Clinical application of high throughput molecular screening techniques for pharmacogenomics

Arun P Wiita1
Iris Schrijver2

1Department of Laboratory Medicine, University of California, San Francisco, CA, USA; 2Department of Pathology and Pediatrics, Stanford University School of Medicine, Stanford, CA, USA

Abstract: Genetic analysis is one of the fastest-growing areas of clinical diagnostics. Fortunately, as our knowledge of clinically relevant genetic variants rapidly expands, so does our ability to detect these variants in patient samples. Increasing demand for genetic information may necessitate the use of high throughput diagnostic methods as part of clinically validated testing. Here we provide a general overview of our current and near-future abilities to perform large-scale genetic testing in the clinical laboratory. First we review in detail molecular methods used for high throughput mutation detection, including techniques able to monitor thousands of genetic variants for a single patient or to genotype a single genetic variant for thousands of patients simultaneously. These methods are analyzed in the context of pharmacogenomic testing in the clinical laboratories, with a focus on tests that are currently validated as well as those that hold strong promise for widespread clinical application in the near future. We further discuss the unique economic and clinical challenges posed by pharmacogenomic markers. Our ability to detect genetic variants frequently outstrips our ability to accurately interpret them in a clinical context, carrying implications both for test development and introduction into patient management algorithms. These complexities must be taken into account prior to the introduction of any pharmacogenomic biomarker into routine clinical testing.

Keywords: high throughput, clinical laboratories, pharmacogenomics, mutation

Introduction

With the completion of the human genome project and much subsequent research, it has become clear that millions of loci within the human genome can vary from person to person.1,2 Nucleotide variation occurs not only within protein coding sequences of known genes but also in upstream regions, intronic sequences, and sequences far from any identified genes. Importantly, only a small minority of these genetic changes are clearly related to a given disease or phenotypic state.3 While some of these alterations are insertions or deletions, the majority of these variations occur as a change in a single nucleotide, known as single nucleotide polymorphisms (SNPs). These inherited SNPs typically involve just two alternative nucleotides at a given genomic position. These variants may be present in either heterozygous or homozygous form on the two chromosomal alleles that determine individual genotype (ie, A/A, A/C, or C/C). The most common nucleotide in a population for a given SNP is termed the major allele while the less frequent variant is termed the minor allele.

Our knowledge of such genomic variation among humans has grown immensely over the past decade. As a result, a major challenge is discerning which of these SNPs are clinically relevant and which are phenotypically silent. One of the major areas
of research into this allelic variation is to determine which genetic variants affect outcomes related to pharmacologic interventions. Certain sequence variations have been known for many years to affect drug metabolism, efficacy, or to increase the risk of adverse reactions, and were either identified by classic genetic studies or targeted mechanistic or experimental approaches; this gene-by-gene approach has traditionally formed the basis of pharmacogenetic studies. More recently, new advances in molecular biology and bioinformatics have allowed for extensive analysis of human nucleic acid sequence and protein expression. For example, genome-wide association studies have taken advantage of methods to analyze hundreds of thousands of SNPs across large populations combined with clinical outcomes data and comprehensive statistical methods. These studies have identified either individual SNPs or combinations of variants associated with drug response. In addition, expression of certain messenger ribonucleic acid or proteins in tumors has led to novel targeted therapies. These new approaches to identifying drug effects in relation to gene sequence and expression are termed pharmacogenomics.

In practice, the terms pharmacogenetics and pharmacogenomics are often used interchangeably. However, it is pharmacogenomic information derived from large-scale genome screening that poses the biggest challenge for applications to the clinical laboratory: our collective capability to analyze SNPs in patients has often surpassed our ability to interpret the data in a clinically meaningful way. Nevertheless, there are many tests related to drug response that have already been incorporated into current clinical testing. Furthermore, it is likely that the variety of such tests will expand in the future as more data connect allelic variation, drug response, and clinical outcomes. As genetic variation becomes more connected to patient management, it is anticipated that the demand for clinical testing for genetic markers will increase. Such increased demand may require the testing of multiple markers in numerous patient samples in a rapid and parallel manner. In this review, we will first discuss various methods of identifying SNPs with high throughput in patient samples. We will then examine these technical advances in the context of practical applications to clinical diagnostics. This review does not attempt to comprehensively cover all available tests and methodologies; rather, it aims to provide a general framework for the current state of clinical pharmacogenomic testing.

High throughput methods of mutation detection

“High throughput” testing of genetic variants or mutations can have a number of meanings, which may variously incorporate: (1) testing hundreds of different patient samples simultaneously (parallelizing), and/or (2) analyzing multiple genetic loci simultaneously for a single patient (multiplexing), and/or (3) obtaining very rapid results. The choice of technological method to be used for a given high throughput genetic test in the clinical laboratory depends strongly on the testing goals. Beyond technical decisions, bringing such a test into routine clinical diagnostics further requires additional, and potentially more important, considerations such as: (1) anticipated clinical application, (2) projected demand for testing (number of samples), (3) cost of testing, (4) complexity of establishing testing, and (5) regulatory status of available tests. Below we review examples of various available methods in the context of these criteria.

Large-scale multiplexing with array-based SNP analysis

Microarray technologies are among the best-known methods by which to determine SNP status, whether for pharmacogenomic status or disease risk factors. This approach uses differential hybridization of nucleic acid oligonucleotides to determine the presence of sequence variants. With high-density chip deposition of standardized oligonucleotides, the Genome Wide SNP Array 6.0 (Affymetrix Inc, Santa Clara, CA) can probe for 906,600 individual SNPs (of over 10 million human SNPs identified) after random amplification and labeling of patient genomic DNA and subsequent hybridization. BeadArray technology (Illumina Inc, San Diego, CA) can probe 1 million SNPs using a different chemistry: microbeads with ligated oligonucleotides complementary to specific SNPs subjected to a single base extension reaction after hybridization. Both of these methods allow for hundreds of thousands of SNPs to be analyzed under a single set of experimental conditions with a time to completion of approximately 1 day.

The main drawback of these approaches is high reagent cost, lack of clinical utility for a subset of the markers analyzed, and difficulty validating each spot on the chip. In particular, the presence of false-negative results is difficult to assess without an independent patient cohort for test validation. In addition, this testing may be time consuming in terms of sample preparation with concomitant increased
Medium-scale multiplexing with low to medium numbers of samples in parallel
The next approaches described are useful for probing approximately 10–200 pharmacogenomic markers in a multiplex fashion. As opposed to the whole genome SNP arrays described above, these tests are ordered to achieve a specific diagnostic goal for a given patient, which, given our current clinical knowledge, involves analyzing a more limited number of genetic changes. These methods are amenable to variable degrees of parallelization to accomplish probing of multiple patient samples simultaneously.

Selected marker microarrays
Using dedicated microarray technologies with a relatively small number of hybridization probes attached to the substrate, it is possible to probe just a few genetic markers at a lower cost than whole genome array technologies. Instead of random amplification of the entire genome, these approaches use an initial polymerase chain reaction (PCR) step with primers specific for genomic areas of interest. Differential hybridization is detected by array reader and proprietary software is used to report a genotype, sometimes with a corresponding clinical interpretation. Some sample modifications can be used to accommodate multiple patient samples in parallel. Examples include the Roche AmpliChip® CYP450 Test (F Hoffmann-La Roche Ltd, Basel, Switzerland) and the Affymetrix DMET™ Plus.

The arrayed primer extension technology uses a similar model of custom-made arrays with oligonucleotides complementary to genetic regions specifically amplified. The oligonucleotides hybridize to a point one nucleotide upstream of the potential variant. A single base pair extension reaction is carried out with fluorescence-based determination of nucleotide incorporation at each specific oligonucleotide probe. This method can be used to potentially genotype dozens of nucleotide variants simultaneously for a single patient sample. The AutoGenomics INFINITI® (Vista, CA) method uses a similar principle of differential hybridization and additional specific primer extension for SNP detection on an array. For these array-based methods the main drawbacks again are related to cost; the reagents and instruments typically are relatively expensive and the available clinical test menus limited.

Targeted bead-based arrays
One of the technologies frequently used is the bead-based array available from Luminex Corporation (Austin, TX). Oligonucleotides targeting the genomic areas of interest are covalently ligated to microspheres impregnated with differentially colored dyes, where each dye combination corresponds to a particular genetic variant. After hybridization with amplified patient DNA and a primer-extension reaction incorporating a labeled mutation-specific oligonucleotide adapter, a genotype can be determined using fluorescent detection and flow cytometry. This technology is available for clinical testing for dozens of multiplexed markers performed on dozens of samples in parallel. While useful for multiplexed assays, one of the main drawbacks is relatively extensive hands-on time for assay preparation. An advantage is that this platform is open and fully customizable for future test development, though the number of widely available clinical tests is currently limited.

The Verigene® System (Nanosphere Inc, Northbrook, IL) is a hybrid microarray and bead-based array. In this system, fragmented genomic DNA is first hybridized to a chip substrate with partially complementary oligonucleotide sequences. Subsequently, the chip is incubated with gold nanoparticles functionalized with complementary DNA specific for the genetic variants; silver staining detects successful binding of gold particles. A similar technology is used in the GenMark Dx eSensor® system (GenMark Diagnostics Inc, Carlsbad, CA) with ferrocene-labeled oligonucleotides leading to electron transfer to a gold substrate if the specific variant is present. Whereas the Verigene system does have the advantage of avoiding target amplification, unlike the GenMark Dx, both systems require stand-alone instrumentation and tests cannot be easily customized.

Targeted DNA sequencing
Direct DNA sequencing is one of the most important methods of probing multiple genetic variants in a single patient sample. Significantly, only DNA sequencing methods allow for the identification of novel genetic variants; all other methods presented here require prior knowledge of the genetic variants to be probed. Given the intense efforts over the past few years to reduce the cost of sequencing in anticipation of widespread whole genome sequencing, a number of different technologies are now becoming available. We include
examples of methods which demonstrate either current utility or promise in the near future for implementation in the clinical laboratory.

The most widespread sequencing technology used in clinical diagnostics is classic Sanger sequencing. Regions of interest are first amplified and polymerase-mediated extension then takes place in the presence of both unlabeled nucleotides and fluorescently labeled, extension-terminating nucleotides. The hundreds of extension fragments produced are subsequently separated by high-resolution capillary electrophoresis, allowing for determination of nucleotide sequence in read lengths over 500 base pairs. A main advantage of Sanger sequencing is its relative ubiquity in many clinical applications and laboratories, as well as its low error rates and relatively long read lengths. Increased throughput is achieved by arraying numerous electrophoresis capillaries in parallel using plate-based systems or by using various electrophoresis polymers allowing for more rapid fragment analysis.

An alternative sequencing method is pyrosequencing, a “sequencing by synthesis” approach as opposed to the chain termination in Sanger sequencing. In this method, nucleotides are sequentially flowed into a chamber containing an immobilized DNA fragment and a polymerase. The release of a pyrophosphate molecule after incorporation into an elongating DNA strand is monitored using a luciferin-luciferase system and a high-sensitivity photodiode. This method can achieve read lengths of 300–500 nucleotides and has been used in many clinical applications as an alternative to Sanger sequencing as it can be more sensitive, less expensive to operate per test, and largely automated to facilitate higher throughput.

“Next generation” sequencing methods hold great promise for the detection of somatic mutations. These approaches are capable of sequencing whole human genomes; this feat can now be accomplished in a few days with a cost in the thousands of dollars as opposed to Sanger methods which took years and millions of dollars. Whole genome sequencing is currently largely beyond the scope of clinical testing, but recent technological advances make this a realistic possibility in the near future.

Though multiple methods exist, here we will discuss two of the most established next generation technologies, produced by 454 Life Sciences (Roche, Branford, CT) and Illumina. The 454 system uses a bead-based approach. Genomic DNA is first fragmented and ligated to adapter oligonucleotides, which are then attached to microbeads. Clonal amplification in an emulsion-PCR “microreactor” system is then performed. The beads, each carrying a clonally amplified DNA fragment, are deposited onto a microtiter plate. Pyrosequencing and high-resolution detection is used to detect incorporated nucleotides for millions of DNA fragments in parallel. This system may allow for hundreds of megabases of sequence read in a 10-hour run. In the Illumina system, fragmented DNA is ligated to primers on a glass substrate, allowing for subsequent clonal “bridge” amplification. As a labeled nucleotide is incorporated into the expanding DNA strand, a camera reads the specific fluorescent color for each base. The dye is then cleaved and the next nucleotide added and read. Like the 454, this system also allows for massively parallel sequencing of thousands to millions of DNA fragments simultaneously.

As compared to Sanger sequencing, these next generation methods allow for much higher throughput of sequencing data. However, these next generation sequencing systems first appeared in versions much too expensive, too powerful, and too cumbersome for the clinical laboratory. Recently, however, both corporations have developed scaled-down versions (the 454 GS Junior and the Illumina MiSeq Personal Sequencing System) as so-called “personal genome” sequencers. These instruments use chemistries similar to the original instruments but feature simplified workflow, less expensive reagents, and a faster turnaround time designed for applications such as targeted sequence analysis, human exome sequencing, or bacterial genome sequencing. Other methods for this smaller-scale genome sequencing are also emerging.

These instruments are designed and marketed to be used for clinical diagnostics in the near future. However, all these next generation sequencing methods remain to be comprehensively proven for clinical applications. The cost of the instruments and the cost of reagents per run are relatively high, and the experimental protocols require considerable hands-on time and experienced technologists. In addition, limitations in sequence read length and the relatively high error rate of base calls are of concern, especially in the detection of relatively rare mutations such as in cancers.

**Medium to high numbers of parallel samples with low- to medium-scale multiplexing**

The methods in this category comprise some of the most established methods for mutation analysis in clinical samples. These techniques are characterized by the use of specific reagents for the detection of one or a few nucleotide variants in genomic DNA. They typically can be performed using
instrumentation present in many clinical laboratories for a relatively low cost per test.

Differential primer hybridization and amplification

TaqMan® Assays (Applied Biosystems, Life Technologies Corporation, Carlsbad, CA) are one of the simplest and most widespread methods of SNP genotyping using differential hybridization.²⁷ This real-time PCR method uses four oligonucleotides: two unmodified PCR primers which flank and amplify the target region of interest, and two short TaqMan (hydrolysis) probes which hybridize with different affinities to the wild-type and mutant allele to be investigated. In the real-time PCR cycle, one probe will bind specifically to its corresponding SNP site while the other probe will be either unstably bound or unbound to the target. Beginning at the PCR primer, nucleotide extension by a DNA polymerase with 5′ exonuclease activity will encounter the stably bound probe and degrade it. A dye and quencher, placed at opposite ends of the probe, will then diffuse apart and the fluorescence color corresponding to that SNP probe will be detected. This method allows for the simple, rapid, and closed-tube detection of SNP genotypes using real-time PCR instrumentation present in many molecular diagnostics laboratories. Analysis of multiple SNPs in the same tube is relatively limited, because spectral overlap between emission dyes and associated detection instrumentation limits large-scale multiplexing.²⁸ However, the use of 96- and 384-well plates allows for hundreds of samples to be easily tested in parallel and results can be obtained in less than 2 hours. Using novel methods of small volume, chip-based reactions, Applied Biosystems has recently released the TaqMan OpenArray™ Genotyping System allowing for high throughput analysis of up to 16 SNPs simultaneously for 144 patients.

Another commonly used real-time PCR platform for mutation analysis is the LightCycler® 480 Real-Time PCR System (Roche).²⁹ Whereas the LightCycler can perform TaqMan-style genotyping it also has the advantage of performing rapid melt curve analysis. In this analysis, one probe overlaps the genetic variant (sensor probe) while another binds to a nearby constant region of DNA. Fluorescent dyes are incorporated into the probes which lead to a fluorescence resonance energy transfer interaction when excited by an appropriate laser wavelength. As the temperature is gradually increased, the probe with a mismatched nucleotide will become unbound at a lower temperature, leading to a loss of fluorescence resonance energy transfer. This loss of fluorescence as a function of temperature allows for easy genotyping. Given the use of fluorescence and multiple oligonucleotides, this system also does not lend itself to large-scale multiplexing of multiple genetic markers in a single patient sample. However, the LightCycler system can be scaled up for parallel sample testing, ranging from 24 to 1536 samples simultaneously.

In the context of high throughput mutation detection in the clinical setting, turnaround time is an important factor. While the above TaqMan- and LightCycler-based assays only require approximately 1–2 hours to produce analyzable data once a sample is placed on the instrument, there are many preceding steps of DNA extraction and reaction preparation. Thus, these tests are typically batched and run relatively infrequently (eg, once a week), leading to relatively long turnaround times from when a sample is received. Recently, technologies have been developed which provide “all-in-one” hands-off sample preparation and real-time PCR analysis with results provided in under an hour from sample receipt. The most widespread technology is the GeneXpert System (Cepheid, Sunnyvale, CA). After minimal processing, a patient sample is added to the proprietary cartridge, where a microfluidic system allows for automated DNA extraction and subsequent real-time PCR detection of a genetic target. Operation of the instrument requires minimal training and could potentially be used either in the clinical laboratory for on-demand testing or even as a point-of-care test in the clinic. The main drawback of this approach is the relatively high cost of consumables with only one patient sample per cartridge, plus a test menu currently focused largely on microbiology applications.³⁰,³¹

Of note, other real-time PCR systems with similar capabilities to those above, with many variations on assay design, also exist and have been applied to clinical diagnostics.³²

Single base extension with oligonucleotide detection

Mass spectrometry methods can determine SNP genotypes as well. The most widespread method, MassARRAY Analyzer 4, is available from Sequenom Inc (San Diego, CA).³³ Target DNA is amplified and hybridized to specifically designed oligonucleotides which are complementary up to one base before the SNP of interest. Single base extension by DNA polymerase is then carried out with mass-modified terminator nucleotides. The resulting DNA fragments can subsequently be resolved by mass spectrometry methods given the known mass of both the oligonucleotide probe and the labeled terminator. Genotyping up to 40 SNPs in a single assay is reportedly possible, though most authors achieve
a lower degree of multiplexing due to primer interactions and inability to resolve all base variants simultaneously. Automated methods can be used to conduct up to 384 assays on a single matrix chip, allowing for the probing of hundreds of SNPs for one patient or a few SNPs for hundreds of patients, facilitating high throughput. Some of the main benefits are the relatively simple sample preparation procedures, lack of fluorescently labeled reagents, and high accuracy of the mass spectrometry methods. Some of the drawbacks, however, include the expense of the necessary instrumentation and chip consumables, as well as the required time and effort to initially develop a clinical assay. In addition, the baseline signal must be carefully modulated to optimize accurate base calling. The SNaPshot® Multiplex System (Applied Biosystems) is another approach for genotyping by single base extension. The principle itself is quite similar to that described for the MassARRAY above, though here terminator nucleotides are modified to incorporate specific fluorescent dyes. These single base extension products can then be resolved using capillary electrophoresis; the color of the dye incorporated corresponds to the nucleotide present at the variant of interest. Extension primers are designed with different lengths to denote analysis of a particular mutation. This system lends itself to multiplexing up to ten genetic variants in a single assay, and plate-based methods can be used for parallelization. The data from this assay are relatively simple to interpret, the reagents are moderately priced, and the assay uses instruments that may already be present in clinical laboratories.

Other methods
While not used in a high throughput context, restriction fragment length polymorphism analysis has widespread use in clinical laboratories. Here, mutation detection relies on restriction enzyme cleavage sites which are present either in the presence of absence of a given mutation in an amplified DNA fragment. In addition, traditional allele-specific PCR, also known as amplification refractory mutations system, based on primers which specifically hybridize and amplify only in the presence of a selected DNA variant at the 3’ end, is commonly used in many laboratories. These methods, though low throughput, have the advantage of being very inexpensive, relatively simple to perform, and easy to interpret, meeting the requirements of many clinical laboratories.

Invader™ Chemistry (Third Wave Technologies Inc, Madison, WI) is used in many diagnostic applications and uses isothermal hybridization of an “invading” probe into the DNA duplex followed by amplification and cleavage of fluorescence resonance energy transfer-based probes to detect the presence of a given genetic variant. This method is used in a number of laboratories, with the main drawback being a relatively limited test menu and relative difficulty of designing new assays using this technology.

Auxiliary methods to facilitate high throughput mutation analysis
Of note, many of the above methods only become “high throughput” when multiple patient samples can be tested in parallel. Key to this approach is the use of laboratory automation systems. One of the important elements is the use of automated DNA extraction techniques, because all of the above methods require purified genomic DNA. Manual methods for DNA extraction are laborious and require significant effort by laboratory personnel. Automated extraction leads to higher sample throughput with less hands-on time. Some of the above systems even can lead to all-in-one systems with automated DNA extraction as well as genotyping reaction preparation. Stand-alone automated liquid handling systems can also lead to higher sample throughput. These instruments can automate reagent pipetting and mixing, leading to decreased amounts of technologist time per assay, often one of the primary costs of any molecular test, and decreasing the scope for potential errors in reaction preparation. In addition, for any assay, instrumentation allowing multiple samples to be analyzed in parallel will lead to higher throughput analysis even if the fundamental reaction chemistry is identical.

Clinical applications of high throughput methods in pharmacogenomics
While the published literature describes the use of a multitude of different techniques for detecting particular genetic variants, many of these applications have only been explored in the research setting. The use of tests in the clinical laboratory requires extensive validation, quality control, and ongoing quality assessment to comply with regulatory mandates and to provide robust, accurate, and reproducible test results for patient samples. Such quality control and data assessment require the standard metrics of clinical validation including test performance compared to existing testing methods, presence and implications of false positives and false negatives, limit of detection, assay limitations, potential interferences, applicability of ongoing quality
control reference materials, indications for use, and clinical test interpretation. Although these validation parameters for laboratory-developed tests are widely used, there are no specific published guidelines for individual assays. The assays using next-generation approaches pose a particular challenge in this respect, given that most of these are currently performed in the research setting and the transition into the clinical setting is only recently coming about. The large amount of generated data in these assays poses a particular challenge, as does the identification of variants of unclear clinical significance. Specific validation guidelines for these novel types of assays are not yet available. In the US, some clinical laboratories only perform laboratory tests that have either been Food and Drug Administration (FDA)-cleared as showing substantial equivalence to a previously approved device or fully approved by the FDA as new in vitro diagnostic devices. This approval process can take months to years and requires a significant financial investment by the manufacturer, but the returns from widespread clinical utilization are expected to justify this upfront cost. In contrast, many other molecular diagnostics providers rely heavily on laboratory-developed tests that have been both designed and validated in a particular clinical laboratory. Some of these tests rely on commercially manufactured, but not FDA-cleared or approved, primers, probes, and other materials such as “analyte specific reagents,” whereas others are implemented with materials that are independently designed by the testing laboratory.

Below we briefly describe some of the pharmacogenomic test methods currently used in practice, concentrating on markers with known clinical utility. These methods are either described in the online test menu of various reference, commercial, and academic clinical laboratories or identified in the FDA in vitro diagnostic database. While this list is not meant to be comprehensive, it provides an overview of the spectrum of validated clinical applications of the molecular diagnostic methods discussed in previous sections.

**Warfarin metabolism/CYP2C9 + VKORC1 genotype**

Warfarin is an anticoagulant used widely to prevent thromboses in patients at elevated risk for such events. Consistent dosing is well known to be difficult and it has subsequently been found that SNPs within the **CYP2C9** gene, encoding one of the many CYP450 enzymes critical in hepatic metabolism, as well as those in the promoter of the **VKORC1** gene, encoding the warfarin therapeutic target vitamin-K epoxide reductase, are strongly related to warfarin metabolism. A recent study showed that incorporating genotype into algorithms for initiation of warfarin is necessary to normalize therapeutic drug levels across patients. Many potentially relevant markers have been identified in both **CYP2C9** and in **VKORC1**, though not all clinically available tests probe all markers. One FDA-cleared assay from TrimGen Corporation (Sparks, MD) incorporates multiplexed melt curve analysis targeting just the three polymorphisms with the most evidence of effect on warfarin metabolism, ie, the *2 and *3 variants of **CYP2C9**, related to warfarin resistance, and −1639G>A in **VKORC1**, related to warfarin sensitivity. Methods with medium levels of multiplexing for a single patient have also been implemented, including the Invader assay and the Luminex xTAG Technology assay. The Verigene® Warfarin Metabolism Nucleic Acid Test (Nanosphere) assay is FDA-cleared for this testing, as is the eSensor® Warfarin Sensitivity Test (GenMark Dx) targeted array. Of note, the ongoing European Pharmacogenetics of Anticoagulation Therapy trial incorporates a novel point-of-care, real-time PCR-based method to assess warfarin genotype at the initiation of therapy. Method comparisons have generally shown similar levels of accuracy between test methods, with the main differences occurring in turnaround time and cost.

**Clopidogrel metabolism/ CYP2C19 genotype**

Clopidogrel is a highly-prescribed antiplatelet agent used to decrease the risk of arterial thromboses. It has been shown that patients with particular variants in **CYP2C19** are at increased risk of thrombosis on a standard dose of clopidogrel. Pharmacogenomic testing has become more in demand since the inclusion of an FDA “black box” warning recommending genetic testing prior to prescription of clopidogrel, though the utility of this recommendation has recently been questioned in one study focused on patients with acute coronary syndromes or atrial fibrillation. Multiple allelic variants have been identified in **CYP2C19** which may be of clinical relevance, with the *1 (wild-type metabolizer), *2 and *3 (poor metabolizers), and *17 (ultra-rapid metabolizer) genotypes having the most supporting clinical data. However, some tests explore other SNPs that are rarely expressed, those that have less clear clinical relevance, or those that may relate to the metabolism of other drugs such as antidepressants. Given the necessity of probing multiple SNPs for each individual patient, multiplexed systems are widely used for this analysis. These include Luminex bead-based targeted arrays which can identify the *1, *2, *3, *4, *5,
CYP2D6 genotyping for multiple drug metabolism

The CYP450 enzyme system is one of the most important determinants of drug metabolism in humans. One of the central components of this pathway is the enzyme encoded by the CYP2D6 gene, which plays a role in metabolizing up to 25% of all available medications including the anticancer drug tamoxifen, antidepressants, antipsychotics, opioids, and beta-blockers, among others. More than 70 alleles, which lead to either enhanced or decreased enzymatic expression or efficiency, have been identified in this gene. For example, studies have found that patients with a CYP2D6 poor metabolizer phenotype demonstrate a higher probability of adverse drug reactions to the antipsychotics haloperidol and risperidone. While the clinical utility of such testing is still debated, the desire to genotype many CYP2D6 markers in parallel has primarily been met with targeted array technologies. The Roche AmpliChip is FDA-approved and offered clinically; it uses a microarray with differential hybridization to probe 27 SNPs in CYP2D6, along with insertions and deletions, as well as two SNPs in CYP2C19. The Luminex xTAG bead-based array system is FDA-cleared to probe a similar number of CYP2D6 SNPs.

Drug hypersensitivity

Some genetic variants relate to serious adverse effects of drug administration. Given the more limited clinical demand for these tests as compared to warfarin and clopidogrel genotyping, they are offered by fewer clinical laboratories and demonstrate less need for high throughput methodologies.

Irinotecan is a topoisomerase II inhibitor used as a chemotherapeutic agent for treatment of a number of solid tumors. Decreased enzymatic activity of the product of the UGT1A1 gene is associated with decreased excretion and toxicity, leading to potentially life-threatening neutropenia. A particular allelic variant, UGT1A1*28, consisting of seven TA repeats in the TATA box of the gene promoter, is most frequently implicated. In clinical testing, this variant is frequently identified by PCR and fragment size discrimination by capillary electrophoresis. Another approach is the Invader assay by Third Wave Technologies, which has been FDA-cleared and is offered clinically. Other variants in this gene also lead to decreased enzyme activity and can be investigated clinically by Sanger sequencing.

Thiopurine methyltransferase is an enzyme, encoded by the gene TPMT, shown to relate to the metabolism of thiopurine drugs used in the treatment of rheumatic disease, organ transplantation, and lymphoblastic leukemia. Deficiencies in enzymatic activity are related to severe hypersensitivity reactions including myelosuppression and gastrointestinal toxicity. Red blood cell enzymatic levels are typically measured prior to initiating therapy, but in some cases TPMT genotyping can be performed instead. TPMT genotyping prior to initiation of therapy has been used to successfully guide drug dosing and avoid adverse effects.

Multiple potentially significant allelic variants have been identified in TPMT, with the *2, *3A, and *3C genotypes most clinically relevant. Both traditional and real-time PCR methods are used for clinical testing, though these techniques cannot easily identify all TPMT variants.

Abacavir is a nucleoside reverse transcriptase inhibitor used to treat human immunodeficiency virus infection. Some patients using this drug develop severe, even life-threatening, hypersensitivity reactions. It was discovered that a large majority of patients with such reactions share a unique allele in HLA-B*5701, and genotyping for this allele has now become standard prior to starting therapy. Pretreatment screening for HLA-B*5701 has significantly decreased the incidence of hypersensitivity reactions. Similarly, hypersensitivity to carbamazepine, an antiepileptic drug, can lead to the life-threatening dermatologic complication of Stevens–Johnson syndrome. Recent studies have found that in East and South Asian populations, the HLA-B*1502 allele is highly correlated with the presence of such severe hypersensitivity reactions. As a result, it is now recommended by the FDA that testing be performed prior to administering carbamazepine to these patients. The most common methods used to identify both the HLA-B*5701 and HLA-B*1502 alleles include an initial PCR step followed by hybridization of sequence-specific oligonucleotide probes or allele-specific PCR.

Hepatitis C therapy

Hepatitis C is estimated to affect approximately 200 million individuals worldwide, frequently leading not only to liver failure but also to the development of hepatocellular carcinoma. It had long been observed that the rate of viral
Table 1 Comparison of current mutation screening technologies

<table>
<thead>
<tr>
<th>Technology</th>
<th>Parallel samples</th>
<th>Multiplex markers</th>
<th>Time to results*</th>
<th>Cost per test</th>
<th>Cost of set-up/instrumentation</th>
<th>Offered clinically for pharmacogenomic testing</th>
<th>FDA-cleared or FDA-approved diagnostic devices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole genome SNP array</td>
<td>Up to 48</td>
<td>&gt; 900,000 SNPs</td>
<td>1–5 days</td>
<td>High</td>
<td>High</td>
<td>Direct-to-consumer only (Affymetrix SNP Array)</td>
<td>No</td>
</tr>
<tr>
<td>Selected marker chip-based microarrays</td>
<td>Up to 8</td>
<td>3–2000</td>
<td>3–6 hours</td>
<td>High</td>
<td>High</td>
<td>Yes</td>
<td>Roche AmpliChip (CYP2D6/CYP2C19); Autogenomics INFINITI (CYP2C19; CYP2C9/VKORC1)</td>
</tr>
<tr>
<td>Selected marker bead-based arrays</td>
<td>Up to 48</td>
<td>10–35</td>
<td>4–8 hours</td>
<td>Medium to high</td>
<td>Medium to High</td>
<td>Yes</td>
<td>Luminex xTAG (CYP2D6); Nanosphere Verigene (CYP2C9/VKORC1); GenMark Dx (CYP2C9/VKORC1)</td>
</tr>
<tr>
<td>Sanger sequencing</td>
<td>Up to 384, typically &lt; 24</td>
<td>Multiple sequences up to 500–800 bp</td>
<td>2–12 hours</td>
<td>Medium to high</td>
<td>Medium</td>
<td>Yes</td>
<td>Majority LDTs</td>
</tr>
<tr>
<td>Pyrosequencing</td>
<td>Up to 96</td>
<td>Multiple sequences up to 300–500 bp</td>
<td>2–6 hours</td>
<td>Medium</td>
<td>Medium to high</td>
<td>Yes</td>
<td>Majority LDTs</td>
</tr>
<tr>
<td>Real-time PCR</td>
<td>Up to 384, typically 24 or 96</td>
<td>Up to 16, typically 1–3</td>
<td>1–3 hours</td>
<td>Low to medium</td>
<td>Low to medium</td>
<td>Yes</td>
<td>Some FDA-cleared, many LDTs</td>
</tr>
<tr>
<td>High resolution melt curve</td>
<td>Up to 1536, typically 32 or 96</td>
<td>Up to 10, typically 1–3</td>
<td>1–3 hours</td>
<td>Low to medium</td>
<td>Low to medium</td>
<td>Yes</td>
<td>Some FDA-cleared, many LDTs</td>
</tr>
<tr>
<td>Single base extension by mass spectrometry (Sequenom)</td>
<td>Up to 384</td>
<td>Up to 1500, typically &lt; 40</td>
<td>10 hours</td>
<td>Low to medium</td>
<td>High</td>
<td>?</td>
<td>No</td>
</tr>
<tr>
<td>Isothermal hybridization and cleavage (Invader)</td>
<td>Up to 96</td>
<td>Up to 15, typically 1–2</td>
<td>2–6 hours</td>
<td>Low to medium</td>
<td>Low to medium</td>
<td>Yes</td>
<td>Invader (UGT1A1)</td>
</tr>
<tr>
<td>Allele-specific PCR</td>
<td>Approx up to 24</td>
<td>Typically I</td>
<td>2–6 hours</td>
<td>Low</td>
<td>Low</td>
<td>Yes</td>
<td>Majority LDTs</td>
</tr>
<tr>
<td>Restriction fragment length polymorphism</td>
<td>Approx up to 24</td>
<td>Typically 1–2</td>
<td>2–10 hours</td>
<td>Low</td>
<td>Low</td>
<td>Yes</td>
<td>Majority LDTs</td>
</tr>
</tbody>
</table>

Note: "Time to results does not include nucleic acid extraction (2–12 hours), which is common to all methods above.

Abbreviations: LDT, laboratory-developed test; FDA, Food and Drug Administration; SNP, single nucleotide polymorphism; PCR, polymerase chain reaction; bp, base pair; approx, approximately.
clearance varied widely from patient to patient using the main treatment modality of ribavirin with pegylated interferon-α. Recently, multiple genome-wide association studies independently identified SNPs upstream of the *IL28b* gene as important predictors of response to therapy.\(^78^\text{-}^80\) Real-time PCR is the most common method for SNP identification in this context, with the majority of clinical laboratories testing for a single SNP, rs12979860, with a strong association to patient outcomes.\(^81^\text{-}^82\)

**Targeted oncologic therapy**

With the advent of newly targeted molecular therapies, selected genetic variants in cancers have been shown to be important indicators of treatment efficacy. However, compared to other pharmacogenomic variants described above, the sensitivity limit of detection becomes much more important as the mutant DNA is typically present in a background of normal cells. For germline DNA, a heterozygous mutation will be represented by 50% of alleles; in cancers, the relevant mutation may be present in under 1% of the examined DNA (or mRNA).\(^83\)

In the treatment of non-small cell lung cancer (NSCLC), activating mutations in the tyrosine kinase domain of the *EGFR* gene are targeted by recently introduced small molecule inhibitors, erlotinib and gefitinib. The most common mutations, found in >90% of cases, are either a deletion present in exon 19 or the point mutation L858R in exon 21.\(^84\) For evaluation of these two specific alterations, many laboratories use a combination of traditional PCR and restriction fragment length polymorphism followed by capillary electrophoresis. Melt curve analysis on the LightCycler is also used for point mutation analysis in exon 21. Sensitivity of these methods needs to be established by each individual laboratory, but may reach approximately 1%–10% of mutant DNA. For evaluation of all described *EGFR* mutations,\(^84\) Sanger sequencing is offered widely to examine all of exons 18, 19, 20, and 21. While the advantage here is detection of many possible mutations, the drawback is the lower sensitivity of sequencing, which requires approximately 20% mutant DNA for accurate detection.

Downstream mutations in the *KRAS* gene have been shown to eliminate the therapeutic benefits derived from the epidermal growth factor receptor inhibitors above in NSCLC as well as the antiepidermal growth factor receptor antibodies used in colorectal cancer.\(^5\) Many alternative approaches are available to identify the most common mutations in *KRAS* codons 12 and 13.\(^85\) These include PCR with single base extension and fluorescent detection by capillary electrophoresis using the Muterctor™ reagents from TrimGen (distributed by Life Technologies), allele-specific real-time PCR using the TheraScreen® reagents from Qiagen Manchester Ltd (Manchester, UK), Sanger sequencing, and pyrosequencing. The sequencing modalities, while typically less sensitive, allow additional clinically significant mutations to be detected, particularly those present in exon 61 of *KRAS*.\(^85\)

Activating mutations in the *BRAF* gene also are of clinical importance in NSCLC as well as in many other tumors, including melanoma and papillary thyroid carcinoma. A highly promising small molecule against melanoma, vemurafenib, specifically targets the *BRAF* V600E mutant.\(^86\) Different methods have been applied to detecting this mutation in *BRAF*, including traditional allele-specific PCR, allele-specific real-time PCR, and real-time PCR with melt curve analysis. While V600E accounts for over 90% of mutations in *BRAF*, others have also been described\(^87\) and can be clinically detected by sequencing methods, such as Sanger sequencing or pyrosequencing.

Of note, a multiplexed panel-based approach using single base pair extension genotyping and SNaPshot reagents is currently used clinically to simultaneously examine mutations in 13 different cancer genes in solid tumors.\(^88\) This panel-based approach may demonstrate additional utility as more targeted therapies become available. Looking forward, as next generation sequencing technologies become widely available in clinical practice, highly parallel resequencing of tumor exomes or targeted genomic regions could lead to the rapid and sensitive identification of numerous genetic variants in a single test.

**Implications for pharmacogenomics and clinical diagnostics**

Given the regular publication of new genome-wide association study analyses, our knowledge of potentially relevant pharmacogenomic markers is constantly evolving. Simultaneously, our ability to custom-tailor clinical analysis of genomic variants continues to improve. Optimization of automation and instrumentation will allow for large numbers of patient samples to be analyzed in a high throughput fashion.

However, while high throughput methodologies for detecting pharmacogenomic markers currently exist, many of these methods have only been implemented in the research setting.\(^89\) As evidenced by this review, for clinical laboratories, current levels of sample demand do not necessitate testing...
hundreds or thousands of samples in parallel. In addition, the clinical utility of most highly multiplexed tests still remains debated. Therefore, lower-cost, lower-throughput methods may suffice at present.

Uncertainty regarding the utility of pharmacogenomic testing poses many complexities for the clinical laboratory, particularly in terms of laboratory economics. New publications are continually arising with the claim that a new SNP or group of SNPs is important in the metabolism or response to a given drug. However, many of these early suggestions from pilot studies do not ultimately prove to be clinically relevant when subjected to larger, more focused trials. The introduction of a new assay for pharmacogenomics-based testing may require the purchase of capital-intensive equipment, expensive reagents, and necessitates additional training of laboratory staff. Therefore, validation of a molecular diagnostic assay for a marker that ultimately proves to be of low clinical utility does not only fail to enhance patient care, but may also represent a significant financial burden and insurers may decline to pay for such testing. In addition, the large majority of clinicians have been slow to incorporate pharmacogenomic testing into their practices. Clinical laboratories will be unlikely to bring these tests online without appropriate demand.

In the absence of significantly increased demand most clinical laboratories may choose not to pursue high throughput analysis methods. Instead, laboratories may look to methods with low or intermediate throughput for the time being, or send pharmacogenomic testing to specialty reference laboratories which can take advantage of economies of scale. A competing force is the potential role of point-of-care molecular testing. One could imagine in particular a utility for pharmacogenomic testing in this context: genotyping could be performed, the results assessed, and the decision to prescribe a given medication made, all as part of one patient visit. These miniaturized and inexpensive point-of-care PCR systems with simple operation and rapid turn-around time are in various stages of development. Although the use of these instruments is in principle appealing, their convenience belies the complexity of the underlying pharmacogenetic reality and its interpretation.

Also emerging is the likely prospect of full human genome sequencing for under US$1000. While many significant hurdles remain to be overcome, we may one day reach a point where the majority of patients have already had their genome fully sequenced before ever being prescribed a medication.

Future directions/Conclusion
The detection of genetic variants is clearly a rapidly evolving field of clinical diagnostics. The most important advances in this field will come from a better understanding of the clinical data rather than further improvements in the technologies for mutation detection alone. Nevertheless, the development of new technologies will in many ways influence and define the role of clinical laboratories with the advent of new and increasingly clinically integrated pharmacogenomic testing. These upcoming changes, encompassing both more complete clinical validation of target biomarkers as well as more widespread implementation of genomic analysis, will greatly affect the options for pharmacogenomic testing. Clinical laboratories must remain central to the health care team through consultation with treating physicians regarding genomic data as these methods become more widespread. One of the most important future developments will be the widespread adoption of pharmacogenomic testing into clinical decision algorithms. At this point the sample demand is likely to increase to where it is economically and practically feasible for a larger number of clinical molecular diagnostic laboratories to incorporate high throughput genome-based testing with a direct impact on clinical patient care.

Disclosure
The authors report no conflicts of interest in this work.

References


