoded amino acid, and are termed ‘synonymous’ cSNPs. Non-synonymous cSNPs can be those that result in conservative substitutions (amino acids with similar sizes or charge) or non-conservative substitutions. Of the 3.5 billion bases of the human genome, there are at least 1.42 million SNPs (ref. 2). These occur most commonly in non-coding regions, but it is estimated that there are at least 60,000 cSNPs, equally represented by synonymous and non-synonymous substitutions.

Studies suggest that SNPs are well tolerated and likely have been retained because they confer survival advantages. However, when combined with other polymorphisms or environmental factors, a SNP may become a low-level disease risk factor or disease modifier. An exciting application of genome variability comes from the fact that SNPs in genes encoding drug targets or drug metabolism pathways can determine the therapeutic utility of pharmacologic agents, a concept known as pharmacogenetics.

Pharmacogenetics is likely to have a significant impact on medicine, analogous to the integration of imaging techniques. Most drugs show significant interindividual variation in therapeutic efficacy. Judgment of the effectiveness and safety of new drugs is still based on the average response of a study group. Inspection of the data from individual subjects, however, usually reveals significant numbers of patients with little or no response, as well as those who have dramatic responses. In cases of complex diseases, this ‘one-drug-fits-all’ attitude subjects patients to empirical trial-and-error periods before acceptable regimens are found. This is often further complicated in diseases where severity of the phenotype waxes and wanes, making it difficult to predict the effects of changing a patient’s medication.

Asthma is a good example of this problem. There are five major classes of asthma drugs, each comprising several different drugs. However, not all patients respond to these drugs, and studies have shown that as many as 50% of patients fail to respond to monotherapy with β2-adrenergic receptor agonists, leukotriene antagonists or inhaled corticosteroids4. It has been estimated that 50% or more of the variability in responses to these drugs can be attributed to genetic factors. Polymorphisms in genes encoding drug targets or metabolic enzymes can positively or negatively affect drug

response. Although many diseases have similar clinical symptoms, genetic variations also underlie differences in disease pathophysiology, rendering drugs aimed at inappropriate pathways ineffective.

Although the potential for individualized therapy—based on a patient’s genetic information—has been accelerated by the availability of the human genome sequence, the polymorphisms associated with drug responsiveness still need to be identified. There are two complimentary approaches for doing this (Fig. 1). One would be to determine the genotype at random SNP loci from a ‘SNP-map’ for each patient (100,000 loci may need to be genotyped). The SNP-map approach relies on SNP-linkage disequilibrium (polymorphisms so close together that they are not separated by recombination events). So, SNP-maps can be used to localize genomic regions where variants that affect drug responses are located. Additional work is then required to identify the actual genes that confer altered drug response from this region. Besides the very large genotyping burden, recent data suggest that this approach is not as reliable as originally believed.

The second approach would be to study polymorphisms in selected genes that are known to be involved in disease and drug response pathways. This also requires obtaining the sequence of the entire gene of interest, which will obviously be facilitated by the completion of the human genome sequence. However, the Human Genome Project revealed that compared with the genomes of other organisms, human genes have short exons and long introns, making the task of identifying protein coding regions or important regulatory regions sometimes difficult.

How many people need to be studied in order to determine the amount of variability in a given gene? The number is actually not very large. We can calculate the number of individuals needed to be screened in order to detect a SNP that occurs at a known frequency. For example, to have a 95% probability of detecting a SNP that occurs at a frequency of 5% in a population requires the screening of 29 individuals. Screening three to four ethnic groups would require genetic testing of about 100 individuals, which is not a major undertaking. The map of human genome sequence variation relied, in part, on sequences derived from a panel of 24 individuals representing an ethnically diverse population. However, the number of samples from some ethnic groups was small. For example, the study included only six samples from African-Americans. With this number of people, 95% probability of detection is found only with polymorphisms occurring at a frequency of 25% or greater. So, less common polymorphisms relatively confined to the African-American population may have been missed.

There is little doubt that it will be important to study genetic variability between different ethnic groups. As an example, a polymorphism in the third intracellular loop of the β2-adrenergic receptor, which is critical for regulation of central and peripheral sympathetic neurotransmission, occurs at a ten-fold higher frequency in African-Americans (~40% allele frequency) as compared with European-Americans. Therefore, we could use genetic analysis to predict that a drug designed to target this receptor may not be effective in a large fraction of African-Americans.

Researchers also debate exactly how much of each gene should be analyzed for variations. It seems reasonable to search for polymorphisms in the entire coding region, as well as the immediate 5’ upstream region (containing the promoter sequence) and the 3’ untranslated region, which both contain elements that regulate expression. Intron/exon boundaries may also have regulatory sequences and should be analyzed. It is necessary to examine multiple regions of a gene, as many genes contain multiple SNPs that can act in combination to influence phenotype.

This concept is best illustrated by a recent study of the β2-adrenergic receptor, expressed in the heart and lungs, which is involved in the pathophysiology and treatment of asthma and congestive heart failure. The gene encoding this receptor contains three non-synonymous SNPs that cause non-conservative amino acid substitutions and affect receptor function, agonist regulation or binding affinity. The gene also contains multiple SNPs in the promoter region that affect receptor expression. In this case, knowing the context of the given SNP, as it occurs in relation to others, offers the best opportunity for predicting drug response. A patient’s haplotype (the combination of SNPs, as arranged on each parental chromosome) offers even greater predictive ability. Several recent reports have shown that certain β2-adrenergic receptor haplotypes correlate with asthmatic phenotypes or responses to β2-adrenergic receptor agonist therapy in instances where individual SNPs do not.

A combination of in vitro studies, animal model experiments and clinical trials is required to fully understand the basis of the effects of various polymorphisms on drug response (Fig. 1). Soon, the process of SNP discovery will be so acceler-
Gazing into a crystal ball—cancer therapy in the post-genomic era

The sequencing of the human genome is likely to speed the discovery of factors involved in cancer pathogenesis and lead to an age of individually tailored anti-cancer drugs. But does the ability to obtain an abundance of genetic information mean that we necessarily know how to use it?

The current therapeutic strategies for cancer include surgery, which can be successful but is often incomplete, and medical approaches such as radiation, chemotherapy and immunotherapy, which kill tumor cells but also cause major damage to normal tissues. As the root causes of cancer lie in genomic abnormalities, the sequencing and analysis of the human genome is likely to help us to overcome these obstacles, revolutionizing cancer diagnosis and treatment.

Right now, it is impossible to predict what cancer researchers will learn from the completed human genome sequence. If we speculate, however, we should bear in mind some basic principles of prognostication, written by Robert X. Cringely. These are: 1) We tend to overestimate the amount of change that will take place in the short term; 2) We tend to underestimate the amount of change that will take place in the long term; 3) The more specific a prediction, the less likely it is to be correct; 4) Past performance is a predictor of future results but not a good one; and 5) The most reliable predictions are those that follow established trends.

In visualizing the long term, the publica-

tion of the completed human genome se-
quence is a necessary first step in optimizing treatment of cancer and other human diseases. Someday, it may be possible to use genetic information to determine which patients will respond or have adverse responses to different anti-cancer drugs, and to specifically tailor their therapy. However much more work beyond the initial sequencing project itself must be completed before this day arrives.

Although an abundance of valuable information will come out of the human genome project, we must remember that the study reports the genomic sequences of only a few volunteers. The ability to design individualized cancer therapies is likely to require DNA sequence information from patients as well as their tumors, requiring individualized genome sequencing. But is this a realistic possibility? Referring back to Cringely’s rules, we would argue that rule #2, which states that we tend to underestimate change in the long term, applies to those who say that this day will never come. In considering this possibility, we need to consider the current pace of genomic technological progress, and today’s cost of sequencing a person’s entire genome. It was recently announced that the Mouse Sequencing Consortium received $58 million to sequence the C57BL/6J mouse. Using this as a guide, the cost of sequencing all 46 chromosomes from a single person would be around $120 million.

How long will it take to reduce sequencing costs to the point where complete genome sequencing could become routine? Sequencing rates and costs have been changing by a factor of two annually. So if we extrapolate forward by 18 years, and if we can continue at our current pace (following rule #5—the most reliable predictions are those that follow established trends), the cost of sequencing a random human would be decreased by a factor of about 250,000, yielding a cost of about $500 (in year 2000 dollars) per person, a reasonable price for a medical procedure. However, we probably also need to apply rule #4, and assume that the rate of change will not be as profound as we expect.

Useful information will also be derived

Mark J. Ratain*3 & Mary V. Relling*3

*Department of Medicine and Pharmacology University of Cincinnati College of Medicine, Cincinnati, Ohio, USA
Email: Stephen.Liggett@UC.Edu


© 2001 Nature Publishing Group http://medicine.nature.com

© 2001 Nature Publishing Group http://medicine.nature.com