Role of Genomics in Identifying New Targets for Cancer Therapy

By Sarah L. Anzick, PhD [3] and Jeffrey M. Trent, PhD [4]

The detailed map of the human genome can potentially transform future cancer therapy by merging genomics with pharmacology, thereby identifying which patients will benefit from specific therapeutic agents. Single-nucleotide polymorphisms (SNPs) provide a valuable tool for this pharmacogenetic approach to cancer therapy.

The recently announced completion of the "working draft" of the human genome sequence is an important milestone in medical science.[1,2] Identifying all the genes and their regulatory regions provides the essential framework for the genetic blueprint of humankind and will facilitate an understanding of the molecular basis of disease. In turn, the postsequence era has brought forth a new foundation for a broad range of genomic tools that can be applied to medical science, which will ultimately change the practice of modern medicine.

Numerous lines of evidence have demonstrated that the risk of developing certain disorders and the metabolism of drugs used to treat these conditions are influenced by one’s genetic makeup.[3] With a detailed structure of the human genome at hand, the future of cancer therapy involves a merging of genomics with pharmacology, in which underlying genomic variations can be used to predict both the efficacy and toxicity of therapeutic agents. This pharmacogenetic approach will likely result in vastly improved patient care.

In this article, we describe how the most abundant class of genetic variants in the human genome, termed single-nucleotide polymorphisms (SNPs), provides a valuable tool for pharmacogenetics in cancer therapy. We also highlight two high-throughput technologies, DNA and tissue microarrays, which have the potential to significantly augment the field of pharmacogenomics, defined by Roses as "the determination and analysis of the genome (DNA) and its products (RNA and proteins) as they relate to drug response."[4]

The application of technologies such as SNP analysis, DNA microarray, and tissue microarray analysis will undoubtedly revolutionize cancer therapy. It will soon be possible to identify patients who respond or fail to respond to treatment early in the clinical drug development process. This information would provide a significant step towards "individualizing" cancer therapy and maximizing the benefits of treatment by tailoring patient therapy.

The human genome, composed of approximately 3 billion base pairs of DNA, is commonly referred to as the "book of life." Chapters of this book represent individual chromosomes, the sentences represent genes, and the words are codons made up of the DNA bases, adenine, cytosine, thymine, and guanine. It is estimated that approximately 99.9% of the genetic makeup of all individuals is identical, leaving genomic sequence variance to less than a fraction of 1% (0.01% or about 3 million bases). Though seemingly negligible, this 0.01% difference is significant indeed, because a single base change/mutation can cause clinical disease. An individual is estimated to carry approximately 300 to 1,200 deleterious mutations.[5]

In addition to deleterious mutations, silent base pair changes (ie, changes that result in no apparent effect in an individual) seem to occur throughout the genome, with an average frequency of 1 per 1,000 to 2,000 bases.[6-8] Single base pair differences that occur when the DNA sequences of individuals are compared are SNPs (Figure 1). Intuitively, a high-density, genome-wide map of all these SNPs would help to create a fingerprint of the polymorphic variants in each individual and would have significant implications for disease gene discovery, diagnosis, and treatment.

Identifying and cataloguing these sequence variations to create a high-density SNP map of the entire human genome are the primary goals of The SNP Consortium and the Human Genome Project (see http://www.ncbi.nlm.nih.gov/SNP/).[9] Recently, a map was published of 1.42 million SNPs distributed...
throughout the human genome (an average density of 1 SNP per 1.9 kilobases), providing one of the first highly detailed marker maps of the sequence variability in human genomes.[7] This valuable resource continues to expand as more SNPs are added to the SNP database.

**Single-Nucleotide Polymorphism Analysis**

Because of their mean density, stability, and high-throughput genotyping capabilities, SNPs have recently emerged as genetic markers of choice for disease gene discovery and mapping.[10] Use of SNPs facilitates disease gene mapping in two ways, genome-wide association studies and linkage disequilibrium analysis. Single-nucleotide polymorphisms may be directly associated with a disease trait by effecting the expression or function of the gene where they are located. These "functional" SNPs may exist in a regulatory region, may result in an amino acid change in a gene product, or may alter the exon-intron splicing pattern. Functional SNPs may be enriched in particular disease populations compared with controls. It has been estimated that individuals are heterozygous for 24,000 to 40,000 polymorphisms that have been found to alter amino acid composition.[11] However, it is thought that single disease-related SNP alleles can increase or modify risk for disease, but are not sufficient to cause disease.[12,13]

Alternatively, SNPs may be used as markers for linkage disequilibrium.[14-17] Linkage disequilibrium is the measure of the degree of association between two or more genetic markers that lie near each other on a chromosome. Studies using linkage disequilibrium can identify regions of the genome associated with a disease in a population. Single-nucleotide polymorphisms that alter the risk of disease outcome will be the most predictive of a possible clinical phenotype.

**Genetic Screening for Treatment of Disease**

Single-nucleotide polymorphism analysis also provides a useful tool in genetic screening for the treatment of disease. There are a number of clinically relevant SNPs that have been shown to be associated with drug response and toxicity.[3] Polymorphisms in genes that encode drug metabolizing enzymes for example, are observed at varying frequencies throughout the human population. Among the most commonly prescribed of all anticancer drugs, the thiopurines (eg, mercaptopurine [Purinethol] and thioguanine) must be converted to thioguanine nucleotides by various enzymes in the body. These nucleotides are then incorporated into the patient’s DNA. Polymorphisms in the drug-metabolizing enzyme thiopurine methyltransferase have been linked to the therapeutic efficacy of mercaptopurine, as well as to its toxicity. Patients with two mutant thiopurine methyltransferase alleles have very low thiopurine methyltransferase activity and, therefore, have an impaired capacity to eliminate mercaptopurine and thioguanine from the body. This results in serious, often life-threatening, toxicity.[18]

Gene-specific polymorphisms have been observed in a number of other drug-metabolizing enzymes, including dihydropyrimidine dehydrogenase, glucuronosyl transferase, Cyp17, glutathione transferase, cytochrome P-450, and 5,10-methylenetetra-hydrofolate reductase. Furthermore, polymorphisms have also been identified in genes that encode proteins involved in drug absorption, distribution, and elimination.[19-22] Thus, we are currently able to identify inherited differences between individuals that may affect patient outcomes with anticancer drug therapies.

Apart from the known gene-specific polymorphisms that have relevance to cancer treatment, how can nondisease/nongene specific SNP analysis be used to predict patient response to medicine? An attractive, evolving model is to obtain genome-wide SNP profiles from large numbers of cancer patients receiving anticancer drugs.[4,23] If a specific SNP pattern from patients who responded to therapy is compared with that of patients who failed to respond, a common set of genetic variants between the two groups might be revealed. Additionally, SNP profiles from patients who experience adverse events during treatment can be compared with those patients who did not suffer adverse events to identify DNA regions associated with drug toxicity. Taken together, these SNP signatures, or "medicine response profiles," provide a potentially powerful tool to predict whether an individual is likely to respond to a drug (Figure 2).[4]

**SNP Integration Into Clinical Trials**
Given this paradigm, SNP analysis has significant implications for examining both drug efficacy and safety in clinical trials (Figure 3).[4,24] Single-nucleotide polymorphism analysis can be implemented into clinical trials in two ways.[4] First, patients should be selected for phase III trials based on the response profiles obtained from high-density SNP scoring of responders and nonresponders in phase II trials. Second, SNP profiles should be identified that characterize patients who suffer serious or common drug adverse events compared with those patients who respond to therapy with no drug adverse events. The goal is to combine these two SNP profiles and generate a comprehensive medical response to drug efficacy.

This SNP signature, or fingerprint, would result in the development of phase III trials that are faster to complete, require fewer patients, and cost less to conduct.[25] Moreover, lead molecules targeted to clinically nonresponsive patients could be more rapidly developed. A focused clinical trial approach is important for individuals—both in terms of response to treatment and in sparing the patient unnecessary adverse treatment effects.

**Expression Analysis**

Presently, the diagnosis and classification of human cancer is based on a pathologic evaluation of the histology and morphology of a tumor, which is an essential step in determining appropriate treatment. The pathologic assessment of tumors, however, has limitations: it is subjective and there are no formal grading systems for many tumor types.[26] In addition, the complex molecular heterogeneity that drives and maintains the neoplastic state limits successful therapy, in that tumor response and clinical outcome can vary considerably despite similar histopathological appearances.

Recent work in our laboratory and in a number of others has shown remarkable progress toward an era in which cancer diagnosis will move from a traditional histologic/morphologic approach to a more molecular-based assessment. The evolution of a new "molecular taxonomy" of cancer is becoming possible through the use of the high-throughput genomic technology known as complementary DNA (cDNA) microarrays.[27,28] DNA microarray techniques afford simultaneous expression monitoring of thousands of genes in a single experiment, engendering a molecular portrait or expression profile of both tumor and normal tissues. This is an eminently valuable step toward the basic understanding of the genetic complexities inherent in cancer and has significant clinical implications for revealing pathways and novel targets related to the neoplastic process.

DNA microarray technology is an RNA-based method of gene expression analysis in which there are several formats for producing microarrays.[29,30] The two most commonly employed methods are oligonucleotide arrays and cDNA arrays. Oligonucleotide arrays, pioneered by Affymetrix, Inc. (Santa Clara, Calif), can be generated using a photolithographic process in which oligonucleotides are synthesized directly onto a glass surface.[31] The current Affymetrix human genome GeneChip arrays contain 1,000,000 unique oligonucleotide features that represent approximately 33,000 characterized human genes. Recently, other manufacturers, such as Motorola Life Sciences (Northbrook, Ill; CodeLink Bioarray System) and Agilent Technologies (Palo Alto, Calif; Custom In Situ Oligo Microarray Kit), have also introduced arrays using mechanical microspotting or ink-jet printing instead of photolithography for deposition of the oligonucleotides onto glass slides.

The National Human Genome Research Institute of the National Institutes of Health and a number of other laboratories use a system in which cDNA microarrays are produced by robotically printing a large number of genes onto glass slides containing a gridded array (see http://www.nhgri.nih.gov/DIR/Microarray/main.html for more information).[32-34] The spotted arrays vary with a "packing density" of up to 50,000 elements possible, although ranges from 5,000 to 30,000 cDNAs are more common.

In general, the procedure requires high-quality mRNA that is isolated and purified from two samples (test and control), differentially labeled using reverse transcription in the presence of fluorescent dyes (eg, Cy3dUTP and Cy5dUTP), and hybridized overnight to the glass slide containing the arrayed cDNAs. Following a series of washes, the slide is scanned and monochromatic images of each fluorescent channel are obtained. The images are merged, pseudo-colored (eg, red and green),
normalized, and the relative expression (level of red versus green fluorescence [R/G ratio]) between compared samples is obtained.

The statistical methods used to profile the vast quantities of expression data generated from DNA microarray analyses must be carefully scrutinized, as they can have significant influence on the interpretation of the results. Several computational tools are available for clustering and organizing the data,[35-38] some of which have been summarized by Quackenbush.[39] It is important to note, however, that no consensus has been reached as to the best method for revealing patterns of gene expression changes. Thus, it is critical that molecular targets identified by DNA array technology be validated (discussed below).

Refined Classification of Cancer Subtypes

The ability to analyze the simultaneous expression of thousands of genes should greatly improve cancer classification and have a profound effect on the way new drug targets are identified. Gene-expression profiling using DNA microarrays has demonstrated unparalleled progress in the ability to subclassify previously unknown or indistinguishable subtypes of lymphoma,[40] leukemia,[41] breast carcinoma,[42-44] melanoma,[35] prostate,[45,46] and ovarian cancer.[47] Other studies utilizing cDNA microarrays have been designed to assess gene expression patterns in estrogen-receptor-positive and -negative breast tumors,[48] as well as familial breast cancers, where hereditary signatures for BRCA1 and BRCA2 mutated tumors were obtained.[43]

The latter study was particularly important in demonstrating the power of gene-expression profiling. In this study, one of the tumors in which no BRCA1 gene mutations were identified displayed a molecular profile that was characteristic of tumors with BRCA1 mutations. Subsequent analysis of this obscure tumor revealed BRCA1 gene silencing through abnormal methylation in the promoter region.[43] These studies clearly illustrate that molecular profiling of tumors using DNA microarrays shows tremendous promise for improving tumor classification.

Implicit in the use of DNA microarrays is the ability to molecularly dissect global gene expression changes that occur in the carcinogenic process. For example, gene expression changes that distinguish benign prostatic hyperplasia from prostate cancer may reveal genes that are important in prostate tumorigenesis.[45,46] Not surprisingly, striking degrees of molecular variation are often revealed among samples, particularly when comparing normal tissue with its cancerous counterpart, making it is essential to establish a database that permits the data management and comparison of large quantities of information. As discussed, the relationships between the gene expression changes require sophisticated data mining techniques, such as hierarchical clustering, multidimensional scaling, artificial neural networks, and self-organizing maps.[39]

Elucidation of Pathways and New Targets

Global gene expression analysis has provided great insight into the molecular heterogeneity of specific types of cancer. However, because it is often difficult to assign functions to a large number of identified genes, cancer pathway discovery remains in its infancy. Still, molecular profiling holds tremendous promise for the identification of novel targets for therapy and the development of drugs tailored to the genetic changes critical for malignant transformation and tumor progression. Several studies are now beginning to use tumor samples with available follow-up information to assess the prognostic significance of gene-expression profiles.

An elegant example illustrating the clinical relevance of DNA microarray technology was provided by Alizadeh et al.[40] Using a "lymphochip" containing 17,856 genes expressed in lymphoid cells, Alizadeh and colleagues performed a global gene expression analysis on diffuse large B-cell lymphoma, a cancer that displays great clinical heterogeneity (only 40% of patients respond to therapy). Here, an unsupervised clustering of the data was able to dissect two molecularly distinct subsets of the disease into germinal center B-like diffuse large B-cell lymphoma and activated B-like diffuse large B-cell lymphoma. More importantly, the authors demonstrated that patients classified with germinal center B-like diffuse large B-cell lymphoma had better overall survival than those patients with the molecular profile characteristic of B-like diffuse large B-cell lymphoma. This study
Role of Genomics in Identifying New Targets for Cancer Therapy

Published on Cancer Network (http://www.cancernetwork.com)

illustrates how specific molecular portraits may be able to predict patient outcome.

Similarly, an important study by Shipp et al utilized gene-expression profiling to predict the outcome in a series of diffuse large B-cell lymphoma patients.\[49\] Thirteen key genes were identified as top prognosis discriminators for patients who had a high chance of dying and patients that demonstrated reduced recurrence following treatment. These results differ from the former study in that Alizadeh et al hypothesized a cell of origin (ie, germinal center B-like or activated B-like) might be predictive of a clinical response, whereas Shipp and colleagues used supervised learning methods to correlate outcome with gene expression.

Another example of the clinical relevance of cDNA microarray technology was recently demonstrated with a set of pediatric tumors, the childhood small-, round-, blue-cell tumors, \[50\] which include Ewing’s sarcoma, neuroblastoma, rhabdomyosarcoma, and non-Hodgkin’s lymphoma. This class of tumors is often difficult to diagnose and shows variable response to chemotherapy.\[51\] Khan et al identified a set of 80 known genes and 13 anonymously expressed sequence tags that discriminate the four distinct subtypes of small-, round-, blue-cell tumors. Of interest, one commonly used marker to diagnose Ewing’s sarcoma, MIC2, was highly expressed in several rhabdomyosarcoma samples. These results suggest MIC2 alone may not represent an appropriate diagnostic marker for Ewing’s sarcoma.

**Profiling Tumor Response/Sensitivity**

DNA microarray technology shows great promise for revealing gene expression changes that are predictive of tumor response and sensitivity to therapeutic agents. However, the majority of drug response studies, including studies of drug resistance, have been limited to cell culture analyses.\[52-54\] A massive effort to identify in vitro growth inhibitory activities of 60,000 compounds has been established at the National Cancer Institute. It is anticipated that a DNA microarray approach to examine gene expression changes in a panel of 60 cancer cell lines exposed to various compounds will help to identify new agents for clinical trials.\[55\]

Several studies have demonstrated the remarkable power of DNA microarrays to reveal the underlying molecular diversity among subtypes of cancer. These applications will undoubtedly have tremendous impact on efforts to individualize cancer therapy and improve treatment success. Molecular targets identified by global gene expression analysis may be useful diagnostic or prognostic indicators and may be important indicators for cancer therapy failure or toxicity. Of significant importance, however, is that the future goal of using high-throughput array technologies as a step toward a tailored treatment approach relies heavily on the careful selection and evaluation of patients who have undergone treatment in clinical trials and on access to large, publicly available data sets that can define clinical objectives.

**Tissue Microarrays**

It is necessary to validate the genes that are leading indicators from array experiments to determine if they can provide meaningful information for clinical applications. One important process may be to verify the gene against archival tumor specimens with known clinical outcome. However, gene-expression profiling reveals large numbers of genes (eg, 10 to 300 genes per tumor) that must be validated in large sets of tumors to obtain statistically significant data. Validation would take years to complete in a traditional setting and could potentially exhaust banked archival specimens. For these reasons, tissue microarray technology is a vital technological advance that permits high-throughput validation of candidate target genes.\[56\]

Tissue microarrays can contain up to 2,500 different archival tumor samples on a single glass slide. In this method, a core biopsy of tissue (0.6 mm in diameter) is punched from a preselected region of a paraffin-embedded tumor (the donor block) and placed into a recipient paraffin block containing premade holes in a defined array (Figure 4).\[56\] Each hole receives a different tumor specimen. The filled recipient block is sectioned and used for the simultaneous in situ analysis at the DNA, RNA, and protein level.\[57\]
Several studies have combined DNA and tissue microarray technologies to validate gene targets. For example, Hedenfalk et al used tissue microarray technology to demonstrate that the protein levels (determined by immunohistochemistry) of two selected genes, cyclin D1 and mitogen-activated protein kinase kinase-1, correlated with the DNA microarray results (Figure 5).[43] Similarly, Bubendorf et al used tissue microarrays containing a broad spectrum of prostate cancer samples to validate overexpression of insulin-like growth factor-binding protein 2 and heat-shock protein (HSP27) genes identified in DNA microarray analysis of hormone-refractory prostate cancer.[58] These results demonstrate the power of combining two high-throughput technologies to validate candidate target genes more efficiently and rapidly.

**Conclusions**

The growing body of data shows convincing evidence that the combined use of the high-throughput technologies described here is certain to unravel the genetic complexities inherent in cancer and revolutionize cancer therapy. Single-nucleotide polymorphism analysis, combined with cDNA microarrays and tissue microarrays, should help to reveal specific genetic profiles and novel molecular targets that have importance in drug efficacy and toxicity. Linking these technologies could be a key element in the quest to broadly and rapidly develop approaches to test individual patients’ tumors for specific therapeutic targets. These technologies may, indeed, revolutionize cancer diagnosis and treatment.

**References:**


Source URL:
http://www.cancernetwork.com/review-article/role-genomics-identifying-new-targets-cancer-therapy

Links:
[1] http://www.cancernetwork.com/review-article
[3] http://www.cancernetwork.com/authors/sarah-l-anzick-phd
[4] http://www.cancernetwork.com/authors/jeffrey-m-trent-phd