

The Role of Single Nucleotide Polymorphisms (SNPs) in Understanding Complex Disorders and Pharmacogenomics

S L Kao*, S S Chong,**^{PhD, DipABMG}, C G L Lee,***^{PhD}

Abstract

Introduction: In the last two years, there has been an increasing interest in single nucleotide polymorphisms (SNPs). They have been hailed as the most common polymorphism found in the human genome and are believed to be responsible for 90% of all inter-individual variation. Efforts are now directed at the large-scale identification and archiving of SNPs in the human genome. Not only are they useful markers for population divergence studies, SNPs can be utilised as markers in studies of complex diseases and pharmacogenomics. **Methods:** Traditional methods for identifying SNPs, as well as methods for large-scale detection and genotyping of SNPs currently being developed, are briefly discussed in this review. Such developments will facilitate and enhance the process of identifying and characterising genes and their functions. **Results:** The utility of SNPs in identifying genes contributing to pharmacogenetic variation and increased risk of a complex disease is discussed. The role of SNPs in influencing drug response in different individuals is also presented. **Conclusions:** In helping to unravel the genetic basis of complex diseases and inter-individual variation in drug response, SNPs will catalyse the transition into a new age of medicine in which medical care is tailored to the individual's genetic profile.

Ann Acad Med Singapore 2000; 29:376-82

Key words: Complex diseases, Drug response, Gene discovery, Pharmacogenetics, Single nucleotide polymorphism

Introduction

With the Human Genome Project well underway and targeted to be completed by 2001, the full set of nucleotide sequences of a reference human genome will soon be available. The realisation now is that there is a need to explore and characterise inter-individual variation. The first peeks into human genome diversity were in the form of studies that examined protein variants. These were followed by studies of DNA sequence variation of specific genes in different individuals. Over the last two decades, we have heard much about restriction fragment length polymorphisms and micro-satellites. These were popularly used as genetic markers for linkage studies. Recently, we have come to realise that 90% of inter-individual variation consists of what are known to be single nucleotide polymorphisms or SNPs (pronounced as "snips").¹

"Single nucleotide polymorphism" is a relatively new term, popularised only in the last two years. One of the first suggestions of the presence of such polymorphisms was

made in a study of 75,000bp (49 genes) of DNA sequences in 5' untranslated regions (UTRs), 3' UTRs and coding regions.² In 1994, a comprehensive study of sequence diversity at the β -globin gene locus was conducted.³ Later in 1998, 79 SNPs were found from the sequencing of the human lipoprotein lipase gene (exons 4 through 9) from 71 unrelated individuals of European, African and Finnish ancestry.⁴ Other studies that searched for SNPs followed⁵ and by 1999, it was believed that there are multiple SNPs distributed throughout the entire human genome.

As of April 1999, there were only 7000 of such polymorphisms available in public databases, but this is all set to change. In December 1997, the National Human Genome Research Institute launched a government-funded effort to create a database of 60,000 to 160,000 SNPs.⁶ The count as of 20 March 2000 is 26,569 SNPs (<http://www.ncbi.nlm.nih.gov/SNP/>). In April 1999, ten major pharmaceutical companies and the Wellcome Trust came together to set up a non-profit SNP consortium aimed at

* Medical III Student

Faculty of Medicine, National University of Singapore

** Senior Lecturer

Department of Paediatrics, Faculty of Medicine, National University of Singapore and
Johns Hopkins School of Medicine, Baltimore, Maryland, USA

*** Assistant Professor

Department of Biochemistry, Faculty of Medicine, National University of Singapore and
Department of Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, USA

Address for Reprints: Dr Caroline G Lee, Department of Biochemistry, Faculty of Medicine, National University of Singapore, MD7, #05-10, 8 Medical Drive, Singapore 117597.

finding, mapping and archiving 300,000 SNPs to be released to the public. Various recent reviews highlighting useful web resources in this field have been published.^{7,8}

What are SNPs?

Single nucleotide polymorphisms are single base-pair positions in genomic DNA that differ from one normal individual to another. Although rare genetic variants exist in the human, the least frequent allele should have a frequency of 1% or greater in the population to satisfy as a polymorphism.^{9,10} This definition of SNPs does not include disease-causing single gene mutations and single gene insertion or deletion variants. Theoretically, there can be a total of 4 alleles of this polymorphism in the population since there can be one transition or two transversions at each nucleotide position. For example, there can be a C to T transition, or a C to A or G transversion. However, the SNPs characterised so far have been bi-allelic in their population distribution.⁸

SNPs are the most abundant type of sequence variation in the human genome. It is estimated that there is about one SNP per 1000 bp in the human genome, giving about 3 million SNPs in the 3 billion bp long human genome. However, SNPs are not equally distributed across the entire genome.¹¹ Some regions have a frequency much lower than the average 0.1%¹² while other regions, like the non-coding HLAs, have frequencies as high as 10%.¹³

SNPs can be located in 3 general areas in the genome. They may be found in coding regions of genes. These are known as cSNPs. It is estimated that there are 340,000 to 400,000 cSNPs in the human genome and 40% of these will be non-synonymous polymorphisms, resulting in a different amino acid coded.⁵ SNPs can also be located in the non-coding regions around the immediate vicinity of genes. These known as non-coding SNPs or nSNPs and they can be located either in the 5'/3' untranslated regions or in introns of genes. cSNPs and nSNPs occur at about the same frequency. SNPs are also found in intergenic "junk" DNA and these are known as random non-coding SNPs or rSNPs. These are believed to be the most common type of SNPs¹⁴ possibly because there is less selective pressure against these polymorphisms as they do not alter gene expression.

Will SNPs Have Any Impact Upon the Practice of Medicine?

One application of SNPs is in the area of population studies where they can be used to study divergence of ethnic groups and population migrations. Mitochondrial nucleotide sequences have been utilised widely in this respect.¹⁵ It is likely that SNPs in nuclear DNA will prove to be useful as well, given its abundance in the human genome. In fact, studies using SNPs to trace population

divergence have already been conducted.¹⁶

In this age of rapid advancements in medical genetics, there has been a lot of speculation and excitement over the development of personalised medicine in which disease prevention, diagnosis and management are tailored to the individual's genetic profile. What has been lacking for many years is a suitable tool to aid in the otherwise tedious methods of gene discovery. With the advent of SNP technology, this is set to change. In this review, we will discuss the impact SNPs are making in the areas of complex disease research and pharmacogenomics.

Complex Diseases

Compared to diseases caused by single gene defects (e.g. cystic fibrosis and sickle cell anaemia), the aetiology and pathogenesis of complex diseases are, as its name suggests, more complicated. No single genetic defect can be isolated in the case of complex diseases. Instead, multiple gene loci interact with one another and with environmental factors to produce disease in the individual. Each of these loci does not cause the disease per se but is considered to modify the risk of developing the disease. The number of interacting genes contributing to a complex disorder can range from a few to a few hundred. In addition, there can be allelic heterogeneity at each locus, resulting in different forms of the gene existing at the same locus and contributing to the genesis of the disease to different degrees. Examples of complex diseases include Alzheimer's disease, common cancers, hypertension, diabetes mellitus and non-syndromic oral clefting. As can be immediately appreciated, complex diseases are considerably more prevalent than diseases due to single gene defects.

One of the more common methods employed in the study of diseases caused by single gene defects is known as positional cloning. This method is based upon the principle that two loci situated close to each other will be inherited together. Identifying polymorphic markers with alleles that cosegregate tightly with the disease in pedigrees allows for the mapping of the disease-causing gene relative to the markers. Theoretically, linkage mapping can also be used to localise genes involved in complex diseases. However, this strategy is technically impractical if there are numerous loci involved, with each locus contributing only a small amount to disease development. With genes that play a minor role in disease predisposition, large numbers of families will need to be studied before a statistically significant linkage relationship can be established between the gene and its marker. This number can easily run into the thousands.¹⁷ The alternative strategy is to conduct association analysis. The basis for such a study is that if a gene is a risk factor for a disease, it should be found in higher frequencies in diseased individuals (cases) than in non-diseased individuals (controls).

There are essentially two approaches employed in association analysis (Fig. 1). The first is known as a linkage disequilibrium-based association analysis (Fig. 1A). The hypothesis is that there is a specific array of polymorphisms known as ancestral haplotypes that flank disease-associated genes in a population.⁸ Linkage disequilibrium-based association analyses involve genome-wide scanning of cases and controls with a large number of markers distributed across the genome to determine if any combination of markers are particularly associated with the disease. There are two main requirements for this method to work. Firstly, there must be a large number of markers distributed across the genome available. Secondly, due to the large number of markers involved in such studies, there must be available technology to assay these markers efficiently and cost-effectively. SNPs serve as suitable markers for such linkage-disequilibrium association analyses studies. They are the most abundant polymorphism in the human genome, thereby providing large sets of markers near or in any locus of interest. They are also more stably inherited from generation to generation than presently used markers like micro-satellites. In addition, present assaying techniques allow for rapid genotyping of SNPs and the development of even cheaper and larger scale techniques are underway.^{18,19}

The disadvantage of linkage disequilibrium-based methods is that a careful selection of population for the study needs to be undertaken because linkage disequilibrium

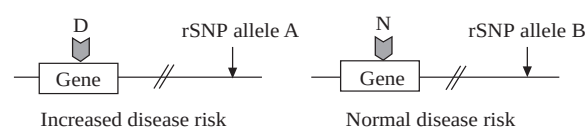
cannot be assured of except in recently expanded populations. As the number of generations increases, so does the chance of recombination occurring within ancestral haplotypes. Some also feel that the genome-wide scanning method may involve too many markers and individuals to be time- and cost-effective to study. This is especially so if the effect of each locus on the disease is small. Therefore, a more efficient type of association study known as a direct association study has been proposed.

The direct association analysis (Fig. 1B) is based on the hypothesis that a given allele of a polymorphism might have pathogenic effects. This is known as the “common disease-common variant” hypothesis. The hypothesis is that if uncommon variants cause rare Mendelian diseases, then it might be likely that common complex diseases are due to common variants.^{5,11} This method involves the studying of polymorphisms in candidate disease-associated genes and testing them for association with disease. It is believed that cSNPs (especially those which are non-synonymous) in candidate genes as well as nSNPs in their regulatory regions are likely to have a pathogenic effect. At the moment, direct association analyses using non-synonymous cSNPs are the more practical approach in the study of complex diseases. In a study on SNPs in coding regions of human genes, SNPs were found in 106 candidate genes relevant to cardiovascular diseases, endocrinology and neuropsychiatry. It is believed that SNPs that affect amino acid sequences in these candidate genes are a likely explanation for the variation in such diseases.⁵ It is also very likely that nSNPs that affect gene regulation and splicing can also contribute to disease risk. However, at the moment, these are harder to characterise due to difficulty in distinguishing them from other non-coding SNPs in intergenic junk.

Identifying the genes that contribute to these complex diseases will aid in the elucidation of molecular mechanisms of gene-gene and gene-environment interaction that lead to pathogenesis. With this is the hope of reducing morbidity and mortality through the design of specific prevention measures and therapy.

There has already been reported success on the use of SNPs in the area of complex diseases. A recent example is the clarification of mechanisms and risk factors in the area of diet and prevention of cancer with the help of SNPs.²⁰ In another study, 75 genes that encoded growth factors, receptors, channels, and other proteins suspected to regulate blood pressure were investigated for SNPs in individuals from the top and bottom 25th percentile of a normalised blood-pressure distribution curve. 874 candidate SNPs were identified in about 190 kb worth of sequences.²¹ SNPs in the myostatin gene have been found and are also being used in the investigation of loss of muscle mass with ageing

(A) Linkage disequilibrium-based association analysis



(B) Direct association analysis

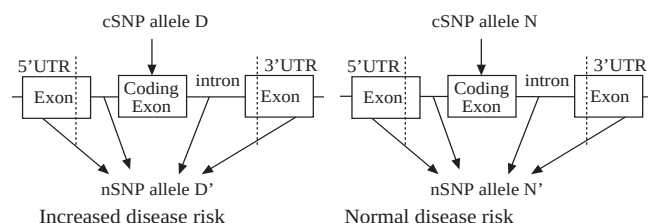


Fig. 1. Approaches to association analyses. (A) Linkage disequilibrium (LD)-based association analysis. D represents disease-associated sequence while N represents normal sequence. This approach relies on random SNP alleles (represented by rSNP allele A or B), which may span over a large genomic region, to exhibit linkage disequilibrium with the disease. LD-based association analyses test for presence of a higher frequency of rSNP A in affected individuals. (B) Direct association analysis. Here the SNPs occur in the candidate disease-associated genes. The SNP can be found either in the coding regions (represented by cSNP allele D or N) or in the non-coding regions (represented by nSNP allele D' or N'). Direct association analyses test for presence of a higher frequency of cSNP D or nSNP D' in affected individuals.

and inter-individual muscle mass differences.²² SNPs related to plasma von Willebrand factor levels have also been characterised.²³ It is anticipated that complex disease research will grow in line with the rapid advancement of SNP technology.

Pharmacogenomics

It was noted by Pythagoras in 510 BC that there was a group of people who developed adverse reactions when they consumed fava beans. We know today that these individuals were deficient in glucose-6 phosphate dehydrogenase and they suffered haemolytic anaemia upon consumption of fava beans.

This could well have been one of the first documented observations in pharmacogenetics—the study of variability in drug responses between individuals due to heredity. A term that has often been used interchangeably with “pharmacogenetics” is “pharmacogenomics”. While the term “pharmacogenetics” was coined in the late 1950s, the concept of pharmacogenomics was only introduced in the last decade. Pharmacogenomics is a rapidly developing field in which the development of drugs is based upon newly discovered genes with the sequencing of the human genome. In the not-so-distant future, it will be possible to tailor each individual’s drug regimen to his genetic profile so that drug therapy is delivered at the optimum dose, with maximal efficacy and minimal adverse effects.

The differences in drug response between individuals can be due to pharmacokinetic as well as pharmacodynamic variation. So far, the majority of pharmacogenetic variation has been discovered to be due to polymorphisms in drug metabolising enzymes (DMEs). One classic example would be the genetic variation in N-acetyl transferase that leads to slow metabolism of isoniazid with resultant peripheral neuropathy.¹⁰ The remaining variations are contributed by differences in receptor affinity, drug transporters, drug uptake and excretion processes. However, these are poorly characterised at the moment¹⁴ but represent an emerging field that is generating great interest.

The study of SNPs will contribute to the advancement of pharmacogenomics in two main ways. Firstly, SNPs can function as genetic markers in the search for genes responsible for pharmacogenetic variation. Secondly, these polymorphisms can also directly contribute by causing functional changes in gene products related to drug response.

Just as SNPs can be used as genetic markers to localise disease-associated genes, they can also be used to identify genes responsible for pharmacogenetic variation. With the characterisation of SNP-based haplotypes associated with certain adverse drug reactions or drug insensitivity, a person’s drug regimen can then be adjusted to meet his individual needs by just simply typing his SNP profile.

Given the abundance of SNPs in the human genome, it is highly likely that some of these may lead to functional changes in gene products involved in drug response. The large number of SNPs also suggests that some of the yet uncharacterised variation in drug receptors, transporters etc. may be encoded in these polymorphisms.¹⁴ Some of these SNPs may be found in coding regions and cause functional gene products by altering the amino acid sequence. They may be non coding SNPs in 5’ and 3’ UTRs that lead to altered mRNA stability or modified transcriptional activity of a gene. This area of research proves to be promising. For example, there is strong evidence to believe that SNPs may contribute to substrate specificity of the *MDR1* multidrug transporter.

The ATP-binding cassette (ABC) superfamily of multidrug transporters exemplified by P-glycoprotein or *MDR1* and includes MRP1, MRP2/cMOAT, MRP3, MRP4 and MRP5 represent important determinants of drug pharmacokinetics and bioavailability.^{24,25} Although interest in the *MDR1* multidrug transporter has mainly centred in the field of cancer chemotherapeutic resistance, *MDR1* has a more wide-ranging effect on transport of other common drugs. In addition to transporting anticancer drugs, it was found to transport a wide variety of different drugs including anti-arrhythmics, antidepressants, antipsychotics and even HIV-1 protease inhibitors (Table I).

Unlike DME, whose major drug metabolising function occurs in the liver, the *MDR1* transporter is expressed at the interface of major organs. This pattern of distribution is suggestive that the *MDR1* transporter may offer two levels of protection for the body against drug toxicity. Its expression at the apical surface of mucosal cells in the small and large intestines; in the brush border of renal proximal tubules and at the biliary surface of hepatocytes²⁶⁻²⁹ provide a first line of defense to prevent absorption of xenobiotics into the body. The capillary endothelial cells of the brain, testis and placenta²⁶⁻²⁹ as well as hematopoietic stem cells³⁰ also contain the *MDR1* transporter and provide a second level of defense to keep cytotoxic products out of sensitive organs. For example, in normal mice expressing *mdr* in the brain, the anti-diarrhoeal narcotic analogue and *mdr* substrate, loperamide, which is unable to cross the blood-brain barrier, is effective in treating the peripheral symptoms of diarrhoea without any CNS side effects. However, in *mdr* ^{-/-} knockout mice, this drug was found to enter the brain and the mice displayed characteristic behavioural abnormalities as if they have received an opiate, e.g. morphine, which enters the central nervous system.³¹

Of greatest relevance to the identification of SNPs in *MDR1* is the experimental observation that single nucleotide changes in the coding region of the *MDR1* transporter

TABLE I: AGENTS INTERACTING WITH THE MULTIDRUG TRANSPORTER

Anticancer drugs	Other drugs
Vinca alkaloids e.g. vinblastine	Calcium channel blockers e.g. verapamil
Anthracyclines e.g. doxorubicin	Anti-arrhythmics e.g. quinidine
Epipodophyllotoxins e.g. etoposide	Antihypertensives e.g. reserpine
Antibiotics e.g. antimycin D	Antibiotics e.g. hydrophobic cephalosporins
Others e.g. taxol, mitomycin C	Antihistamines e.g. terfenadine
	Immunosuppressants e.g. cyclosporin A
	Steroid hormones e.g. progesterone
	Modified steroids e.g. tamoxifen
	Lipophilic cations e.g. tetraphenylphosphonium
	Diterpenes e.g. forskolins
	Detergents e.g. Tween 80
	Antidepressants e.g. tioperidone
	Antipsychotics e.g. phenothiazines
	HIV-1 protease inhibitors e.g. ritonavir
	Other hydrophobic drugs and their analogs
Other cytotoxic drugs	
Antimicrotubule drugs e.g. colchicine	
Protein synthesis inhibitors e.g. puromycin	
Toxic peptides e.g. valinomycin	
DNA intercalators e.g. EtBr	

changes its profile of substrate specificity. For example, a nucleotide change at the 185 amino acid position resulting in a change from glycine to valine (G185V) led to increased resistance to colchicine but not other drugs.³² Hence, it would be worthwhile to examine the genetic variation in the population and correlate SNPs identified with drug transport.

The study of SNPs may well decrypt the yet unsolved mysteries of inter-individual drug response variation. Once we understand the genetic basis of inter-individual variation, with the development of rapid genotyping methods, we may soon be able to administer individualised drug regimens according to the patient's genetic profile.

Developing SNP Markers

The first step in the basic approach towards SNP discovery is obtaining the DNA sequence surrounding the polymorphism. A polymerase chain reaction (PCR) is designed to amplify the segment of the DNA from the genomic DNA of different individuals. SNPs in these segments are identified by a variety of methods. These can be classified into small-scale gel-based methods and the large-scale chip-technology-based methods.

At one end is the traditional method of DNA sequencing followed by multi-alignment of sequences. Newer gel-based methods utilise the differences in hybridisation stability of perfectly matched and mismatched sequences of DNA. By denaturing and reannealing a mixture of PCR amplified products, a hybrid between two different alleles will be formed if there is a nucleotide polymorphism within the length of DNA amplified. Because this reannealed hybrid is not 100% complementary, it would exhibit less stability than a fully complementary double-stranded DNA. Therefore, the hybrid can be detected using a Denaturing Gradient Gel Electrophoresis or a Temperature Gradient

Gel Electrophoresis. Another gel-based method called Single Strand Conformation Polymorphism is based on the principle that a denatured single-stranded DNA will assume different secondary structures if it has a different nucleotide sequence. Hence, different alleles of SNPs will form different secondary structures that can be differentiated due to the different speeds of migration during electrophoresis. Restriction enzyme digestion has also been utilised to differentiate alleles when a certain allele of an SNP alters a restriction enzyme site. PCR reactions that make use of allele-specific primers have also been used.

The methods mentioned above are largely for small-scale detection of SNPs. With the increasing need for large-scale screenings for SNPs, more efficient high-technology methods have been developed. A high-throughput strategy for SNP screening utilises "DNA chips". These chips are mainly glass slides containing arrays of specific oligonucleotide probes covalently bound to their surfaces. Biotin-labelled DNA sequences amplified by PCR are annealed to these oligonucleotides and differences in strengths of hybridisation detected by confocal microscopy. This method is known as an Array Hybridisation Assay. Recently, 3200 candidate SNPs have been reported in 2.4 Mb of DNA using this technology.¹¹ Another highly efficient strategy exploits real-time PCR and involves hybridisation-based allele-discrimination during the process of PCR.¹⁹ Two main methods based on this principle have been developed. The Taqman™ assay utilises allele-specific probes known as Taqman™ probes that are conjugated to donor-acceptor dye pairs that interact via fluorescence resonance energy transfer (FRET). Since the donor is in close proximity to the acceptor in an intact Taqman™ probe, fluorescence is quenched. This assay exploits the 5' nuclease activity of Taq DNA polymerase to digest away the specific Taqman™ probe that has hybridised to specific

amplification products during PCR amplification. This cleavage of the Taqman™ probe will dissociate the donor from its acceptor and fluorescence will increase. By labelling different Taqman™ probes representing different alleles with different fluorophores, it will be possible to discriminate different alleles. Another alternative procedure that also makes use of FRET involves molecular beacons. These beacons are oligonucleotide probes conjugated with a pair of fluorophores. If a probe fails to hybridise with the PCR amplification product, it will adopt a hairpin-loop conformation that bring the two fluorophores together and the fluorescence from each will be quenched due to FRET.

Once an SNP is identified, it needs to be mapped to a certain position in the human genome before it can be used as a marker. Once it is mapped, a genotyping assay is developed to genotype large numbers of individuals from the population. A wide variety of strategies have been employed for large scale, high throughput assaying of SNPs. Most of these methods being developed make use of “chip” technology.

Various strategies have been introduced to cut down the amount of work required in SNP discovery by skipping some of the steps in the SNP marker development process. Utilising sequence tagged sites to screen for SNPs gives the advantage that their sequences are already known, mapped and the PCR reactions optimised. By comparing overlapping BAC clones and PACs obtained from the Human Genome Project, the steps of sequencing, finding the SNP and mapping it can be skipped.

Conclusion

By the end of year 2000, more than 500,000 SNPs in the human genome will be available to the public for free usage in research. This number is expected to increase rapidly with the advancement and maturation of SNP discovery methods. This would also mean that the disease-gene discovery process and the development of pharmacogenomics will also be accelerated. It is clear that the discovery and utilisation of SNPs will catalyse the approach of a new era of personalised medicine.

REFERENCES

- Collins F S, Brooks L D, Chakravarti A. A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res* 1998; 8:1229-31.
- Li W H, Sadler L A. Low nucleotide diversity in man. *Genetics* 1991; 129:513-23.
- Fullerton S M, Harding R M, Boyce A J, Clegg J B. Molecular and population genetic analysis of allelic sequence diversity at the human beta-globin locus. *Proc Natl Acad Sci U S A* 1994; 91:1805-9.
- Clark A G, Weiss K M, Nickerson D A, Taylor S L, Buchanan A, Stengard J, et al. Haplotype structure and population genetic inferences from nucleotide-sequence variation in human lipoprotein lipase. *Am J Hum Genet* 1998; 63:595-612.
- Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, et al. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat Genet* 1999; 22:231-8.
- Marshall E. Drug firms to create public database of genetic mutations. *Science* 1999; 284:406-7.
- Kwok P Y, Gu Z. Single nucleotide polymorphism libraries: why and how are we building them? *Mol Med Today* 1999; 5:538-43.
- Brookes A J. The essence of SNPs. *Gene* 1999; 234:1771-86.
- Lewin B. Isolating the gene. In: Lewin B, editor. *Genes VI*. Chap 6. New York: Oxford Univ Press, 1997:123.
- Nebert D W. Polymorphisms in drug-metabolizing enzymes: what is their clinical relevance and why do they exist? *Am J Hum Genet* 1997; 60:265-71.
- Wang D G, Fan J B, Siao C J, Berno A, Young P, Sapolsky R, et al. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 1998; 280:1077-82.
- Nachman M W, Bauer V L, Crowell S L, Aquadro C F. DNA variability and recombination rates at X-linked loci in humans. *Genetics* 1998; 150:1133-41.
- Guillaudeux T, Janer M, Wong G K, Spies T, Geraghty D E. The complete genomic sequence of 424,015 bp at the centromeric end of the HLA class I region: gene content and polymorphism. *Proc Natl Acad Sci U S A* 1998; 95:9494-9.
- Nebert D W. Pharmacogenetics and pharmacogenomics: why is this relevant to the clinical geneticist? *Clin Genet* 1999; 56:247-58.
- van Holst Pellekaan S, Frommer M, Sved J, Boettcher B. Mitochondrial control-region sequence variation in aboriginal Australians. *Am J Hum Genet* 1998; 62:435-49.
- Zietkiewicz E, Yotova V, Jarnik M, Korab-Laskowska M, Kidd K K, Modiano D, et al. Nuclear DNA diversity in worldwide distributed human populations. *Gene* 1997; 205:161-71.
- Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996; 273:1516-7.
- Taillon-Miller P, Piernot E E, Kwok P Y. Efficient approach to unique single-nucleotide polymorphism discovery. *Genome Res* 1999; 9:499-505.
- Landegren U, Nilsson M, Kwok P Y. Reading bits of genetic information: methods for single-nucleotide polymorphism analysis. *Genome Res* 1998; 8:769-76.
- Ames B N. Cancer prevention and diet: help from single nucleotide polymorphisms. *Proc Natl Acad Sci U S A* 1999; 96:12216-8.
- Halushka M K, Fan J B, Bentley K, Hsie L, Shen N, Weder A, et al. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat Genet* 1999; 22:239-47.
- Ferrell R E, Conte V, Lawrence E C, Roth S M, Hagberg J M, Hurley B F. Frequent sequence variation in the human myostatin (GDF8) gene as a marker for analysis of muscle-related phenotypes. *Genomics* 1999; 62:203-7.
- Keightley A M, Lam Y M, Brady J N, Cameron C L, Lillicrap D. Variation at the von Willebrand factor (vWF) gene locus is associated with plasma vWF:Ag levels: identification of three novel single nucleotide polymorphisms in the vWF gene promoter. *Blood* 1999; 93:4277-83.
- Lee C G, Gottesman M M, Cardarelli C O, Ramachandra M, Jeang K T, Ambudkar S V, et al. HIV-1 protease inhibitors are substrates for the MDR1 multidrug transporter. *Biochemistry* 1998; 37:3594-601.
- Lee C G, Gottesman M M. HIV-1 protease inhibitors and the MDR1 multidrug transporter. *J Clin Invest* 1998; 101:287-8.

26. Cordon-Cardo C, O'Brien J P, Casals D, Rittman-Grauer L, Biedler J, Melamed M R, et al. Multidrug-resistance gene (P-glycoprotein) is expressed by endothelial cells at blood-brain barrier sites. *Proc Natl Acad Sci U S A* 1989; 86:695-8.
 27. Cordon-Cardo C, O'Brien J P, Boccia J, Casals D, Bertino J R, Melamed M R. Expression of the multidrug resistance gene product (P-glycoprotein) in human normal and tumor tissues. *J Histochem Cytochem* 1990; 38:1277-87.
 28. Thiebaut F, Tsuruo T, Hamada H, Gottesman M M, Pastan I, Willingham M C. Immunohistochemical localization in normal tissues of different epitopes in the multidrug transport protein P170: evidence for localisation in brain capillaries and crossreactivity of one antibody with a muscle protein. *J Histochem Cytochem* 1989; 37:159-64.
 29. Thiebaut F, Tsuruo T, Hamada H, Gottesman M M, Pastan I, Willingham M C. Cellular localisation of the multidrug-resistance gene product P-glycoprotein in normal human tissues. *Proc Natl Acad Sci U S A* 1987; 84:7735-8.
 30. Chaudhary P M, Roninson I B. Expression and activity of P-glycoprotein, a multidrug efflux pump, in human hematopoietic stem cells. *Cell* 1991; 66:85-94.
 31. Schinkel A H, Wagenaar E, Mol C A, van Deemter L. P-glycoprotein in the blood-brain barrier of mice influences the brain penetration and pharmacological activity of many drugs. *J Clin Invest* 1996; 97: 2517-24.
 32. Cardarelli C O, Aksentijevich I, Pastan I, Gottesman M M. Differential effects of P-glycoprotein inhibitors on NIH3T3 cells transfected with wild-type (G185) or mutant (V185) multidrug transporters. *Cancer Res* 1995; 55:1086-91.
-