



DNA Technologies in Precision Medicine and Pharmacogenetics

8

Seyedeh Sedigheh Abedini, Niloofar Bazazzadegan,
and Mandana Hasanzad

What Will You Learn in This Chapter?

A wide range of biological fields, including medicine and precision medicine, can benefit from next-generation sequencing. DNA technologies and their applications to pharmacogenomics, and drug prescribing in the context of genome genotyping, are discussed in this chapter. This chapter discusses a variety of NGS technologies, including whole-genome sequencing (WGS), clinical exome sequencing (CES) and whole-exome sequencing (WES), whole transcriptome sequencing (WTS), targeted sequencing (TS), single-cell sequencing (SCS), and DNA microarrays and their clinical and medical applications.

Rationale and Importance

Due to the genetic diversity, some individuals might show unexpected side effects and even drug resistance. Therefore, the genetic profile of these patients must be analyzed to determine molecular biomarkers and genetic data for prescription medicine. DNA technologies are

enabled to elucidate the profile of the human genome, which could result in improved drug treatments. Pharmacogenomics (PGx) is a critical component of personalized medicine since genomic information enables the development of safer, more effective, and more affordable drugs. Because of their low cost and accuracy, genotyping technologies like next-generation sequencing (NGS), microarrays, and bead arrays are expected to make their way into clinical application. These techniques provide a lot more information than other types of genetic testing, which could be extremely useful when attempting to figure out what is wrong with a patient.

Researchers and medical professionals will be able to use the NGS on precision medicine routinely due to developing chemistries, lowering cost, and the newest tools available to facilitate the analysis of genotyping data based on PGx. As a result, we expect genotyping technologies to have strong capabilities and more guidelines for personalized medicine and pharmacogenetics in the next years, leading to improved healthcare.

S. S. Abedini (✉) · N. Bazazzadegan
Genetics Research Center, University of Social
Welfare and Rehabilitation Sciences, Tehran, Iran

M. Hasanzad
Personalized Medicine Research Center,
Endocrinology and Metabolism Clinical Sciences
Institute, Tehran University of Medical Sciences,
Tehran, Iran

Medical Genomics Research Center, Tehran Medical
Sciences, Islamic Azad University, Tehran, Iran

8.1 Introduction

Many studies indicate that drug-related genes, also referred to as “pharmacogenes,” in the human genome contain extensive functional genetic variations (FGVs). Different alleles are associated with diverse outcomes of drug treat-

ments [1–3]. Genetic variations and level of expression in drug-targeted molecules, containing membrane and nuclear receptors, signal transduction components, and enzymes, moreover drug transporters and drug-metabolizing enzymes may affect the incidence of the individual variations in response to a drug [4]. Around 97–98% of people have at least one actionable FGV in their drug-related genes. In addition, the possibility of the presence of a genetic variant that could result in a loss of function (LOF) variant in pharmacogenes is 93% for every individual [5]. Hence, identifying the different genetic variants associated with the drug metabolism would affect medication prescription, allowing for selecting the right drug and dose, thereby reducing the potential adverse effects or therapeutic inefficacy.

Breakthrough of NGS platforms throughout this decade now provides affordable and reliable high-throughput sequencing for assessment of functional DNA variations in many diseases, including both monogenic and polygenic phenotypes such as diabetes, cancer, and cardiovascular and neurological disorders as well as in the regulation of physiological conditions such as height, blood pressure, and body mass index [6–11]. Thus, there is an increasing excitement to apply individual genome sequencing for predicting disease risk, lifelong well-being of individuals, medical care, and response to drugs in the era of personalized medicine. Currently, the area of PGx is shifting from reactive testing of a single gene toward scanning a whole panel of genes concerned with drug absorption, distribution, metabolism, and excretion (ADME) before prescribing (preemptive genotyping) by using different types of next-generation sequencing (NGS) platforms [12]. DNA technologies have been used to detect variants that affect the drug's toxicity and efficacy. The properties of NGS technolo-

gies make them an exciting approach to performing clinical PGx testing. Several investigators have recently explored approaches utilizing NGS platforms, namely, targeted sequencing, whole-exome sequencing (WES), and whole-genome sequencing (WGS) in pharmacogenomics. Microarrays enable gene expression profiling of thousands of genes in tens of samples by research. Also, different gene clusters showed a correlation with discrete phenotypes in tumors offering that tumor grades are associated with distinct gene expression [13].

8.2 NGS Technology Procedures: Its Important Applications and Related Different Databases

Figure 8.1 indicates NGS steps, improvement in biochemical steps, the kind of machine types, performance of each phase, mechanism of preparation, and how each step is carried out for complete huge parallel sequencing [14–16].

NGS has diverse biological applications; nevertheless, the applications in precision medicine have extended tremendously in the last few years. Figure 8.2 represents some DNA technology applications in the field of precision medicine [17].

A growing number of databases can be used for showing disease associations. Population databases supply information concerning the frequencies of variants in populations. These databases might contain healthy and diseased patients and must be used cautiously to rely on the purpose of the association study. On the other hand, the disease databases include variants in patients suffering from a particular disease and relevant information about its pathogenicity. An outline of the most used databases is provided in Table 8.1.

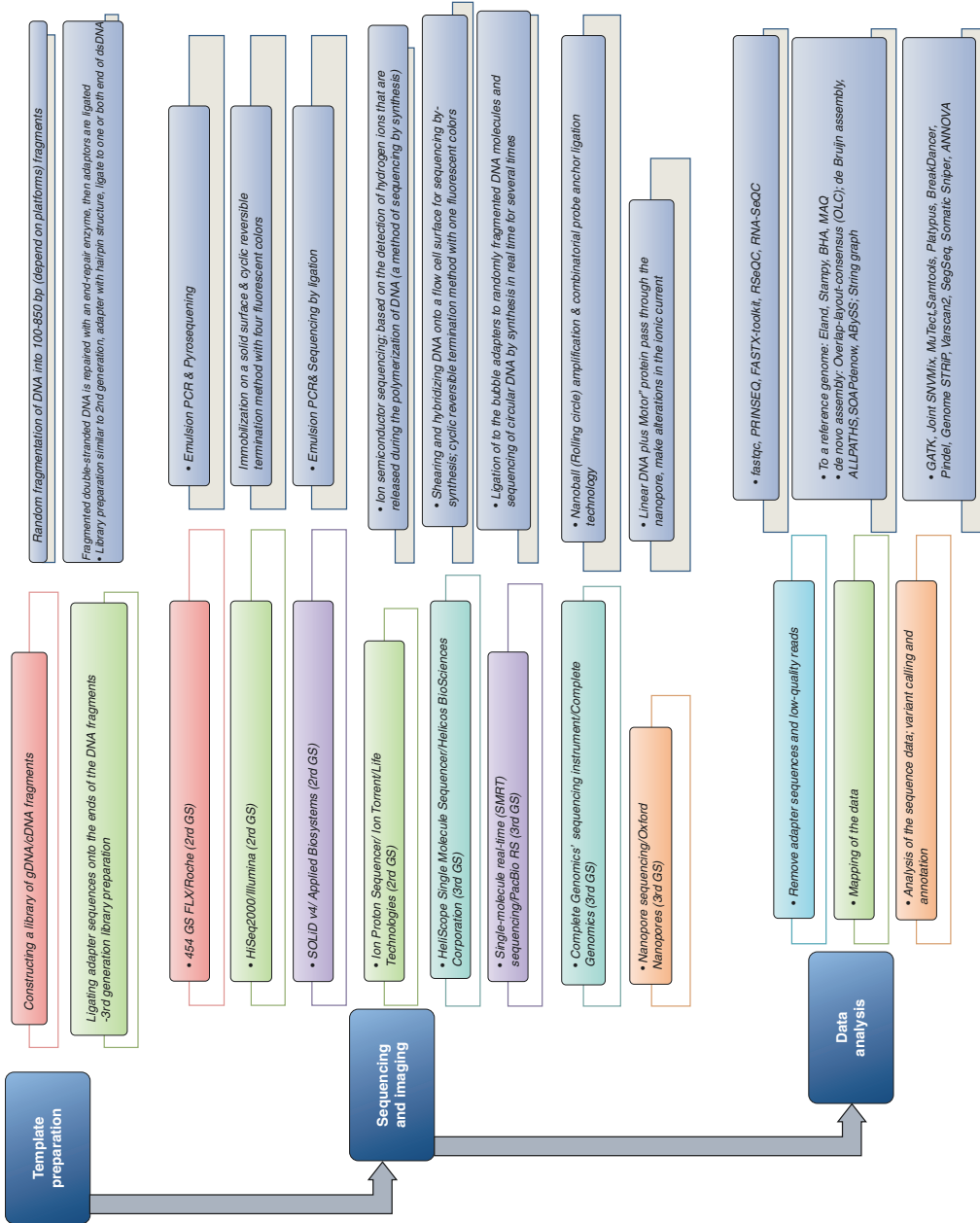


Fig. 8.1 NGS step, development in biochemical steps, machine types, the performance of each step, and mechanism of preparation and how each step is performed for complete massive parallel sequencing is shown. “This figure is reprinted with permission from source Rabbani B, et al. Mol Biosyst. 2016. PMID: 27066891 Review”

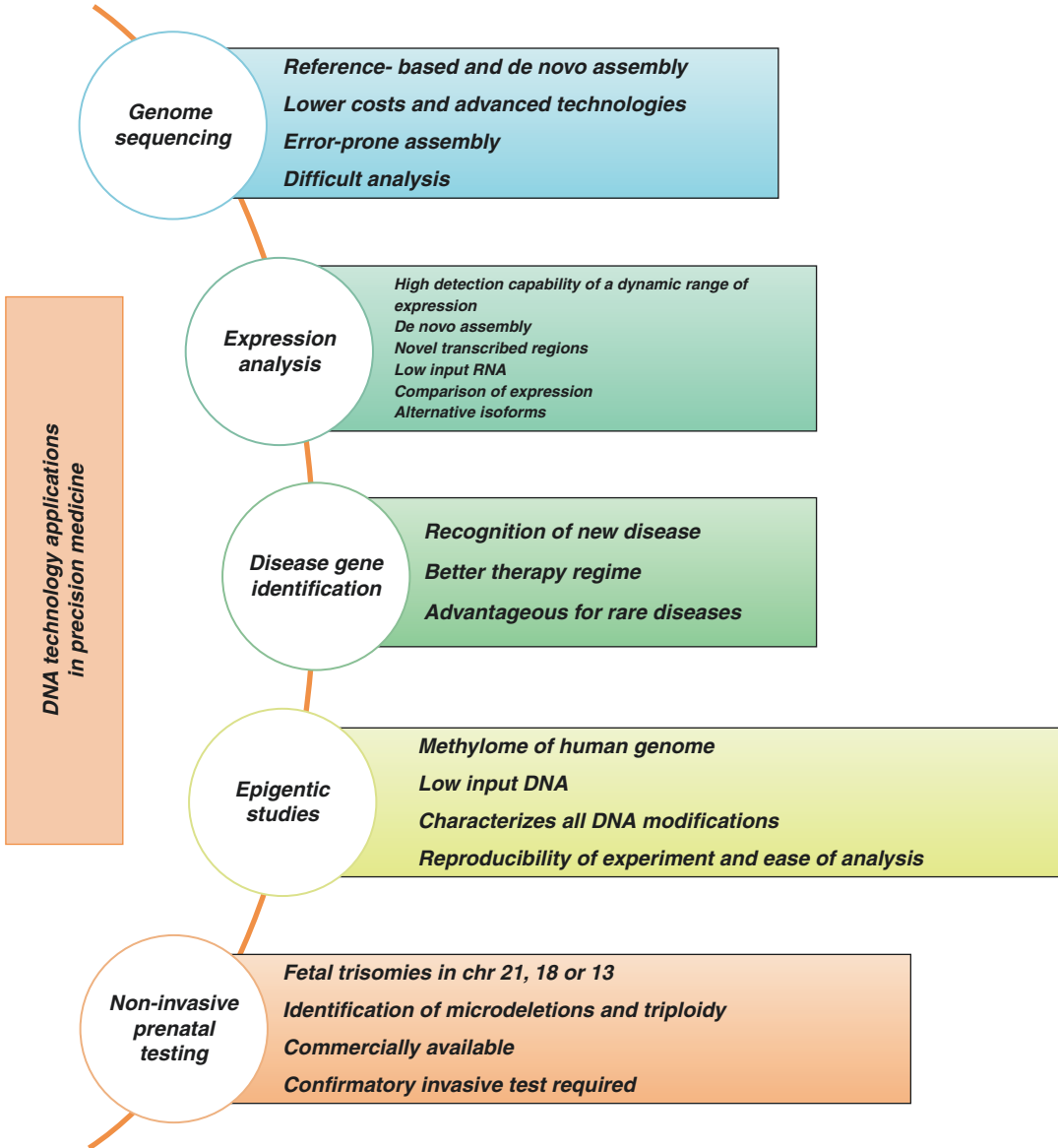


Fig. 8.2 Overview of some major applications of next-generation sequencing

Table 8.1 Summary of useful databases in medicine and pharmacogenomics

Databases	Description
Data interpretation and genomic test databases	
Evidence aggregator http://64.29.163.162:8080/GAPPKB/evidenceStartPage.do	An application that makes finding evidence reports, systematic reviews, and guidelines for using genetic tests and other genomic applications easier
FDA Pharmacogenomic Biomarkers http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm	A list of FDA-approved medications that include pharmacogenomic data on the label
PharmGKB http://www.pharmgkb.org	Clinical information, such as dose guidelines and prescription labels, as well as potentially clinically actionable gene-drug connections and genotype-phenotype interactions, are all included in this pharmacogenomics knowledge repository
EGAPP http://www.egappreviews.org/about.htm	A collection of information on the validity and value of genetic testing in clinical practice, as well as advice from professional organizations and advisory committees on how to utilize genetic tests
GAPP Finder http://64.29.163.162:8080/GAPPKB/topicStartPage.do	A searchable collection of genetic testing and genomic applications that are in the process of transitioning from research to clinical and public health use. Disease, genes, drugs, tests, and other terms can be used in the search query
Genetic testing registry http://www.ncbi.nlm.nih.gov/gtr/	Information about test technique, validity, proof of the test's utility, and laboratory contacts and credentials are available in one central spot
Locus/disease/ethnic/other-specific databases	
DECIPHER http://decipher.sanger.ac.uk	DECIPHER provides information on over 19,000 patients (with more on the way) and allows for phenotype-genotype comparisons
Human Genome Variation Society (HGVS) http://www.hgvs.org/locuspecific-mutation-databases	HGVS gives data on variants that are unique to the locus of interest
Leiden Open Variation Database (LOVD) http://www.lovd.nl	LOVD is an open source for analyzing sequence variants and storing NGS data
Sequence databases	
Locus Reference Genomic (LRG) http://www.lrg--sequence.org	EBI and NCBI assemble and preserve LRG sequences, which have documented mutations with a permanent ID
NCBI Genome http://www.ncbi.nlm.nih.gov/genome	This NCBI database contains genome-related information like sequences, chromosomes, and annotations
RefSeqGene http://www.ncbi.nlm.nih.gov/refseq/rsg	By offering sequences that may be used as reference standards, RefSeqGene assists researchers in matching their data
MitoMap-Human Mito Seq http://www.mitomap.org/MITOMAP	The database contains data on human mitochondrial DNA variation from over 31,000 human mitochondrial DNA sequences, both published and unpublished (with more on the way)

(continued)

Table 8.1 (continued)

Databases	Description
Disease databases	
Human Gene Mutation Database (HGMD) http://www.hgmd.org	HGMD is a licensed database that contains up-to-date information on mutations that cause human hereditary illnesses
Clin Var http://www.ncbi.nlm.nih.gov/clinvar	ClinVar is a free resource that contains information on genetic variants and phenotypes
Online Mendelian Inheritance in Man (OMIM) http://www.omim.org	OMIM has up-to-date information on over 15,000 genes and their genotype-phenotype relationships
Population databases	
VarSome https://varsome.com/	A search engine, aggregator, and impact analysis tool for human genetic variation and a community-driven project aiming at sharing global expertise on human variants
dbSNP http://www.ncbi.nlm.nih.gov/snp	SNPs (single nucleotide polymorphisms) and other small-scale variants are stored in the dbSNP database
Exome Variant Server http://evs.gs.washington.edu/EVS	Through next-generation sequencing data sets, this database provides information on new genes and pathways that contribute to heart, lung, and blood illnesses
Exome Aggregation Consortium (ExAC) http://exac.broadinstitute.org/	For the scientific community, this consortium provides exome sequencing data for over 60,000 individuals from large-scale sequencing efforts
dbVar http://www.ncbi.nlm.nih.gov/dbvar	dbVar collects data on structural variants such as translocations, duplications, deletions, insertions, and more
1000 Genomes Project http://browser.1000genomes.org	This database is an extension of the Ensembl human genome browser, and it shows data on genetic variation in a large number of people

8.3 Whole-Genome Sequencing

WGS is the sequencing of an entire organism's genome at a single time. WGS includes the sequencing of **chromosomal DNA**, **mitochondrial DNA**, and chloroplast DNA in plants [18]. Complete genomic variants (including PGx-related markers) for an individual would be available by utilizing the WGS approach. WGS was introduced to clinics in 2014 and has been chiefly used as a research tool [19–21]. Although the widespread data interpretation of such tests is still challenging, a reduction in sequencing costs alongside the comprehensiveness of WGS may also result in the method turning into a widespread platform for clinical PGx tests.

Using phase I WGS data from the 1000 Genome Project followed by annotations, the variant minor allele frequency was >1%, of which 8207 resulted in strong linkage disequilibrium (LD) ($r^2 > 0.8$) with known PGx variants. Differences were distributed in various genome components, introns, coding, and 5'-upstream and 3'-downstream regions. Finally, the authors identified putative functional variants within the known pharmacological genomics loci underlying the drug response phenotype and suggested direct testing instead of relying on LD, which will be different among populations [22].

Yang et al. conducted a three-way analysis with the Directorate of Medical Education and Training (DMET), WES, and WGS, to examine the concordance between PGx genotyping calls based on these various technologies. They showed a 94% concordance between the DMET and WES and a 96% concordance between the DMET and WGS [23]. The functional copy number variation (CNV) of the ADME genes was distributed in different populations at significantly different frequencies [24, 25]. NGS data can also be used for CNV calls with different ethnic backgrounds.

Researchers used integrated WGS and WES data from 1000 genomes and ExAC repositories for CNV identification of 208 pharmacogenes. Novel CNVs (deletion in 84% and duplications in 91% of genes) over six distinct populations of non-Finnish Europeans, Africans, Finns, East

Asians, South Asians, and mixed Americans were decoded effectively. The ultimate result highlighted the need for the comprehensive NGS-based genotyping of the pharmacogenes for the CNV distinguishing proof nearby their allele frequencies. The evaluation of the commitment of such CNVs to the medicate response results is additionally conceivable through a population-specific analysis of uncommon variants [26].

WGS has been performed on a patient with a family history of vascular disease and sudden early death in a study. However, no clinically significant medical records predict the potential risk of coronary artery disease and the cause of sudden cardiac death [27]. Rare variants of three genes, including *TMEM43* (MIM # 612048), *DSP* (MIM # 125647), and *MYBPC3* (MIM # 600958), have been found to be clinically associated with sudden cardiac death. This patient was heterozygous for a null mutation in the *CYP2C19* (MIM # 124020) gene, suggesting possible resistance to clopidogrel. The authors suggested that WGS could provide helpful and clinically relevant information for individual patients. Knowledge of pharmacogenetic variants may be essential for future personalized medicine of patients. Overall, personal genomic analysis is a field of genomics that ultimately provides experimental medical treatment to individuals based on genomic analysis. Predictive and preventive care results in advanced healthcare.

WES and WGS may give a promising approach to recognize low-frequency (1–5%) and uncommon (<1%) variations. The suitable medicate reaction loci with a genome-wide approach can be found instead of finding one gene [28]. Researchers at Washington University have utilized WGS to analyze a challenging leukemic disorder, appearing that this information can be assembled and analyzed within a time frame consistent with clinical decision-making [29].

Most WGS is generated using methods that the user can modify concerning laboratory standards [30]. Many scientists feel that their responsibility is to modify sequencing protocols to provide the best sequencing results possible. Just as microarray data depends on the method used

to isolate mRNA and generate labeled cDNA [31], WGS results depend on changes made to the protocol developed, tested, and validated by the manufacturer [32, 33]. This variability leads to significant differences between WGS produced by different laboratories [32].

8.4 Whole-Exome Sequencing

Whole-exome sequencing (WES) is one of the primary applications of high-throughput DNA sequencing methodology for detecting different variations in the coding sequence, and additional relevant adjacent and untranslated regions of the genome. WES is a progressively critical technology and molecular diagnostic tool in rare diseases and drug response genetics [34, 35].

Since 2011, WES as a useful diagnostic tool has been routinely offered in clinical genetics laboratories [36]. Then, it has been consolidated into National Heart, Lung, and Blood Institute (NHLBI) “Grand Opportunity” Exome Sequencing Project (GOESP) (more than 6500 patients), DiscovEHR study (functional variants in 50,726 human genome), the 1000 Genome Project (variants in 1092 individuals) and the Exome Aggregation Consortium (ExAC) projects (pathogenic variants in 60,706 individuals) [37–41]. Distinct evaluation of the WES data in the different populations indicates the frequency and potential functional association of rare variants, almost novel SNVs, among many pharmacogenes [40, 42]. The recent studies especially phase I and II drug transporters and metabolic enzymes consisting of exome sequencing and SNVs data revealed that approximately 93% of all identified variants are rare (minor allele frequency [MAF] < 1%) or very rare (MAF < 0.01%) [42]. Another study, involving 14,002 subjects for investigation of a rare genetic variant by sequencing of 202 drug target genes, explored that almost variants have MAF below 0.5% and had not been previously identified. In addition, many of these variants are harmful which are associated with risk factors for developing a disease and drug response. Also, at least 5–10% of them have a critical role in the PGx panel [43].

Through the years, WES has dramatically made improvements in the robustness of research procedures and laboratory tests, dataset uniformity, and advances in filtering and interpretation of variants [44]. To date, greater than 80–85% of pathogenic mutations in Mendelian disorders and complex disorders have been detected within the exomes, and which WES technique offers a fair-minded approach to identify these variations and provides additional information about them in the era of personalized medicine and PGx profiling [14, 43–47]. This information leads to achieving maximum benefits for particular patients by adapting unique treatment based on genetic makeup. For example, affected individuals with pyridoxine-dependent epilepsy-*ALDH7A1* (OMIM ID, 266100) are commonly resistant to therapy with anticonvulsants; however, using massive doses of pyridoxine (vitamin B6) can be treated efficiently [48]. Prevention of futile drug use and consideration of the treatment strategy in these non-insensitive patients are done according to WES data as “game-changing technology” [48, 49].

WES technology maturity and process standardization exhibit impressive improvements in the different eras of personalized medicine including targeted therapeutic agents based on tumor biomarkers like PD-L1, larotrectinib (Vitrakvi) and olaparib, vemurafenib (BRAF-positive tumors), imatinib (KIT-positive tumors), and monoclonal antibody pembrolizumab in which all of them have been approved by the US Food and Drug Administration (FDA) [50–53]. It is anticipated that the digital genome market and the personalized medicine market will attain over \$45 billion and \$87.7 billion by 2024, respectively [35].

In recent years, the potential effectiveness of WES has been additionally studied for a wide range of pharmacogenes’ profiling, drugs’ pharmacokinetics (PK), and pharmacodynamics (PD) from absorption to excretion of certain diseases, such as nervous system, seizures, kidney transplantation, cancer, infectious diseases, and autoimmune disorders [40, 54, 55]. The first stage of xenobiotics metabolism involves cytochrome P450 (CYP450 from the CYP1 enzyme family) activities that may change by way of genetic vari-

ants positioned in their related genes. Therefore, identifying the genetic variation using drugs' PK and PD helps the clinician to choose suited therapeutic without toxicity [35, 56]. As a result, WES data has the ability to revolutionize the prevention and even therapy of human disease. In addition to the prediction of common drug reactions, genetic information and WES data are also used to select small molecular inhibitors and analyze different variants (somatic and germline) in various diseases.

According to WES analysis study performed by Van der Lee and his colleagues for providing the PGx panel with actionable Ubiquitous Pharmacogenomics (U-PGx; www.upgx.eu) panel, Clinical Pharmacogenomics Implementation Consortium (CPIC), and Dutch Pharmacogenetics Working Group (DPWG) guidelines, 39 out of 42 variants (86% of total) in 11 pharmacogenes represent linking genotype to drug response phenotypes. Recently, more than 21 important genes and 50 drugs have been recommended by a combination of CPIC and the DPWG. This data highlighted the ability of WES data to create a significant PGx panel based on critical pharmacogenes (7 out of 11 genes). However, this group did not identify any structural variations (SVs) in *CYP2C19*, *UGT1A1*, *CYP3A5*, and *CYP2D6* genes due to a limited sample size [57]. Proper coverage of variants of the *CYP2D6* gene is clinically important because the *CYP2D6* enzyme accounts for 25–30% of commonly prescribed drugs. Mutations in the *VKOR* and *CYP2C9* genes lead to different metabolic capacities of the coding enzyme in the drug metabolism pathway [57].

The PGx panel is only informative when specific drugs are used and are expounded regarding each patient's genetic information and their medicine desires. Therefore, the results of PGx are beneficial when gene-drug interactions are considered and completely interpreted according to the obtained sequencing data of each patient. This genotype-phenotype correlation substantially decreases the risk of revealing unsolicited finding and accelerates the treatment processing.

Cousin and his colleagues revealed that a considerable percentage of patients had actionable

PGx profiles based on current drug intake. They investigated 94 patients for PGx variants in the three important pharmacogenes (*CYP2C19*, *CYP2C9*, and *VKORC1* genes) using WES data and detected at least one actionable variant in 91% of all subjects. Twenty percent of total patients showed an immediate impact on current medicinal drug use (warfarin and clopidogrel) through the PGx finding [58]. The proper interpretation of PGx variants in this study was the key to inhibiting drug adverse effects and making individualizing prescribing decisions.

Several studies have been carried out to determine the accuracy and the concordance rate of WES technology in PGx and its application in precision medicine. Rennert et al. investigated 337 cancer patients with Exome Cancer Test v1.0 (EXaCT-1), and causative genetic mutation has been detected in 82% of all cases. The results suggest accurate cancer treatment and provide utilized information for precision medicine cancer care. The positive predictive value, specificity, and sensitivity were 99.2%, 99.9%, and 95.7%, respectively. This emphasizes the accuracy of WES for mutation detection and prescribed medications with improvement in saving the cost and time [59]. Yang et al. simultaneously examined three technologies, clinical genotyping (DMET array-based), WES, and WGS, for comparing PGx variants obtained from sequencing of 13 valuable pharmacogenes with ICI guidelines. The contradiction genotyping was observed between 4 out of 68 loci by DMET and WES and 3 out of 66 loci by DMET and WGS. They reported the concordance rate between WES and DMET and WGS and DMET is 94% and 96%, respectively. They confirmed that WES and WGS are capable of providing worthy and usability data for most pharmacogenes and prepare further validation of genomic sequences in clinical laboratories [23]. Another study for the assessment of the WES variant's integrity was performed by Chua et al. They used cross-comparison between the MiSeqR amplicon sequencing data and WES for two important pharmacogenes: *CYP2D6* and *CYP2C1*. They indicated the error rate is less than 1% and WES is a pioneer tool in providing PGx profiling, even if complex loci have been

studied [45]. Other researchers have published similar results that the most useful outcomes are obtained from sequencing data compared to orthogonal tests [40, 52, 54, 57].

The improvement of WES accuracy and its cost make it as a usable molecular diagnostic tool for the evaluation of genetic disorders and pharmacogenetic tests. However, the obtained WES variants, read length, depth of coverage, and variant interpretation regarding the PGx panel for each patient to avoid any futile drugs should be considered in more detail.

8.5 Clinical Exome Sequencing

The clinical value of WES and WGS as a general test for mutation findings is now appropriate for the almost genetic diagnostic query. Although whole-exome sequencing and whole-genome sequencing are emerging, panel-based testing (based on clinical) is more practical for clinical annotation in the human genome and has a strong position in precision medicine. Clinical exome sequencing (CES) has become more viable – and possibly cost-effective – as a first-line diagnostic, rather than an alternative to exploring if other types of testing fail to offer a diagnosis, due to the rapid improvement of high-throughput sequencing technology in speed and cost. This approach concentrates on genes in which disease-causing mutations have been discovered and documented in the Human Mutation Database. Ambray Genetics (Aliso Viejo, CA, USA) was the first CLIA laboratory to use NGS technology for establishing a “Clinical Diagnostic Exome” in 2011 [34, 47, 60].

In comparison to WES and WGS, the CES dataset is substantially smaller, but it offers several advantages: firstly, not generating excessive numbers of uncertain significant variants, which simplifies genetic counseling; secondly, putting by the emphasis on clinically indicated genes, achieving trio analyses, and obtaining high-quality data (deep coverage) which is cost-effective; and finally, using an instrument such as the Illumina MiSeq, which can be used at a benchtop scale for data analyzing [17]. Therefore,

it helps to facilitate the identification of actionable variants for applicability in precision medicine and therapeutic decision-making. Many firms have offered different panels for some genetic disorders such as cancer, hearing loss, and cardiomyopathy that are used by researchers and clinicians. To date, different panels have been established from actionable gene panels, hotspot panels, and disease focus panels to comprehensive multigene panels. The panels will allow us to detect genetic variants responsible for diseases and predict treatment regimens that will be effective, leading to better and more prompt patient management. More recently, the CES application was mainly used for determining the risk of hereditary malignancies and drug decision-making for somatic cancers [60, 61].

Using a hotspot panel (comprised of common hotspot mutations), clinicians can identify mutations in regions of the genome relevant for treatment, diagnostics, or prognosis. The first commercially available hotspot panel was the AmpliSeq cancer panel V1 which covers 46 cancer genes (tumor suppressor and oncogenes) with 739 actionable mutations. The number of hotspot mutations is increased to 2855 from 50 cancer genes in the new version of this panel. In contrast with hotspot panels, actionable gene panels cover all exon and targeted genes to identify other harmful mutations outside of hotspot variants. The most common target genes of these panels are FDA-approved genes such as *BRAF*, *PIK3CA*, *KIT*, *ALK*, *NRAS*, *KRAS*, and *EGFR*. The TruSight Tumor panel was the first commercial actionable gene panel that covers 26 genes involved in melanoma and ovarian, gastric, lung, and colon cancers. The majority of these panels look into somatic mutations to help determine therapeutic options. Unlike the gene panels, the disease focus panels are mostly focused on identifying inherited diseases or detecting suspected genetic disorders based on germline mutations. The sensitivity, specificity, and depth of coverage can be increased via these panels consisting of a limited set of genes, while the cost is reduced [17, 34, 36, 47, 60–62].

While disease-specific panels have become increasingly popular, some potential barriers are

propounded: (1) the limited quantity of samples for clinical testing, (2) the process of developing and validating the panels according to ACMG guidelines, and (3) the need to keep current panels up to date. These challenges have led investigators and clinicians to explore the new utilizing panel, comprehensive panels, which include different actionable genes associated with their related disorders. By using this panel, disease-specific testing would be simplified, while the medical significance of most variants could avoid interrogation. More than 60 valuable genes with 4813 causative genes for genetic disorders have been listed in Illumina's TruSight panels as a known and popular comprehensive multigene panel [60, 61].

Despite the potential advantages of CES for patients whose diseases are undiagnosed or whose results from disease-focused panels are inaccurate, we do not expect the full scale of this approach in clinical trial tests due to the limitation of the restricted number of specified genes and complex bioinformatics pipelines.

8.6 Whole Transcriptome Sequencing

Transcriptome sequencing, or gene expression arrays in general, has been a well-established diagnostic tool for characterizing and quantifying gene expression profiles and detecting fusion transcripts. Improvements in RNA sequencing (RNA-Seq), including polyA selection and WTS, can be used to develop analytical spectra that cover multiple transcriptional events (chimeric transcripts, isoform switching, expression, etc.) in a single approach. RNA-Seq provides single base-pair resolution and significantly less background noise, enabling distortion-free transcriptome evaluation compared to expression arrays [63]. Currently, the diagnosis of acute lymphoblastic leukemia patients requires several analyses encompassing morphology, immunophenotyping, molecular evaluation of gene fusions and mutations, and detection of numerical and structural abnormalities based totally on chromosomal banding analysis and

fluorescence in situ hybridization [64]. WTS parallel analysis of gene expression profiles allows for changes in fusion transcripts and copy numbers, leading to the specific characterization of patients' genetic profiles as the basis for disorder classification based on a single method dataset. Understanding the transcriptome is necessary to interpret the genome's functional elements and apprehend the underlying development and disease mechanisms.

Advancements in large-scale parallel DNA sequencing technology have enabled transcriptome sequencing (RNA-Seq) through cDNA sequencing. RNA-Seq quickly replaced microarray technology due to its high resolution and reproducibility. This method can be used to expand knowledge about alternative splicing events [65], new genes and transcripts [66], and fusion transcripts [67].

One difficulty involving the utility of RNA-Seq is estimating abundance at the gene level and differential expression at the transcriptional level under various conditions. RNA-Seq can determine the expression profile of normal and affected cells and tissues [68].

The etiology of Alzheimer's disease (AD) is complicated and remains challenging to research efforts worldwide. In the absence of a greater understanding of AD pathogenesis, cure strategies do not supply a treatment, however only deal with symptoms or decrease the price of onset. The transcriptome displays cellular activity within the tissue at a given time. Genome-wide expression studies, which are no longer influenced by deductive assumptions, provide an independent strategy for investigating the etiology of complicated ailments such as AD. Transcriptome analyses have been performed using transgenic animal models of AD and patient-derived cell lines [69, 70].

In contrast to these approaches, an autopsy of brain tissue is challenging to obtain, and some RNA quality issues can affect transcriptome studies [71, 72]. Nonetheless, the same postmortem brain tissue as the tissue affected by the disease remains the gold standard for evaluating against all other model systems. However, transcriptome studies of AD using brain tissue have

yielded almost contradictory results. The latest improvement in next-generation sequencing offers a more complete and accurate tool for transcriptome analysis of this invaluable resource [73, 74].

Jinquan et al., in 2014, sought to identify differences in *ATRX* mRNA expression that extended the biological understanding of astrocytic tumors and supplied new possible markers of prognosis. They used RNA-Seq in 169 astrocyte tumor samples in which three levels of different *ATRX* mRNA expressions have been detected [75]. Their approach identified *ATRX* as a prognostic marker and highlighted the power of RNA-Seq technology in characterizing three subsets of astrocytic tumors [76, 77].

Information about environmental and other influences is partly captured in the transcriptome, which can be explored through RNA-Seq, part of the Large-scale Unbiased Sequencing (LUS) family of technologies. RNA-Seq signatures are currently under investigation (e.g., in some breast cancers) and may provide a previous opportunity to combine genomic and transcriptomic data [78]. RNA-Seq analysis of individual subsets of peripheral blood mononuclear cells in patients with autoimmune disease has potential future research.

The importance of RNA-Seq in drug development is becoming increasingly apparent to clinicians and drug developers. Divergence in the expression levels and splicing of drug-metabolizing enzymes, transporters, and targets, such as receptors and ion channels, have been associated with inter-individual differences in optimal drug dose, drug effectiveness, and adverse drug events [79, 80].

Therefore, a comprehensive study of variation in the transcriptome profiles of pharmacologically relevant tissues promises to yield significant insights into the molecular basis of variation in drug response. In pharmacogenomics, polymorphisms that influence the expression levels or effect in alternative splicing of drug-metabolizing enzymes significantly affect drug disposition and

response. For instance, *UGT1A1*28* (rs8175347), with seven thymine-adenine 13 repeats in the promoter region, leads to decreased transcription rates of this enzyme and substantial toxicity in patients obtaining the topoisomerase inhibitor, irinotecan [81, 82]. Also, alternative splicing of *CYP2D6* often arises in human populations and is liable for the reduced activity of the enzyme [83]. Given these significant and clinically meaningful effects in drug-metabolizing enzymes, a systematic study of the transcriptome focusing on pharmacogenes is needed. With the support of the NIH, the Pharmacogenomics Global Research Network (PGRN) has launched a transcriptome sequencing project to catalog differences in gene expression and splicing between individuals within tissues and pharmacologically significant genes. They used this approach to represent the expression of 389 genes of pharmacologic significance in some human tissue types and lymphoblastoid cell lines (LCLs). Different from many other transcriptome profiling studies using RNA-Seq, this study showed findings for numerous samples across tissues, authorizing the capture of inter-individual divergence in expression levels in addition to comparison of expression and splicing across various tissues [84]. It was possible that peripheral B-lymphocytes, the primary cells from which LCLs were derived, also displayed various expression patterns from the other four physiological tissues (human liver, heart, kidney, adipose tissue) included in the study. These results proposed that considering the phenotype and the gene of interest is essential when utilizing LCLs as a substitute for other tissues in pharmacogenetic consideration and when using tissues as proxies for each other. This study also showed significant variability in gene expression, particularly among drug transporters and drug-metabolizing enzymes. Several cytochrome P450 (CYP) enzymes revealed substantial variability in expression levels between individuals in the liver; such variability can cause differences in drug metabolism across individuals, directing to divergence in drug effectiveness and vulnerability to toxicity [85].

8.7 Targeted Sequencing

WES and WGS techniques can be combined in a targeted sequencing (TS) approach to preserve the accuracy and abundance of WGS data while lowering costs. Both coding and noncoding regions of interest genes are captured by this technology. The selected genes are sequenced at a high coverage level typically more than 30-fold, thus leading to improving genotype calling accuracy by reducing error rates and uncertainty in genotype analysis, which are commonly encountered in short-read sequences. Target-enrichment approaches provide rapid detection and analysis of common and rare genetic variations that affect response to therapeutic drugs or adverse effects. This information is critical and fundamental for tailoring personalized pharmacotherapy [40, 57, 61].

There are several custom pharmacogenetic panels including drug target genes and other pharmacogenes that are involved in ADMET (absorption, distribution, metabolism and excretion, and toxicity), such as the PGRNseq panel, xGen Pan-Cancer Panel v2.4, and CleanPlex NGS Panel which cover 84 pharmacogenes, 532 responsible genes for cancer, and 180 pharmacogenes, respectively [57, 61, 86]. The PGRNseq platform is utilized in different PGx profiling researches, and it has the ability to cover all complex variations in different regions of the most important pharmacogenes such as *CYP2A6*, *CYP2D6*, and *HLA-B* genes [40, 44, 86]. Similarly, valuable and comprehensive panels for studying PGx genes have been developed by other research groups. The accuracy and coverage in these panels for more than 100 PK/PD-related genes were higher than 99% [87, 88]. It has been demonstrated this approach is especially relevant in the area of pharmacogenetic research as well as to actionable clinical targets for individualized pharmacotherapy, since particular noncoding sequences of genes encoding phase I and II enzymes (*CYP2C19*, *CYP3A5*, *CYP3A4*, and *UGT1A1*) can be enriched and targeted [40, 87, 88]. Over 5000 patients have been sequenced by PGRNseq with collaboration between the Electronic Medical Records and

Genomics (eMERGE) network, and most of identified variants are related to CPIC [40].

Target enrichment can be achieved through multiple strategies such as molecular inversion probe (MIP), polymerase chain reaction (PCR)-based, or hybridization capture-based, but the results can vary considerably among these approaches. Among these methods, the MIP-based approach has been used for large-scale sequencing in several versions with considerable advancements, and the modest quantities of input DNA can be captured with high specificity [89].

Han et al. developed an MPI strategy with capturing improvement based on the PharmaADME database and ADMET-PGx to detect relevant and rare variants of 114 PK/PD-related genes in 375 Korean subjects. According to their finding, widespread profiling of pharmacogenes which are important for personalized medicine approaches can be easily accomplished using this method [88]. In addition to screening patients for functional variations, these panels aimed to be a diagnostic tool for rapid and reliable identification of rare, potentially clinically significant variations across a population.

The other research group designed an exome panel of capture probe (PGxseq panel) for 100 pharmacogenes including all SNVs and CNVs in 235 patients. They confirmed that a technique like the PGxseq panel can be used as a robust, fast, and accurate method to identify common as well as novel SNVs alongside CNVs in drug target genes, which will provide insights into the area of precision medicine [87]. In this study, the noncoding region has been not sequenced. Klein et al. performed a comprehensive study of 340 ADMET-related pharmacogenes using a targeted NGS-PGx panel with coverage at least 100-fold, and all SNVs, small Indel, and large structural variants were analyzed with MAF below 2%. Similar to other studies, they found that deleterious variations are more prevalent in less common variants, and also they demonstrated that this approach can provide a more accurate pharmacogenetic framework for the prediction of toxicity and adverse effects of drugs [57, 87, 88, 90].

Overall, targeted sequencing is the more cost-effective and higher level of coverage in comparison to other advanced techniques, providing valuable information for uncommon variants and unbiased PGx profiling. But despite this fact, there are some limitations and challenges with the TS technique. One of the major limitations is that only a tiny percentage of all medications' metabolism is regulated by a few genetic variations, while most of pharmaceuticals exhibit only minor effects of multiple variants (most of which are still unknown to date). The other is sometimes difficult to determine large Indels (more than 1 kb) with short-read sequencing since Indel length might exceed the length of the read. Besides this, developing a close collaboration network between clinicians and analysts can be a challenging task.

8.8 Single-Cell Sequencing

The current development of SCS techniques has led to a paradigm shift in genomics, away from bulk tissue analysis and toward distinctive and comprehensive research of individual cells. A significant milestone occurred in 2005 with the development of the first NGS technologies, which enabled genome-wide sequencing of DNA and RNA [91]. The highest point of these technologies led to the invention of the first genome-wide single-cell DNA [92] and RNA [93] techniques for mammalian cells. These preliminary studies led to the establishment of a new discipline of biology: single-cell sequencing.

The improvement of DNA SCS techniques has been established to be extra challenging than RNA. A single cell comprises only two copies of each DNA molecule but many copies of most RNA molecules. Due to the restrained amount of WGA input material, some technical errors such as coverage nonuniformity, allele dropout (ADO) events, false-positive (FP) errors, and false-negative (FN) errors occur [94, 95]. The first SCS method was developed for genomic DNA combined with degenerative oligonucleotide PCR with flow-sorting nuclei and NGS to create high-

resolution copy number profiles for single mammalian cells [92, 96]. Since then, many SCS with higher coverage technologies have been used.

Single-cell RNA sequencing technology has shown remarkable progress in recent years. RNA must first be amplified by amplifying the entire transcriptome to sequence a single-cell transcriptome. This step is needed because a typical mammalian cell carries only 10 pg of total RNA and 0.1 pg of mRNA [97].

Epigenomic profiling of single cells remains one of the most significant technical challenges in the area. The difficulty is that standard epigenomic sequencing methods need a pool of DNA split into two separate fractions for treatment with bisulfite or methylation restriction enzymes before sequencing. The other technical barrier is that epigenetic DNA modifications cannot be amplified with DNA polymerases. Despite these technical hurdles, studies have made initial progress [98, 99]. SCS methods have impacted many broad fields of biology. They include microbiology, neurobiology, tissue mosaicism, germline transmission, organogenesis, immunology, cancer research, and clinical applications. Figure 8.3 represents the clinical applications of SCS in various fields [100]. Single-cell DNA and RNA sequencing methods supply a powerful new approach to unraveling microbial genomes and depicting intercellular diversity within different populations. However, bacteria and other microorganisms often have only femtograms of DNA

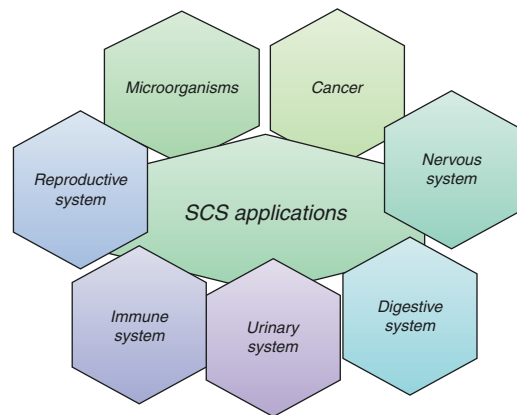


Fig. 8.3 The clinical applications of SCS in different fields

and RNA, making it even more challenging to amplify than mammalian cells [95].

Single-cell RNA sequencing presents a practical and impartial technique for categorizing neurons based on transcriptional profiles. In a study by Qiu and his coworkers, the RNA sequencing of a single neuron was combined with electrophysiology to obtain transcriptional profiles of mouse embryonic hippocampus and neocortical neurons [101]. In a study done by Usoskin et al. using single-cell RNA sequencing, 622 sensory neurons in mice were profiled, revealing 11 novel expression classes of sensory neuron cell types [102].

A new method for studying the mechanisms that cause germline variation was presented as single-cell DNA sequencing. In a study on this topic, after single sperm cell sequencing, the results consisted of ~22.8 recombination events, 5–15 gene conversion events, and 25–36 de novo mutations in each sperm cell [103]. Copy number profile calculation showed that 7% of the single sperm cells had aneuploid genomes.

RNA SCS was used to analyze transcriptional reprogramming *in vitro* during the transition from the inner cell mass of blastocysts to pluripotent embryonic stem cells [104]. Also, it was used to study transcriptome dynamics from oocyte to morula development in human and mouse embryos, which delineated a stepwise advancement of pathways that regulate the cell cycle, gene regulation, translation, and metabolism [105].

RNA SCS methods provide a robust new fair approach to perform transcriptional profiling and determine groups of cells that share standard expression programs, representing specific cell types. In another research, RNA SCS was used to analyze lung epithelium development in the first study to apply this approach [106]. These data chased the development of lung progenitor cells that form the alveolar air sac that regulates gas exchange. Also, the authors recognized lots of novel markers for distinguishing the four essential cell types and used them to reconstruct the cell lineage throughout alveolar sac differentiation.

The primary immune cell types have been known for decades, but little is known about transcriptional heterogeneity within cell types that

respond to antigens. One study used RNA-SCS to analyze bone marrow-derived dendritic cells from mice stimulated *in vitro* under various conditions, with individual cells showing different responses mediated by paracrine interferon signaling [107].

Most SCS studies of cancer research have centered on intra-tumor heterogeneity and clonal evolution in primary tumors. The first study used single-nucleus RNA sequencing (SNS) to observe the improvement of aneuploidy expansion in single cells of sufferers with triple-negative (ER/PR/HER2) breast cancer [92]. These data indicated that copy number aberrations developed in punctuated bursts of evolution, tracked by steady clonal expansions to form the tumor mass. In summary, SCS procedures have already substantially enhanced our fundamental understanding of intra-tumor heterogeneity, clonal evolution, and metastatic dissemination in human cancers [108, 109].

SCS techniques have direct translational applications in cancer therapy and prenatal genetic diagnosis (PGD) in clinical applications. In cancer research, intra-tumor heterogeneity shows a considerable challenge for clinical diagnostics because single samples may not represent the tumor as a whole. SCS supplies a potent tool for determining intra-tumor heterogeneity and steering targeted treatment toward the most malignant clones. SCS can also be utilized to estimate a diversity index for each cancer patient, which may have prognostic utility for predicting poor survival and unsatisfactory response to chemotherapy.

Single-cell sequencing translation applications in precision cancer treatment can improve cancer diagnosis, prognosis, targeted therapy, early detection, and noninvasive monitoring [110]. Single-cell sequencing enables sensitive detection of rare mutations and cell-specific gene expression profiles. This method can identify rare tumor tissue variants that may promote drug resistance or act as biomarkers for successful treatment, ultimately advancing cancer genomics [111]. Drug resistance dynamics have been formerly modeled in metastatic breast cancer cell lines using RNA-Seq technology [111].

Treatment of metastatic breast cancer cells with paclitaxel causes the stressed cells to stop and die, but those rare drug-resistant cells resume proliferation, and clones expand. The strength to profile the genome and transcriptome of the same cell can potentially unravel heterogeneity at the genomic, epigenomic, and transcriptomic levels. SCS in drug development grows on bulk genomic data by proposing a more complete and comprehensive picture of responders' underlying genetics, epigenetics, and transcriptomics versus nonresponders at an individual cell level. Applications of SCS in pharmaceutical development include identifying drug candidates and drug targets, drug resistance, and drug reactions and toxicities [112].

8.9 DNA Microarray

DNA microarray technology has the potential to be a swift, reliable, and affordable technique for pharmacologic research and clinical activities by allowing investigators to study the expression of the entire human genome simultaneously. DNA microarrays are commonly used to analyze changes in gene expression patterns across the genome to link genes or proteins to drug responses. Disease prevention, drug response prediction, personalized medicine, and the molecular fingerprints of different genetic diseases such as nervous system disorders and cancer could be as results of studying gene expression profiles. Gene expression changes are hierarchic, regulated, and compatible with the phenotypic and physiological responses to medication.

The most widely used platforms for measuring gene expression are Affymetrix and Illumina. The GeneChip was created by Affymetrix using a photolithographic process. By doing microarray experiments, GeneChip is performed to analyze the expression levels of genes in samples. GeneChips are constructed from the sequence repetition of the multiple probes in which a hundred repetitions is usually sufficient. Photolithography and *in situ* solid-phase oligonucleotide DNA synthesis were used to create the GeneChip array. In contrast, silica microbeads

are used in Illumina microarrays. Several copies of an oligonucleotide probe are coated on the silica beads that are placed in microwells targeted at specific genes in the genome. The matrix component can be presynthesized oligonucleotides or PCR-amplified cDNA inserts obtained from expressed sequence tags (ESTs) using high-speed robotics. The density characteristics, tiny features, and the ability to analyze multiple samples simultaneously are some of Illumina's strengths and also is one of the cheapest available techniques [113, 114].

Today, different advancements have been achieved using these platforms in the area of pharmacogenomics, toxicogenomics, gene discovery, and discriminating between responders and nonresponders to prescribed drugs. Several commercial arrays containing pharmacogenetic content are available such as Agena's iPLEX PGx Pro panel, Infinium Global Screening Array (GSA), VeraCode ADME core panel, Affymetrix's Drug Metabolizing Enzymes and Transporters panel, and ADMET arrays, which could be used for detection of PGx variants which are related to drug response based on PharmGKB or PGx guidelines [44, 47, 57, 87, 115].

A large number of variations are included in some commercial arrays and most of them are interested in research in personalized medicine and clinical activities. For example, the AmpliChip™ CYP450 test from Affymetrix microarray technology was authorized by FDA to analyze 27 CYP2D6 alleles (including seven duplications) and three CYP2C19 alleles linked to distinct metabolizing phenotypes [116].

For novel gene and pharmacological target identification, increasing the number of targets on the array to include anonymous ESTs, ESTs with functional orthologues, and homologies to known genes of other animal models can be useful. When assessing microarray approaches, some major considerations should be evaluated: (1) array content, (2) ability to change the content, (3) array expenditure, (4) cycle times for DNA sample, (5) sample size of each array, and (6) technician associated with creating the data. The expression data of each sample should be compared to a publicly available database, and

expression profiles will be used to study the biological effect of cytotoxic agents, therapeutic drugs, environmental toxins, and adverse effects of different drugs used for genetic disorders, especially cancers. Using gene expression profiles, it is now possible to identify patients at risk and leukemia subtypes with poor prognoses that may end up failing therapy [57, 113, 115, 117].

Chine et al. analyzed the expression pattern of breast cancer MCF-7 cells chosen for anti-doxorubicin resistance or treated with doxorubicin using DNA microarray. They observed transient alterations in the expression of a significant number of genes in MCF-7 cells treated with doxorubicin. Some of these genes such as XRCC1 and microsomal epoxide hydrolase 1 have a critical role in drug resistance, which may lead to accelerated doxorubicin metabolism and reduce medication availability. According to this data, they were able to define the treatment plan and anticipate clinical results for that patient [114]. Stanford University and the University of Florida collaborated in 2012 to establish an SNP microarray panel from 120 genes, including 25 genes involved in drug metabolism and 12 drug transporter genes based on PharmGKB. In total, 256 SNPs are screened including 252 “PGXs SNPs,” two quality-control duplicates, and two sex markers. This panel is used for disease risk prediction, improving patient care based on genetic information, and providing pharmacogenetic data in clinical activities [117].

Clinical PGx implementation studies frequently employ these types of arrays. The VeraCode ADME core panel was used in Vanderbilt Electronic Systems for Pharmacogenetic Assessment study and the PREDICT project to generate extensive information for PGx and precision medicine. The St. Jude’s Children’s Research Hospital used the DMET array through the PG4KDS protocol (step-by-step approach to deploying gene/drug pairings, collecting data, and getting patient and family permission) in 1559 patients that the role of four genes (*CYP2D6*, *CYP2C1*, *TPMT*, and *SLCO1B1*) has been highlighted in PGx implementation [57, 118]. The genome-wide and pharmacogene coverage of the commercial microarray

panel has been investigated in the comprehensive study consisting of 15 Affymetrix genome-wide and 18 Illumina arrays. They analyzed more than 20,000 variants in 3146 genes and demonstrated these panels provide low coverage for genome-wide study, but they could be implemented as complementary assays in pharmacogenomics investigations [115].

The developments in DNA microarray allow comparative measurement of all genes and their products and also evaluate the alteration of gene expression levels in response to drug treatment. In combination with PGx approaches in the pre-clinical phase of drug discovery, this high throughput provides insight into the cytotoxic effects of drugs before they are clinically tested. Besides this, linking the in vivo PK/PD experiments and modeling/stimulation data with an expression profile database would shed light on knowledge of pharmacological mechanisms and therapeutic effects and accelerate the speed of drug discovery. Therefore, comprehensive information was obtained across any chemical compounds and drug targets chosen from monitoring in the combination of expression profile and chemical genomics.

8.10 Conclusion

NGS in combination with innovative technologies such as DNA microarray and transcriptome sequencing created a new window for the investigation of genetic disorders. The discovery of the human genome, alongside the development of high-throughput technologies, provides a strong potential for the detection of complex genetic variants, especially with the advent of the era of personalized medicine. In this era, genetic profiling as a useful diagnostic tool enables to offer pharmacological therapy with greater efficacy and fewer unwanted side effects. In recent years, the variety of genotyping technologies for PGx has grown dramatically and keeps rising.

NGS technologies are becoming more common in clinics and PGx research studies and, as costs fall, make it a routine part of medical care and treatment. These technologies provide a lot

more information and also are easier, faster, and more targeted than other types of genetic testing, which may be extremely useful when attempting to figure out what's wrong with a patient. The widespread adoption of DNA technologies in various clinical settings will be due to the quick development of component and bioinformatics tools, as well as the lower cost and technical innovation that will allow for testing of a larger number of drug-related genes and biomarkers.

Despite some limitations of DNA technologies such as time-consuming for data analysis, huge storage capacity, the requirement for complicated bioinformatics processes, identifying the high number of VUS and their management, poor coverage of some sequences by various platforms, and the limitation of functional analysis through bioinformatics tools for variants, the use of actionable pharmacogenetic variations and PGx testing in clinical practice and researches is growing. These approaches provide opportunities for PGx variant discovery, more accurate prediction of specific drug phenotypes in individuals, and, as a result, more appropriate genotype-based treatment modifications and a promising future for pharmacogenomics-guided medicine.

References

- Madian AG, et al. Relating human genetic variation to variation in drug responses. *Trends Genet.* 2012;28(10):487–95.
- Guchelaar H-J. Pharmacogenomics, a novel section in the European Journal of Human Genetics. Berlin: Nature Publishing Group; 2018. p. 1399–400.
- Suarez-Kurtz G, Parra EJ. Population diversity in pharmacogenetics: a Latin American perspective. *Adv Pharmacol.* 2018;83:133–54.
- Evans WE, Johnson JA. Pharmacogenomics: the inherited basis for interindividual differences in drug response. *Annu Rev Genomics Hum Genet.* 2001;2(1):9–39.
- Schärfe CPI, et al. Genetic variation in human drug-related genes. *Genome Med.* 2017;9(1):1–15.
- Rabbani B, Tekin M, Mahdih N. The promise of whole-exome sequencing in medical genetics. *J Hum Genet.* 2014;59(1):5–15.
- Rabbani B, et al. Next-generation sequencing: impact of exome sequencing in characterizing Mendelian disorders. *J Hum Genet.* 2012;57(10):621–32.
- Diaz-Horta O, et al. Whole-exome sequencing efficiently detects rare mutations in autosomal recessive nonsyndromic hearing loss. *PLoS One.* 2012;7(11):e50628.
- Kingsmore S, Saunders C. Deep sequencing of patient genomes for disease diagnosis: when will it become routine? *Sci Transl Med.* 2011;3:87ps23.
- Bamshad MJ, et al. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet.* 2011;12(11):745–55.
- Metzker M. Sequencing technologies—the next generation. *Nat Rev Genet.* 2010;11(1):31–46.
- Bielinski SJ, et al. Preemptive genotyping for personalized medicine: design of the right drug, right dose, right time—using genomic data to individualize treatment protocol. In: Mayo clinic proceedings. Elsevier; 2014.
- Gabrovska P, et al. Gene expression profiling in human breast cancer—toward personalised therapeutics? *Open Breast Cancer J.* 2010;2:46–59.
- Rabbani B, et al. Next generation sequencing: implications in personalized medicine and pharmacogenomics. *Mol Biosyst.* 2016;12(6):1818–30.
- Hintzsche JD, Robinson WA, Tan AC. A survey of computational tools to analyze and interpret whole exome sequencing data. *Int J Genomics.* 2016;2016:7983236.
- Slatko BE, Gardner AF, Ausubel FM. Overview of next-generation sequencing technologies. *Curr Protoc Mol Biol.* 2018;122(1):e59.
- Dhawan D. Chapter 2 – Clinical next-generation sequencing: enabling precision medicine. In: Verma M, Barh D, editors. Progress and challenges in precision medicine. Academic Press: New York; 2017. p. 35–54.
- Health, N.I.o., *NCI dictionary of cancer terms.* National Institutes of Health, Bethesda, MD, USA; 2018.
- Gilissen C, et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature.* 2014;511(7509):344–7.
- Nones K, et al. Genomic catastrophes frequently arise in esophageal adenocarcinoma and drive tumorigenesis. *Nat Commun.* 2014;5(1):1–9.
- van El CG, et al. Whole-genome sequencing in health care: recommendations of the European Society of Human Genetics. *Eur J Hum Genet.* 2013;21(6):580–4.
- Choi J, Tantisira KG, Duan QL. Whole genome sequencing identifies high-impact variants in well-known pharmacogenomic genes. *Pharmacogenomics J.* 2019;19(2):127–35.
- Yang W, et al. Comparison of genome sequencing and clinical genotyping for pharmacogenes. *Clin Pharmacol Ther.* 2016;100(4):380–8.
- He Y, Hoskins JM, McLeod HL. Copy number variants in pharmacogenetic genes. *Trends Mol Med.* 2011;17(5):244–51.
- Martis S, et al. Multi-ethnic cytochrome-P450 copy number profiling: novel pharmacogenetic alleles and

- mechanism of copy number variation formation. *Pharmacogenomics J.* 2013;13(6):558–66.
26. Santos M, et al. Novel copy-number variations in pharmacogenes contribute to interindividual differences in drug pharmacokinetics. *Genet Med.* 2018;20(6):622–9.
 27. Ashley EA, et al. Clinical assessment incorporating a personal genome. *Lancet.* 2010;375(9725):1525–35.
 28. Daly AK. Genome-wide association studies in pharmacogenomics. *Nat Rev Genet.* 2010;11(4):241–6.
 29. Welch JS, et al. Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. *JAMA.* 2011;305(15):1577–84.
 30. Jorgenson E, Witte JS. A gene-centric approach to genome-wide association studies. *Nat Rev Genet.* 2006;7(11):885–91.
 31. Badiie A, et al. Evaluation of five different cDNA labeling methods for microarrays using spike controls. *BMC Biotechnol.* 2003;3(1):1–5.
 32. Lam HY, et al. Performance comparison of whole-genome sequencing platforms. *Nat Biotechnol.* 2012;30(1):78–82.
 33. Zook JM, et al. Integrating human sequence data sets provides a resource of benchmark SNP and indel genotype calls. *Nat Biotechnol.* 2014;32(3):246–51.
 34. Klein H-G, Bauer P, Hambuch T. Whole genome sequencing (WGS), whole exome sequencing (WES) and clinical exome sequencing (CES) in patient care. *LaboratoriumsMedizin.* 2014;38 <https://doi.org/10.1515/labmed-2014-0025>.
 35. Suwinski P, et al. Advancing personalized medicine through the application of whole exome sequencing and big data analytics. *Front Genet.* 2019;10:49.
 36. Yang Y, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med.* 2013;369(16):1502–11.
 37. McVean GA, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature.* 2012;491(7422):56–65.
 38. Tennessen JA, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science.* 2012;337(6090):64–9.
 39. Lek M, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536(7616):285–91.
 40. Schwarz UI, Gulilat M, Kim RB. The role of next-generation sequencing in pharmacogenetics and pharmacogenomics. *Cold Spring Harb Perspect Med.* 2019;9(2):a033027.
 41. Dewey FE, et al. Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study. *Science.* 2016;354(6319):aaf6814.
 42. Kozyra M, Ingelman-Sundberg M, Lauschke VM. Rare genetic variants in cellular transporters, metabolic enzymes, and nuclear receptors can be important determinants of interindividual differences in drug response. *Genet Med.* 2017;19(1):20–9.
 43. Katsila T, Patrinos GP. Whole genome sequencing in pharmacogenomics. *Front Pharmacol.* 2015;6:61.
 44. Tafazoli A, et al. Applying next-generation sequencing platforms for Pharmacogenomic testing in clinical practice. *Front Pharmacol.* 2021;12(2025):693453.
 45. Chua EW, et al. Cross-comparison of exome analysis, next-generation sequencing of amplicons, and the iPLEX(®) ADME PGx panel for Pharmacogenomic profiling. *Front Pharmacol.* 2016;7:1.
 46. Mueller JJ, et al. Massively parallel sequencing analysis of mucinous ovarian carcinomas: genomic profiling and differential diagnoses. *Gynecol Oncol.* 2018;150(1):127–35.
 47. Pereira M, et al., Application of next-generation sequencing in the era of precision medicine; 2017.
 48. Costain G, et al. Clinical application of targeted next-generation sequencing panels and whole exome sequencing in childhood epilepsy. *Neuroscience.* 2019;418:291–310.
 49. Raffan E, Semple RK. Next generation sequencing—implications for clinical practice. *Br Med Bull.* 2011;99(1):53–71.
 50. Khoja L, et al. Pembrolizumab. *J Immunother Cancer.* 2015;3:36.
 51. Prasad V, Kaestner V, Mailankody SJJ. Cancer drugs approved based on biomarkers and not tumor type—FDA approval of pembrolizumab for mismatch repair-deficient solid cancers. *JAMA Oncol.* 2018;4(2):157–8.
 52. Jørgensen JT. Twenty years with personalized medicine: past, present, and future of individualized pharmacotherapy. *Oncologist.* 2019;24(7):e432–40.
 53. Demkow U. Chapter 11 - next generation sequencing in pharmacogenomics. In: Demkow U, Ploski R, editors. *Clinical applications for next-generation sequencing.* Boston: Academic Press; 2016. p. 217–40.
 54. Cohn I, et al. Genome sequencing as a platform for pharmacogenetic genotyping: a pediatric cohort study. *npj Genomic Med.* 2017;2(1):19.
 55. Zhang Q, et al. Clinical application of whole-exome sequencing: a retrospective, single-center study. *Exp Ther Med.* 2021;22(1):753.
 56. Carter TC, He MM. Challenges of identifying clinically actionable genetic variants for precision medicine. *J Healthc Eng.* 2016;2016:3617572.
 57. van der Lee M, et al. Technologies for pharmacogenomics: a review. *Genes (Basel).* 2020;11:12.
 58. Cousin MA, et al. Pharmacogenomic findings from clinical whole exome sequencing of diagnostic odyssey patients. *Mol Genet Genomic Med.* 2017;5(3):269–79.
 59. Rennert H, et al. Development and validation of a whole-exome sequencing test for simultaneous detection of point mutations, indels and copy-number alterations for precision cancer care. *Npj Genomic Med.* 2016;1(1):16019.
 60. Dong L, et al. Clinical next generation sequencing for precision medicine in cancer. *Curr Genomics.* 2015;16(4):253–63.
 61. Tilleman L, et al. Pan-cancer pharmacogenetics: targeted sequencing panels or exome sequencing? *Pharmacogenomics.* 2020;21(15):1073–84.

62. Guidugli L, et al. Clinical utility of gene panel-based testing for hereditary myelodysplastic syndrome/acute leukemia predisposition syndromes. *Leukemia*. 2017;31(5):1226–9.
63. Conesa A, et al. A survey of best practices for RNA-seq data analysis. *Genome Biol*. 2016;17(1):1–19.
64. Hoelzer D, et al. Acute lymphoblastic leukaemia in adult patients: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2016;27:v69–82.
65. Wang ET, et al. Alternative isoform regulation in human tissue transcriptomes. *Nature*. 2008;456(7221):470–6.
66. Denoeud F, et al. Annotating genomes with massive-scale RNA sequencing. *Genome Biol*. 2008;9(12):1–12.
67. Maher CA, et al. Transcriptome sequencing to detect gene fusions in cancer. *Nature*. 2009;458(7234):97–101.
68. Carey MF, Peterson CL, Smale ST. Chromatin immunoprecipitation (chip). *Cold Spring Harb Protoc*. 2009;2009(9):pdb.prot5279.
69. Soldner F, et al. Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. *Cell*. 2009;136(5):964–77.
70. Matigian N, et al. Disease-specific, neurosphere-derived cells as models for brain disorders. *Dis Model Mech*. 2010;3(11–12):785–98.
71. Atz M, et al. Methodological considerations for gene expression profiling of human brain. *J Neurosci Methods*. 2007;163(2):295–309.
72. Monoranu CM, et al. pH measurement as quality control on human post mortem brain tissue: a study of the BrainNet Europe consortium. *Neuropathol Appl Neurobiol*. 2009;35(3):329–37.
73. Courtney E, et al. Transcriptome profiling in neurodegenerative disease. *J Neurosci Methods*. 2010;193(2):189–202.
74. Janitz M. Next-generation genome sequencing: towards personalized medicine. Hoboken, NJ: John Wiley & Sons; 2011.
75. Cai J, et al. ATRX mRNA expression combined with IDH1/2 mutational status and Ki-67 expression refines the molecular classification of astrocytic tumors: evidence from the whole transcriptome sequencing of 169 samples. *Oncotarget*. 2014;5(9):2551.
76. Kridel R, et al. Whole transcriptome sequencing reveals recurrent NOTCH1 mutations in mantle cell lymphoma. *Blood J Am Soc Hematol*. 2012;119(9):1963–71.
77. Greif P, et al. Identification of recurring tumor-specific somatic mutations in acute myeloid leukemia by transcriptome sequencing. *Leukemia*. 2011;25(5):821–7.
78. Creighton CJ, Reid JG, Gunaratne PH. Expression profiling of microRNAs by deep sequencing. *Brief Bioinform*. 2009;10(5):490–7.
79. Wang L, McLeod HL, R.M.J.N.E.J.o.M. Weinshilboum. Genomics and drug response. *N Engl J Med*. 2011;364(12):1144–53.
80. Evans WE, McLeod HL. Pharmacogenomics—drug disposition, drug targets, and side effects. *N Engl J Med*. 2003;348(6):538–49.
81. Iyer L, et al. UGT1A1* 28 polymorphism as a determinant of irinotecan disposition and toxicity. *Pharmacogenomics J*. 2002;2(1):43–7.
82. Tukey RH, Strassburg CP, Mackenzie PI. Pharmacogenomics of human UDP-glucuronosyltransferases and irinotecan toxicity. *Mol Pharmacol*. 2002;62(3):446–50.
83. Wang D, et al. Common CYP2D6 polymorphisms affecting alternative splicing and transcription: long-range haplotypes with two regulatory variants modulate CYP2D6 activity. *Hum Mol Genet*. 2014;23(1):268–78.
84. Chhibber A, et al. Transcriptomic variation of pharmacogenes in multiple human tissues and lymphoblastoid cell lines. *Pharmacogenomics J*. 2017;17(2):137–45.
85. Zanger UM, Schwab M. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol Ther*. 2013;138(1):103–41.
86. Gordon AS, et al. PGRNseq: a targeted capture sequencing panel for pharmacogenetic research and implementation. *Pharmacogenet Genomics*. 2016;26(4):161–8.
87. Gulilat M, et al. Targeted next generation sequencing as a tool for precision medicine. *BMC Med Genet*. 2019;12(1):81.
88. Han SM, et al. Targeted next-generation sequencing for comprehensive genetic profiling of pharmacogenes. *Clin Pharmacol Ther*. 2017;101(3):396–405.
89. Mamanova L, et al. Target-enrichment strategies for next-generation sequencing. *Nat Methods*. 2010;7(2):111–8.
90. Klein K, et al. A new panel-based next-generation sequencing method for ADME genes reveals novel associations of common and rare variants with expression in a human liver cohort. *Front Genet*. 2019;10:7.
91. Mardis ERJN. A decade's perspective on DNA sequencing technology. *Nature*. 2011;470(7333):198–203.
92. Navin N, Hicks J. Future medical applications of single-cell sequencing in cancer. *Genome Med*. 2011;3(5):1–12.
93. Tang F, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods*. 2009;6(5):377–82.
94. Dean FB, et al. Comprehensive human genome amplification using multiple displacement amplification. *Proc Natl Acad Sci*. 2002;99(8):5261–6.
95. Lasken RS. Single-cell genomic sequencing using multiple displacement amplification. *Curr Opin Microbiol*. 2007;10(5):510–6.

96. Baslan T, et al. Genome-wide copy number analysis of single cells. *Nat Protoc.* 2012;7(6):1024–41.
97. Van Gelder RN, et al. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc Natl Acad Sci.* 1990;87(5):1663–7.
98. Nagano T, et al. Single-cell hi-C reveals cell-to-cell variability in chromosome structure. *Nature.* 2013;502(7469):59–64.
99. Guo H, et al. Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res.* 2013;23(12):2126–35.
100. Tang X, et al. The single-cell sequencing: new developments and medical applications. *Cell Biosci.* 2019;9(1):1–9.
101. Qiu S, et al. Single-neuron RNA-Seq: technical feasibility and reproducibility. *Front Genet.* 2012;3:124.
102. Usoskin D, et al. Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nat Neurosci.* 2015;18(1):145–53.
103. Wang J, et al. Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. *Cell.* 2012;150(2):402–12.
104. Tang F, et al. Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. *Cell Stem Cell.* 2010;6(5):468–78.
105. Xue Z, et al. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature.* 2013;500(7464):593–7.
106. Treutlein B, et al. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature.* 2014;509(7500):371–5.
107. Shalek AK, et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature.* 2014;510(7505):363–9.
108. Navin NE. Cancer genomics: one cell at a time. *Genome Biol.* 2014;15(8):1–13.
109. Van Loo P, Voet T. Single cell analysis of cancer genomes. *Curr Opin Genet Dev.* 2014;24:82–91.
110. Navin NE. The first five years of single-cell cancer genomics and beyond. *Genome Res.* 2015;25(10):1499–507.
111. Lee M-CW, et al. Single-cell analyses of transcriptional heterogeneity during drug tolerance transition in cancer cells by RNA sequencing. *Proc Natl Acad Sci.* 2014;111(44):E4726–35.
112. Xu X, et al. Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell.* 2012;148(5):886–95.
113. Fajriyah R. Paper review: an overview on microarray technologies. *Bull Appl Math Math Educ.* 2021;1(1):21–30.
114. Chin K-V, Kong ANT. Application of DNA microarrays in pharmacogenomics and Toxicogenomics. *Pharm Res.* 2002;19(12):1773–8.
115. Lemieux Perreault LP, et al. Pharmacogenetic content of commercial genome-wide genotyping arrays. *Pharmacogenomics.* 2018;19(15):1159–67.
116. de Leon J, Susce MT, Murray-Carmichael E. The AmpliChip CYP450 genotyping test: integrating a new clinical tool. *Mol Diagn Ther.* 2006;10(3):135–51.
117. Johnson JA, et al. Implementing personalized medicine: development of a cost-effective customized pharmacogenetics genotyping array. *Clin Pharmacol Ther.* 2012;92(4):437–9.
118. Hoffman JM, et al. PG4KDS: a model for the clinical implementation of pre-emptive pharmacogenetics. *Am J Med Genet C Semin Med Genet.* 2014;166C(1):45–55.