Pharmacogenomics:
The relevance of emerging genotyping technologies

By Tina Hernandez-Boussard, PhD; Teri E. Klein, PhD; and Russ B. Altman, MD, PhD

Analysis of genotype-phenotype relationships, particularly drug effects due to genetic variation — pharmacogenomics — has greatly evolved over the past several years. Advances in genomic technologies, best defined as methods used to manipulate and analyze genomic information, have catalyzed this evolution. Before 1980, few human genes had been identified as genetic risk factors for hereditary disorders; and few links between ethnicity or inheritance and deviant drug responses to a gene were made by biochemical studies. These disorders and drug responses were mostly monogenic. Analyses of these monogenic relationships relied on the knowledge of the disease-associated or drug-associated gene located in a specific chromosomal region.

With cloning of DNA in the early 1980s, molecular genetics vastly expanded the opportunity to study variations in DNA and their effects on disease and drug reactions. From about 1992 to 2002, microsatellite markers were the “hot” technology used to study these disease- or drug-associated genes, particularly in linkage-analysis studies. In these time-consuming and costly studies, about 300 to 800 microsatellite markers were identified throughout a genomic region and studied in a single DNA sample. Between microsatellites and arrays was the five-year-long “re-sequence” era — expensive but useful for discovering single nucleotide polymorphisms (SNP).

More recently, high-density genotyping arrays have become available for the study of genetic variation. These genome-wide technologies use in situ hybridization for SNP detection and contain anywhere from 1,000 to 500,000 SNP markers, with about one marker per five kilobase, or 5 kb. These arrays, in conjunction with the abundance of genomic-sequence information now available, have lead to a new era in biology.

What will new genotyping technologies bring to pharmacogenomics, and how will this information eventually lead into the clinical arena to produce individualized drug therapy?

Where are we today

We are in an era of post-genome biology. The human genome has been sequenced, and genetic variation is being characterized. We have an understanding of what different parts of the...
With the completion of the Human Genome Project and the recent advances in genome technology, we now have the capabilities to analyze pharmacogenomic data in a timely and cost-efficient manner.

There is a growing list of gene variants that contribute to aberrant drug responses, from drug-metabolizing enzymes to drug transporters to drug targets as well as variants that are disease-related. Before genome-wide analyses were available, most aberrant gene-variant correlations were monogenic, such as those found in Mendelian-inheritance patterns. Candidate gene approaches have been used for these tedious, costly studies. Single genes related to an adverse drug reaction were based on clinical pharmacological studies of proteins (receptors) and pathways known to be involved in a drug pharmaco-kinetic or -dynamic response. An example is thiopurine S-methyltransferase (TPMT) and its enhancement of the S-methylation of thiopurine drugs — an effect due to genetic variation that can be responsible for thiopurine toxicity. It is clear, however, that monogenic relations are not the case for most diseases, such as diabetes, cancer, heart disease, and pharmacological deviants. With the completion of the Human Genome Project and the recent advances in genome technology, we now have the capabilities to analyze pharmacogenomic data in a timely and cost-efficient manner.

Types of studies being performed

Several different types of studies addressing genetic variation are being performed today on these emerging genome-wide technologies, depending on the marker density of a genotyping array. Broadly speaking, whole genome genotyping arrays containing 3,000 to 10,000 markers are sufficient for linkage-analysis studies; arrays containing 10,000 to 100,000 are sufficient in marker density to study loss of heterozygosity and comparative genomic hybridization. Association studies, however, require a large number of study subjects (cases and controls) and high-density arrays, such as those with 100,000 to 500,000 markers on a single array.

Linkage-analysis studies with surrogate markers have been successfully used to track genomic regions as they co-segregate with a disease through a pedigree. Before high-density arrays were available, hundreds of extremely laborious, expensive microsatellite markers were individually analyzed to address Mendelian diseases and other monogenic phenotypes. Common practice for these studies now is to use genome-wide SNP arrays, which contain thousands of SNP markers in one assay; and, therefore, thousands of markers are tested on a single sample. SNP markers on genome arrays — used successfully in

Continues on page 27
many linkage-analysis studies — are biallelic and evenly distributed throughout the genome. For example, in a matter of days, Puffenburger, et al, mapped sudden infant death with dysgenesis of the testes syndrome, or SIDD T, to a novel loss of function in the TSPYL gene using a 10K SNP genome array.6 Another linkage-analysis study using a 10K SNP array showed that a Mendelian locus on chromosome 16 determines susceptibility to doxorubicin nephropathy in the mouse.5 Both of these studies identified novel genes that can be used as possible drug targets for the related disease.

The great benefit of using these new technologies is that these studies are being produced in days instead of months and years. The saved time alone reduces the cost of the overall experiment. Additionally, several comparison studies have been performed between traditional microsatellite markers and SNP array markers, and have shown that the overall information content is increased by at least 20%, with a mean SNP call rate of 95% and a false-positive rate of 0.05% to 0.08% using the SNP arrays.6

The great benefit of using these new technologies is that these studies are being produced in days instead of months and years.

Genome-wide molecular biological analyses, such as loss of heterozygosity (LOH) profiling and comparative genomic hybridization (CGH) have significantly enhanced our ability to detect chromosomal aberrations in cancer cells and assess their role in tumorigenesis; certainly larger genome-wide SNP arrays have been useful. In these genome-wide studies, whole regions of the chromosome that are amplified carry candidate oncogenes, and whole regions that are deleted carry potential tumor suppressor genes — giving a better understanding of the complex events that occur during tumorigenesis. LOH studies are useful in identifying new tumor-suppressor genes based on the model that one inherited allele is mutated and the other is lost somatically, which is often the case in cancer tissue.

For example, Garraway, et al, ran 100,000 genome-wide SNP arrays on NCI60 cell lines. They first looked at chromosome copy number changes and then added the loss of heterozygosity information to further analyze certain regions of amplification. From their analyses, they found a novel therapeutic pathway into the treatment of melanoma targeting a novel gene, microphthalmia-associated transcription factor, or MITF, which was identified in one of the amplified regions on the genome-wide SNP array.7 This study and other like studies demonstrate that the use of genome-wide arrays to analyze LOH and CGH gives greater insight into the underlying genetic alterations in cancer cells with identification of complex events, including loss and reduplication of loci.

Most diseases are influenced by both genetic and environmental factors. To study these complex diseases, a large sample of well-placed markers must be taken throughout the genome in either a case-control association study or a transmission/disequilibrium study. Using high-density SNP arrays is ideal, as the approximately 500,000 markers on the array have been strategically placed and expected to fall within close proximity to true disease-causing variants. It is not expected that these 500,000 SNPs genotyped on the array are the actual disease-causing variant, but that they are in linkage disequilibrium with a disease-causing mutant. These technologies have proven fruitful in several association studies. For example, a recent association study on a Japanese population yielded several new candidate loci for Type 2 diabetes,8 a complex disease that is the result of many environmental factors as well as several expected genetic variants.

**Future directions**

There are still many challenges in this field, and we still do not fully understand the contribution of all genetic variants to individual drug responses or how to make this information useful at the clinical level. Most SNPs that are disease-causing are very rare events and, therefore, difficult to capture in a case-control study. SNPs that enhance an individual’s risk for disease are more abundant, but the relationship between

---

**Reach Lower... and Reduce the Strain**

With the New Portable Pipet-Aid® XL Pipetting in a Hood Doesn’t Have to be a Stretch

- Longer lightweight handle lowers the arm position required to perform the same task as with a conventional pipettor
- New quiet, more efficient pump extends operation period between charges
- Can be charged while in use to eliminate downtime
- Three-speed settings provide more precise control for both aspirate and dispense modes

For more information on the New Drummond Pipet-Aid® XL call 1-800-523-7480 or visit our web site at www.drummondscl.com

From the Developers of the Original Pipet-Aid®

www.mlo-online.com
PharmGKB catalyzes pharmacogenomics research and deployment

The Pharmacogenetics and Pharmacogenomics Knowledge Base (PharmGKB) is a public resource that promotes research into the relationships between human genotypes, phenotypes, and clinical outcomes by publishing, linking, and annotating primary datasets from ongoing research and established data from the literature. PharmGKB has worked with others to develop a number of disease and drug-centered pathways to identify target genes and investigate functional polymorphisms. PharmGKB pathways are generated by collaboration of investigators to link data, either novel or in the public domain, centered on a particular drug [the thiopurine pathway is shown in Figure 2], and the representation is a consensus of the opinions of the authors. PharmGKB and its associated network have developed a public database to catalyze pharmacogenetics research and deployment, as outlined in the accompanying article.

Figure 2. The thiopurine pathway from PharmGKB. Drugs are depicted by purple boxes, transporter genes by green ovals, and genes coding for metabolic enzymes by blue ovals.

The next and more powerful step that PharmGKB catalyzes pharmacogenomics research and deployment is the Pharmacogenetics and Pharmacogenomics Knowledge Base (PharmGKB). This public resource promotes research into the relationships between human genotypes, phenotypes, and clinical outcomes by publishing, linking, and annotating primary datasets from ongoing research and established data from the literature. PharmGKB works with others to develop disease and drug-centered pathways to identify target genes and investigate functional polymorphisms. These pathways are generated by collaboration of investigators to link data, either novel or in the public domain, centered on a particular drug [the thiopurine pathway is shown in Figure 2], and the representation is a consensus of the opinions of the authors. PharmGKB and its associated network have developed a public database to catalyze pharmacogenetics research and deployment, as outlined in the accompanying article.

Figure 2. The thiopurine pathway from PharmGKB. Drugs are depicted by purple boxes, transporter genes by green ovals, and genes coding for metabolic enzymes by blue ovals.
genotyping strategies will help develop personalized medicine and individualization of diagnosis and therapy. Genotype studies have a significant value to medicine, since a better understanding of a patient’s genetic makeup can allow physicians to avoid drugs with deleterious effects. If a patient is found to have a particular variant in a drug gene target, dosage can be modified or medication can be individualized to avoid the aberrant drug effects that could possibly occur due to this gene variant.

Obvious benefits from pharmacogenomic information are still not apparent in a clinical setting; however, before this translation can occur, several factors need to be changed on the healthcare side, as outlined recently by Evans, et al.13 Here, Evans points out that, first, healthcare professions will need to be re-educated in the way that they prescribe medication, such as dosing based in genotype, ethnicity, or age of the patient, rather than a mean dose for an entire population. Additional genotyping tests will need to be performed before the proper medication can be administered; this is not common practice today. Both the clinician and the patient will have to accept these extra tests that will be required to obtain one’s genotype before a drug can be administered. In addition, clinical trials will also need to take genotyping into consideration, and perhaps this would save certain drugs from abandonment. For example, drug X has adverse effects in patients with genotype Y but is effective for patients who have genotype Z. The deleterious drug effects are only seen in patients with the genotype Y.

Pharmacogenetics holds promise as a form of individualized medicine, but it is not yet a routine component in patient diagnoses.14 In this post-genome era, many informatics challenges require the marriage of bioinformatics and clinical informatics. We are one step closer to a more unified laboratory and clinical setting, merging pharmacogenomic information into the clinical world.

**Acknowledgement:** This work is supported by NIH U01GM61374.


Continues on page 30


