

Genomics, Other “Omic” Technologies, Personalized Medicine, and Additional Biotechnology-Related Techniques

Robert D. Sindelar

INTRODUCTION

The products resulting for biotechnologies continue to grow at an exponential rate, and the expectations are that an even greater percentage of drug development will be in the area of the biologics. In 2011, worldwide there were over 800 new biotech drugs and treatments in development including 23 antisense, 64 cell therapy, 50 gene therapy, 300 monoclonal antibodies, 78 recombinant proteins, and 298 vaccines (PhRMA 2012). Pharmaceutical biotechnology techniques are at the core of most methodologies used today for drug discovery and development of both biologics and small molecules. While recombinant DNA technology and hybridoma techniques were the major methods utilized in pharmaceutical biotechnology through most of its historical timeline, our ever-widening understanding of human cellular function and disease processes and a wealth of additional and innovative biotechnologies have been, and will continue to be, developed in order to harvest the information found in the human genome. These technological advances will provide a better understanding of the relationship between genetics and biological function, unravel the underlying causes of disease, explore the association of genomic variation and drug response, enhance pharmaceutical research, and fuel the discovery and development of new and novel biopharmaceuticals. These revolutionary technologies and additional

biotechnology-related techniques are improving the very competitive and costly process of drug development of new medicinal agents, diagnostics, and medical devices. Some of the technologies and techniques described in this chapter are both well established and commonly used applications of biotechnology producing potential therapeutic products now in development including clinical trials. New techniques are emerging at a rapid and unprecedented pace and their full impact on the future of molecular medicine has yet to be imagined.

Central to any meaningful discussion of pharmaceutical biotechnology and twenty-first century health-care are the “omic” technologies. The completion of the Human Genome Project has provided a wealth of new knowledge. Researchers are turning increasingly to the task of converting the DNA sequence data into information that will improve, and even revolutionize, drug discovery (see Fig. 8.1) and patient-centered pharmaceutical care. Pharmaceutical scientists are poised to take advantage of this scientific breakthrough by incorporating state-of-the-art genomic and proteomic techniques along with the associated technologies utilized in bioinformatics, metabonomics/metabolomics, epigenomics, systems biology, pharmacogenomics, toxicogenomics, glycomics, and chemical genomics into a new drug discovery, development, and clinical translation paradigm. These additional techniques in biotechnology and molecular biology are being rapidly exploited to bring new drugs to market and each topic will be introduced in this chapter.

It is not the intention of this author to detail each and every biotechnology technique exhaustively, since numerous specialized resources already meet that need. Rather, this chapter will illustrate and enumerate various biotechnologies that should be of key interest to pharmacy students, practicing pharmacists, and pharmaceutical scientists because of their effect on many aspects of pharmacy, drug discovery, and drug development.

R.D. Sindelar, Ph.D., FCAHS
Department of Pharmaceutical Sciences
and Department of Medicine,
Providence Health Care and
The University of British Columbia,
1081 Burrard Street, Vancouver, BC V6Z 1Y6, Canada
e-mail: rsindelar@providencehealth.bc.ca,
robert.sindelar@ubc.ca

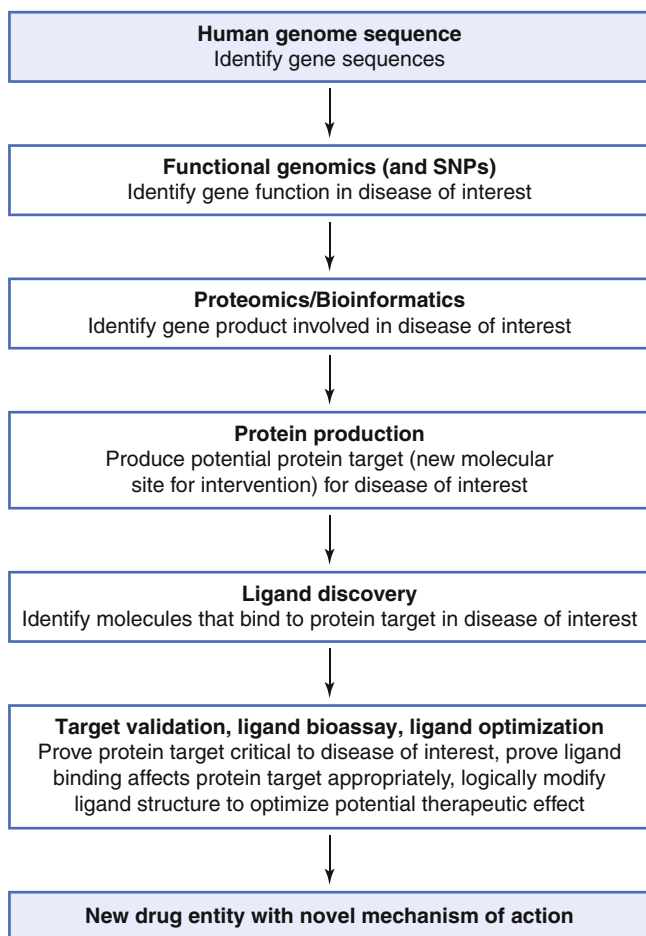


Figure 8.1 ■ The genomic strategy for new drug discovery.

AN INTRODUCTION TO “OMIC” TECHNOLOGIES

Since the discovery of DNA’s overall structure in 1953, the world’s scientific community has rapidly gained a detailed knowledge of the genetic information encoded by the DNA of a cell or organism so that today we are beginning to “personalize” this information. In the 1980s and 1990s, biotechnology techniques produced novel therapeutics and a wealth of information about the mechanisms of various diseases such as cancer at the genetic and molecular level, yet the etiologies of other complex diseases such as obesity and heart disease remained poorly understood. Recently, however, researchers utilizing exciting and groundbreaking “omic” technologies and working closely with clinicians have begun to make serious progress not only toward a molecular-level understanding of the etiology of complex diseases but to clearly identify that there are actually many genetically different diseases called by the single name of cancer, diabetes, depression, etc. Later in this chapter, we will explore the concepts of phenotype. Important here is that

most human diseases are manifested through very complex phenotypes that result from genetic, environmental, and other factors. In a large part, the answers were hidden in what was unknown about the human genome. Despite the increasing knowledge of DNA structure and function in the 1990s, the genome, the entire collection of genes and all other functional and nonfunctional DNA sequences in the nucleus of an organism, had yet to be sequenced. DNA may well be the largest, naturally occurring molecule known. Successfully meeting the challenge of sequencing the entire human genome is one of history’s great scientific achievements and heralds enormous potential (Venter et al. 2001; International Human Genome Sequencing Consortium 2001). While the genetic code for transcription and translation has been known for years, sequencing the human genome provides a blueprint for all human proteins and the sequences of all regulatory elements that govern the developmental interpretation of the genome. The potential significance includes identifying genetic determinants of common and rare diseases, providing a methodology for their diagnosis, suggesting interesting new molecular sites for intervention (see Fig. 8.1), and the development of new biotechnologies to bring about their eradication. Unlocking the secrets of the human genome may lead to a paradigm shift in clinical practice toward true targeted molecular medicine, better disease taxonomy, and patient-personalized therapy.

■ Genomics

The term genomics is the comprehensive analysis and understanding of DNA structure and function and broadly refers to the analysis of all genes within the genome of an organism. Sequencing the human genome and the genomes of other organisms has led to an enhanced understanding of not only DNA structure and function but also a fundamental understanding of human biology and disease. While it is a complex and complicated journey from DNA sample to DNA sequence stored in a database, a multitude of technologies and approaches along with impressive enhancements in instrumentation and computation have been employed to sequence genomic DNA faster and less expensively. While many industry analysts predicted a tripling of pharmaceutical R&D productivity due to the sequencing of the human genome, it is the “next-generation” genome sequencing technology and the quest for the “\$1000 genome” that will move genomic technology effectively into the clinic (Davies 2010).

Likewise, the field of genomics is having a fundamental impact on modern drug discovery and development. While validation of viable drug targets identified by genomics has been challenging, great progress has occurred (Yang et al. 2009). No matter

Human genome project goals	Base pair resolution
<i>Detailed genetic linkage map</i> Comments: poorest resolution; depicts relative chromosomal locations of DNA markers, genes, or other markers and the spacing between them on each chromosome	2 Mb
<i>Complete physical map</i> Comments: instead of relative distances between markers, maps actual physical distance in base pairs between markers; lower resolution = actual observance of chromosomal banding under microscope; higher resolution is “restriction map” generated in presence of restriction enzymes	0.1 Mb
<i>Complete DNA sequence</i> Comments: the ultimate goal; determine the base sequence of the genes and markers found in mapping techniques along with the other segments of the entire genome; techniques commonly used include DNA amplification methods such as cloning, PCR and other techniques described in Chap. 1 along with novel sequencing and bioinformatics techniques	1 bp

Mb megabase = 1 million base pairs, *bp* base pair

Table 8.1 ■ The increasing levels of genetic resolution obtained from structural genomic studies of the HGP.

whether it is a better understanding of disease or improved drug discovery, the genomic revolution has been the foundation for an explosion in “omic” technologies that find applications in research to address poorly treated and neglected diseases.

Structural Genomics and the Human Genome Project

Genetic analysis initially focused on the area of structural genomics, essentially, the characterization of the macromolecular structure of a genome utilizing computational tools and theoretical frameworks. Structural genomics intersects the techniques of DNA sequencing, cloning, PCR, protein expression, crystallography, and big data analysis. It focuses on the physical aspects of the genome through the construction and analysis of gene sequences and gene maps. Proposed in the late 1980s, the publicly funded Human Genome Project (HGP) or Human Genome Initiative (HGI) was officially sanctioned in October 1990 to map the structure and to sequence human DNA (US DOE 2012a). As described in Table 8.1, HGP structural genomics was envisioned to proceed through increasing levels of genetic resolution: detailed human genetic linkage maps [approximately 2 megabase pairs (Mb = million base pairs) resolution], complete physical maps (0.1 Mb resolution), and ultimately complete DNA sequencing of the approximately three billion base pairs (23 pairs of chromosomes) in a human cell nucleus [1 base pair

(bp) resolution]. Projected for completion in 2003, the goal of the project was to learn not only what was contained in the genetic code but also how to “mine” the genomic information to cure or help prevent the estimated 4,000 genetic diseases afflicting humankind. The project would identify all the approximately 20,000–25,000 genes in human DNA, determine the base pair sequence and store the information in databases, create new tools and improve existing tools for data analysis, and address the ethical, legal, and societal issues (ELSI) that may arise from the project. Earlier than projected, a milestone in genomic science was reached on June 26, 2000, when researchers at the privately funded Celera Genomics and the publicly funded International Human Genome Sequencing Consortium (the international collaboration associated with the HGP) jointly announced that they had completed sequencing 97–99 % of the human genome. The journal *Science* rates the mapping of the human genome as its “breakthrough of the year” in its December 22, 2000, issue. The two groups published their results in 2001 (Venter et al. 2001; The Genome International Sequencing Consortium 2001).

While both research groups employed the original cloning-based Sanger technique for DNA sequencing (now approximately 30 years old), the genomic DNA sequencing approaches of the HGP and Celera Genomics differed. HGP utilized a “nested shotgun” approach. The human DNA sequence was “chopped” into segments of ever decreasing size and the segments put into rough order. Each DNA segment was further divided or blasted into smaller fragments. Each small fragment was individually sequenced and the sequenced fragments assembled according to their known relative order. The Celera researchers employed a “whole shotgun” approach where they broke the whole genome into small fragments. Each fragment was sequenced and assembled in order by identifying where they overlapped. Each of the two sequencing approaches required unprecedented computer resources (the field of bioinformatics is described later in this chapter).

Regardless of genome sequencing strategies, the collective results are impressive. More than 27 million high-quality sequence reads provided fivefold coverage of the entire human genome. Genomic studies have identified over one million single-nucleotide polymorphisms (SNPs), binary elements of genetic variability (SNPs are described later in this chapter). While original estimates of the number of human genes in the genome varied consistently between 80,000–120,000, the genome researchers unveiled a number far short of biologist’s predictions; 32,000 (Venter et al. 2001; The International Human Genome Sequencing Consortium 2001). Within months, others suggested

that the human genome possesses between 65,000 and 75,000 genes (Wright et al. 2001). Approximately 20,000–25,000 genes is now most often cited number (Lee et al. 2006).

Next-Generation Genome Sequencing (NGS) and the \$1,000 Genome

The full spectrum of human genetic variation ranges from large chromosomal changes down to the single base pair alterations. The challenge for genomic scientists is to discover the full extent of genomic structural variation, referred to as genotyping, so that the variations and genetic coding may be associated with the encoded trait or traits displayed by the organism (the phenotype). And they wish to do this using as little DNA material as possible, in as short time and for the least cost, all important characteristics of a useful point-of-care clinical technology. The discovery and genotyping of structural variation has been at the core of understanding disease associations as well as identifying possible new drug targets (Alkan et al. 2011). DNA sequencing efficiency has facilitated these studies. In the decade since the completion of the HGP, sequencing efficiency has increased by approximately 100,000-fold and the cost of a single genome sequence has decreased from nearly \$1 million in 2007 to \$1,000 just recently (Treangen and Salzberg 2011). The move toward low-cost, high-throughput sequencing is essential for the implementation of genomics into personalized medicine and will likely alter the future clinical landscape. Next-generation genome sequencing methodologies, which differ from the original cloning-based Sanger technique, are high-throughput, imaging-based systems with vastly increased speeds and data output. There is no clear definition for *next-generation genome sequencing*, also known generally as NGS, but most are characterized by the direct and parallel sequencing of large numbers of amplified and fragmented DNA without vector-based cloning. The fragmented DNA tends to have sequence reads of 30–400 base pairs. There are now numerous examples of single-molecule techniques utilizing commercially available DNA sequencers (Cherf et al. 2012; Woollard et al. 2011). Early in 2012, the DNA sequencing companies Illumina and Life Technologies each announced new products that can sequence an entire human genome in 1 day for approximately \$1,000 (BusinessWeek 2012).

Functional Genomics and Comparative Genomics

Functional genomics is the subfield of genomics that attempts to answer questions about the function of specific DNA sequences at the levels of transcription and translation, i.e., genes, RNA transcripts, and protein products (Raghavachari 2012). Research to relate genomic sequence data determined by structural

genomics with observed biological function is predicted to fuel new drug discoveries through a better understanding of what genes do, how they are regulated, and the direct relationship between genes and their activity. The DNA sequence information itself rarely provides definitive information about the function and regulation of that particular gene. After genome sequencing, a functional genomic approach is the next step in the knowledge chain to identify functional gene products that are potential biotech drug leads and new drug discovery targets (see Fig. 8.1).

To relate functional genomics to therapeutic clinical outcomes, the human genome sequence must reveal the thousands of genetic variations among individuals that will become associated with diseases or symptoms in the patient's lifetime. Sequencing alone is not the solution, simply the end of the beginning of the genomic medicine era. Determining gene functionality in any organism opens the door for linking a disease to specific genes or proteins, which become targets for new drugs, methods to detect organisms (i.e., new diagnostic agents), and/or biomarkers (the presence or change in gene expression profile that correlates with the risk, progression, or susceptibility of a disease). Success with functional genomics will facilitate the ability to observe a clinical problem, take it to the benchtop for structural and functional genomic analysis, and return personalized solutions to the bedside in the form of new therapeutic interventions and medicines.

The face of biology has changed forever with the sequencing of the genomes of numerous organisms. Biotechnologies applied to the sequencing of the human genome are also being utilized to sequence the genomes of comparatively simple organisms as well as other mammals. Often, the proteins encoded by the genomes of more simple organisms and the regulation of those genes closely resemble the proteins and gene regulation in humans. Now that the sequencing of the entire genome is a reality, the chore of sorting through human, pathogen, and other organism diversity factors and correlating them with genomic data to provide real pharmaceutical benefits is an active area of research. Comparative genomics is the field of genomics that studies the relationship of genome structure and function across different biological species or strains and thus, provides information about the evolutionary processes that act upon a genome (Raghavachari 2012). Comparative genomics exploits both similarities and differences in the regulatory regions of genes, as well as RNA and proteins of different organisms to infer how selection has acted upon these elements.

Since model organisms are much easier to maintain in a laboratory setting, researchers are actively pursuing "comparative" genomic studies between

multiple organisms. Unlocking genomic data for each of these organisms provides valuable insight into the molecular basis of inherited human disease. *S. cerevisiae*, a yeast, is a good model for studying cancer and is a common organism used in rDNA methodology. For example, it has become well known that women who inherit a gene mutation of the *BRCA1* gene have a high risk, perhaps as high as 85 %, of developing breast cancer before the age of 50 (Petrucci et al. 2011). The first diagnostic product generated from genomic data was the *BRCA1* test for breast cancer predisposition. The gene product of *BRCA1* is a well-characterized protein implicated in both breast and ovarian cancer. Evidence has accumulated suggesting that the Rad9 protein of *S. cerevisiae* is distantly, but significantly, related to the *BRCA1* protein. The fruit fly possesses a gene similar to *p53*, the human tumor suppressor gene. Studying *C. elegans*, an unsegmented vermiform, has provided much of our early knowledge of apoptosis, the normal biological process of programmed cell death. Greater than 90 % of the proteins identified thus far from a common laboratory animal, the mouse, have structural similarities to known human proteins.

Similarly, mapping the whole of a human cancer cell genome will pinpoint the genes involved in cancer and aid in the understanding of cell changes and treatment of human malignancies utilizing the techniques of both functional and comparative genomics (Collins and Barker 2007). In cancer cells, small changes in the DNA sequence can cause the cell to make a protein that doesn't allow the cell to function as it should. These proteins can make cells grow quickly and cause damage to neighboring cells, becoming cancerous. The genome of a cancer cell can also be used to stratify cancer cells identifying one type of cancer from another or identifying a subtype of cancer within that type, such as HER2+ breast cancer. Understanding the cancer genome is a step toward personalized oncology. Numerous projects are underway around the world. Two such projects include the US NIH Cancer Genome Atlas Project (U.S. NIH 2012) and the Sanger Institute Cancer Genome Project (Sanger Institute 2012).

Comparative genomics is being used to provide a compilation of genes that code for proteins that are essential to the growth or viability of a pathogenic organism, yet differ from any human protein (cf. Chap. 22). For example, the worldwide effort to rapidly sequence the severe acute respiratory syndrome (SARS)-associated coronavirus genome to speed up diagnosis, prevent a pandemic, and guide vaccine creation was a great use of genomics in infectious disease. NGS will likely provide an opportunity to place genomics directly into the clinic to enable infectious disease point-of-care applications and thus, selective and superior patient outcomes. Also, genomic mining

of new targets for drug design using genomic techniques may aid the quest for new antibiotics in a clinical environment of increasing incidence of antibiotic resistance.

A valuable resource for performing functional and comparative genomics is the “biobank,” a collection of biological samples for reference purposes. Repositories of this type also might be referred to as biorepositories or named after the type of tissue depending on the exact type of specimens (i.e., tissue banks). Genomic techniques are fostering the creation of DNA banks, the collection, storage, and analysis of hundreds of thousands of specimens containing analyzable DNA. All nucleated cells, including cells from blood, hair follicles, buccal swabs, cancer biopsies, and urine specimens, are suitable DNA samples for analysis in the present or at a later date. DNA banks are proving to be valuable tools for genetics research (Thornton et al. 2005). While in its broadest sense such repositories could incorporate any collection of plant or animal samples, some of the most developed biobanks in the world are devoted to research on various types of cancer. While DNA banks devoted to cancer research have grown the fastest, there also has been an almost explosive growth in biobanks specializing in research on autism, schizophrenia, heart disease, diabetes, and many other diseases.

■ “Omic”-Enabling Technology: Bioinformatics

Structural genomics, functional genomics, proteomics, pharmacogenomics, and other “omic” technologies have generated an enormous volume of genetic and biochemical data to store and analyze. Living in an era of faster computers, bigger and better data storage, and improved methods of data analysis have led to the bioinformatics superhighway that has facilitated the “omic” revolution. Scientists have applied advances in information technology, innovative software algorithms, and massive parallel computing to the ongoing research in biotechnology areas such as genomics to give birth to the fast growing field of bioinformatics (Lengauer and Hartmann 2007; Singh and Somvanshi 2012). The integration of new technologies and computing approaches in the domain of bioinformatics is essential to accelerating the rate of discovery of new breakthroughs that will improve health, well-being, and patient care. Bioinformatics is the application of computer technologies to the biological sciences with the object of discovering knowledge. With bioinformatics, a researcher can now better exploit the tremendous flood of genomic and proteomic data, and more cost-effectively data mine for a drug discovery “needle” in that massive data “haystack.” In this case, data mining refers to the bioinformatics approach of “sifting” through volumes of raw data, identifying and

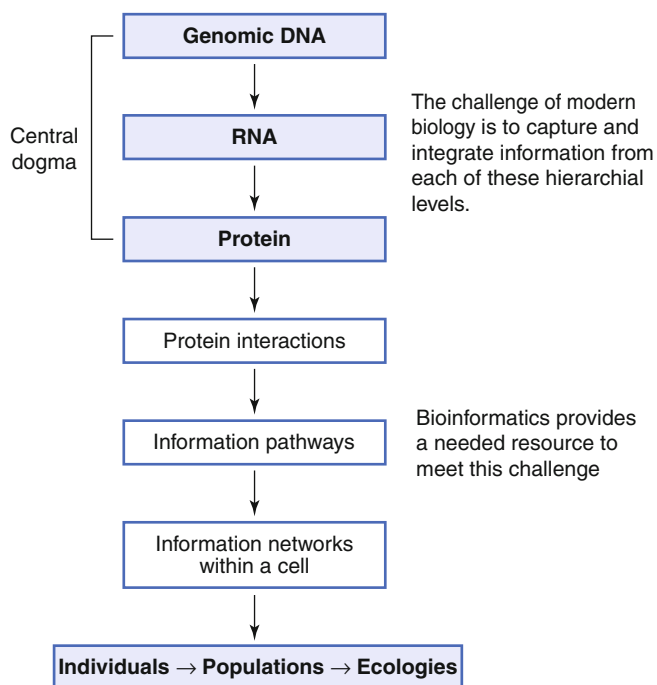


Figure 8.2 ■ The information challenges of systems biology in the genomic era.

extracting relevant information, and developing useful relationships among them.

Modern drug discovery and the commensurate need to better understand and define disease is utilizing bioinformatics techniques to gather information from multiple sources (such as the HGP, functional genomic studies, proteomics, phenotyping, patient medical records, and bioassay results including toxicology studies), integrate the data, apply life science developed algorithms, and generate useful target identification and drug lead identification data. As seen in Fig. 8.2, the hierarchy of information collection goes well beyond the biodata contained in the genetic code that is transcribed and translated. A recent National Research Council report for the US National Academies entitled “Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease” calls for a new data network that integrates emerging research on the molecular basis of diseases with the clinical data from individual patients to drive the development of a more accurate taxonomy of disease that ultimately improves disease diagnosis and patient outcomes (U.S. National Academies 2011). The report notes that challenges include both scientific (technical advances needed to correlate genetic and environmental findings with incidence of disease) and legal and ethical challenges (privacy issues, electronic health records or EHR, etc.).

The entire encoded human DNA sequence alone requires computer storage of approximately 10^9 bits of information: the equivalent of a thousand 500-page books! GenBank (managed by the National Center for Biotechnology Information, NCBI, of the National Institutes of Health), the European Molecular Biology Laboratory (EMBL), and the DNA Data Bank of Japan (DDBJ) are three of the many centers worldwide that collaborate on collecting nucleic acid sequences. These databanks (both public and private) store tens of millions of sequences (Wu et al. 2011a). Once stored, analyzing the volumes of data (i.e., comparing and relating information from various sources) to identify useful and/or predictive characteristics or trends, such as selecting a group of drug targets from all proteins in the human body, presents a Herculean task. This approach has the potential of changing the fundamental way in which basic science is conducted and valid biological conclusions are reached.

Bioinformatics in its multifaceted implementations may be thought of as a technique of “electronic biology” (eBiology), conceptual biology, in silico biology or computational biology. A data-driven tool, the integration of bioinformatics with functional knowledge of the complex biological system under study, remains the critical foundation of any of the omic technologies described above and to follow.

The profession of pharmacy has readily recognized that optimal patient-centered care requires an effective integration of drug information and patient information into a system now known as “pharmacy informatics” (Anderson et al. 2010). Patient information includes data from genomics, proteomics, individual patient characteristics, patient safety, evidence-based medicine, and electronic health records. Drug information includes that found in the primary literature, drug information databases, internet resources, hospital information systems, pharmacy information systems, drug discovery literature, and pharmacogenomic studies. While it is beyond the scope of this chapter to explore pharmacy informatics further, this is becoming an important area for pharmacists to be knowledgeable (Fox 2010).

■ Transcriptomics

Remember that the central dogma of molecular biology is DNA to RNA via the process of transcription and RNA to protein via the process of translation. The transcriptome is the collection of all RNA transcribed elements for a given genome, not only the collection of transcripts that are subsequently translated into proteins (mRNAs). Noncoding transcripts such as noncoding microRNAs (miRNAs) are part of the transcriptome (cf. Chap. 23). The transcriptome represents just a small part of a genome, for instance, only

5 % of the human genome (Lu et al. 2005). The term *transcriptomics* refers to the omic technology that examines the complexity of RNA transcripts of an organism under a variety of internal and external conditions reflecting the genes that are being actively expressed at any given time (with the exception of mRNA degradation phenomena such as transcriptional attenuation) (Subramanian et al. 2005). Therefore, the transcriptome can vary with external environmental conditions, while the genome is roughly fixed for a given cell line (excluding mutations). The transcriptomes of stem cells and cancer cells are of particular interest to better understand the processes of cellular differentiation and carcinogenesis. High-throughput techniques based on microarray technology are used to examine the expression level of mRNAs in a given cell population.

■ Proteomics, Structural Proteomics, and Functional Proteomics

Proteomics is the study of an organism's complete complement of proteins. Proteomics seeks to define the function and correlate that with expression profiles of all proteins encoded within an organism's genome or "proteome" (Veenstra 2010). While functional genomic research will provide an unprecedented information resource for the study of biochemical pathways at the molecular level, certainly a vast array of the approximately 20,000 genes identified in sequencing the human genome will be shown to be functionally important in various disease states (see druggable genome discussion above). These key identified proteins will serve as potential new sites for therapeutic intervention (see Fig. 8.1). The application of functional proteomics in the process of drug discovery has created a field of research referred to as pharmacoproteomics that tries to compare whole protein profiles of healthy persons versus patients with disease. This analysis may point to new and novel targets for drug discovery and personalized medicine (D'Alessandro and Zolla 2010). The transcription and translation of approximately 20,000 human genes can produce hundreds of thousands of proteins due to posttranscriptional regulation and posttranslational modification of the protein products. The number, type, and concentration may vary depending on cell or tissue type, disease state, and other factors. The proteins' function(s) is dependent on the primary, secondary, and tertiary structure of the protein and the molecules they interact with. Less than 30 years old, the concept of proteomics requires determination of the structural, biochemical, and physiological repertoire of all proteins. Proteomics is a greater scientific challenge than genomics due to the intricacy of protein expression and the complexity of 3D protein structure (structural proteomics) as it

relates to biological activity (functional proteomics). Protein expression, isolation, purification, identification, and characterization are among the key procedures utilized in proteomic research.

To perform these procedures, technology platforms such as 2D gel electrophoresis, mass spectrometry, chip-based microarrays (discussed later in this chapter), X-ray crystallography, protein nuclear magnetic resonance (nmr), and phage displays are employed. Initiated in 2002, the Human Proteome Organization (HUPO) completed the first large-scale study to characterize the human serum and plasma proteins, i.e., the human serum and plasma proteome (States et al. 2006). They have spent the past 3 years developing a strategy for the first phase of the Human Proteome Project (Paik et al. 2012). The international consortium Chromosome-Centric Human Proteome Project is attempting to define the entire set of encoded proteins in each human chromosome (Paik et al. 2012). Pharmaceutical scientists anticipate that many of the proteins identified by proteomic research will be entirely novel, possessing unknown functions. This scenario offers not only a unique opportunity to identify previously unknown molecular targets, but also to develop new biomarkers and ultrasensitive diagnostics to address unmet clinical needs (Veenstra 2010). Today's methodology does not allow us to identify valid drug targets and new diagnostic methodologies simply by examining gene sequence information. However, "in silico proteomics," the computer-based prediction of 3D protein structure, intermolecular interactions, and functionality is currently a very active area of research.

Often, multiple genes and their protein products are involved in a single disease process. Since few proteins act alone, studying protein interactions will be paramount to a full understanding of functionality. Also, many abnormalities in cell function may result from overexpression of a gene and/or protein, underexpression of a gene and/or protein, a gene mutation causing a malformed protein, and posttranslational modification changes that alter a protein's function. Therefore, the real value of human genome sequence data will only be realized after every protein coded by the approximately 20,000 genes has a function assigned to it.

■ "Omic"-Enabling Technology: Microarrays

The biochips known as DNA microarrays and oligonucleotide microarrays are a surface collection of hundreds to thousands of immobilized nucleic acid sequences or oligonucleotides in a grid created with specialized equipment that can be simultaneously examined to conduct expression analysis (Amaratunga et al. 2007; Semizarov 2009a). Biochips

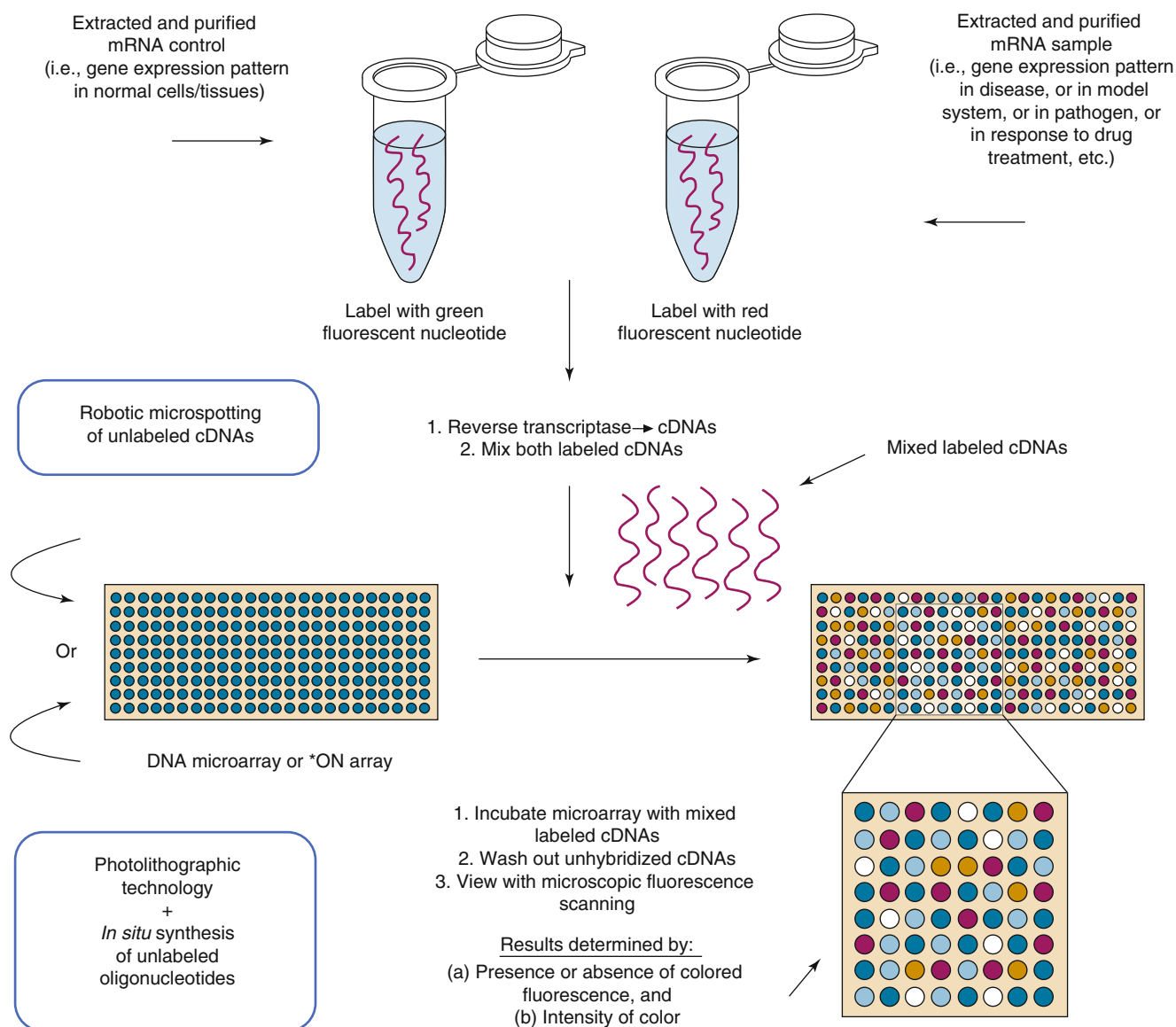


Figure 8.3 ■ Principle of operation of a representative DNA microarray or oligonucleotide (*ON) microarray.

may contain representatives of a particular set of gene sequences (i.e., sequences coding for all human cytochrome P450 isozymes) or may contain sequences representing all genes of an organism. They can produce massive amounts of genetic information (Semizarov 2009a). While the *in vitro* diagnostics market has been difficult to enter, Roche Diagnostics AmpliChip CYP 450 is a FDA-approved diagnostic tool able to determine a patient's genotype with respect to two genes that govern drug metabolism. This information obtained may be useful by a physician to select the appropriate drug and/or dosage for a given patient in the areas of cardiovascular disease, high blood pressure, depression, and others (according to the company).

Commonly, arrays are prepared on nonporous supports such as glass microscope slides. DNA microarrays generally contain high-density microspotted cDNA sequences approximately 1 kb in length representing thousands of genes. The field was advanced significantly when technology was developed to synthesize closely spaced oligonucleotides on glass wafers using semiconductor industry photolithographic masking techniques (see Fig. 8.3). Oligonucleotide microarrays (often called oligonucleotide arrays or DNA chips) contain closely spaced synthetic gene-specific oligonucleotides representing thousands of gene sequences. Microarrays can provide expression analysis for mRNAs. Screening of DNA variation is also possible. Thus, biochips can provide polymorphism

detection and genotyping as well as hybridization-based expression monitoring (Semizarov 2009a).

Microarray analysis has gained increasing significance as a direct result of the genome sequencing studies. Array technology is a logical tool for studying functional genomics since the results obtained may link function to expression. Microarray technology's potential to study key areas of molecular medicine and drug discovery is unlimited at this stage of development. For example, gene expression levels of thousands of mRNA species may be studied simultaneously in normal versus cancer cells, each incubated with potential anticancer drug candidates. Related microarray technologies include protein microarrays, tissue microarrays, cell microarrays (also called transfection microarrays), chemical compound microarrays, and antibody microarrays. The principles are the same, while the immobilized collections differ accordingly.

■ “Omic”-Enabled Technology: Brief Introduction to Biomarkers

Biomarkers are clinically relevant biological features used as indicators of a biologic state, a disease, predisposition to a disease, disease progression, or disease regression (DePrimo 2007). Detection of or concentration change of a biomarker may indicate a particular disease state (e.g., the presence of a certain antibody may indicate an infection), physiology, or toxicity. A change in expression or state of a protein biomarker may correlate with the risk or progression of a disease, with the susceptibility of the disease to a given treatment or the drug's safety profile. Implemented in the form of a medical device, a measured biomarker becomes an *in vitro* diagnostic tool (Williams et al. 2006). While it is well beyond this chapter to provide a detailed discussion of biomarkers, it is important to note that omic technologies including omic-enabled technologies such as microarrays are being developed as clinical measuring devices for biomarkers. Biomarkers enable characterization of patient populations undergoing clinical trials or drug therapy and may accelerate drug development. Modern drug discovery often simultaneously involves biomarker discovery and diagnostic development (Frank and Hargreaves 2003). Drug development scientists are hopeful that the development of appropriate biomarkers will facilitate “go” and “no go” decisions during a potential therapeutic agents development process (Pritchard and Jurima-Romet 2010). Biomarker discovery is closely tied to the other applications of genomics previously described in this chapter. As an indicator of normal biological processes, pathogenic processes, or pharmacological responses to therapeutic intervention, biomarkers may serve as a substitute for a clinical

end point and thus be a surrogate end point (Semizarov 2009b). Biomarkers are now available for a wide range of diseases and conditions including Alzheimer's and Parkinson's disease (Maetzler and Berg 2010), cardiac injury (McLean and Huang 2010), lung injury (Kodavanti 2010), drug-induced liver injury (Ozer et al. 2010), acute kidney injury (Dieterle and Sistare 2010), immunotoxicity (Dieter 2010), various cancers (Kelloff and Sigman 2012), pediatric care (Goldman et al. 2011), and a host of other diseases and biological conditions.

A “theranostic” is a rapid diagnostic, possibly a microarray, measuring a clinically significant biomarker, which may identify patients most likely to benefit or be harmed by a new medication (Warner 2004). Bundled with a new drug (and likely developed in parallel with that drug), the theranostic's diagnosis of the requisite biomarker (e.g., the overexpression of the HER2 gene product in certain breast cancer patients) influences the physician's therapeutic decisions [i.e., prescribing the drug trastuzumab (Herceptin) for HER2 receptor-positive breast cancer patients]. Thus, the diagnostic and the therapy are distinctly coupled = theranostic. The theranostic predicts clinical success of the drug. This example used to introduce the concept of a theranostic is possibly the best example of personalized medicine (see later in this chapter), achieving the best medical outcomes by choosing treatments that work well with a person's genomic profile or with certain characteristics.

■ Metabonomics and Metabolomics

The metabolome consists of the complete set of small molecules that are involved in the energy transmission in the cells by interacting with other biological molecules following metabolic pathways. These metabolites may be metabolic intermediates, hormones and other signaling molecules, and secondary metabolites (Nicholson and Wilson 2003; Patti et al. 2012). The techniques and processes for identifying clinically significant biomarkers of human disease and drug safety have fostered the systematic study of the unique chemical fingerprints that specific cellular processes leave behind, specifically their small molecule metabolite profiles. In January 2007 scientists at the University of Alberta and the University of Calgary finished a draft of the human metabolome (Wishart et al. 2007). They have catalogued and characterized 2,500 metabolites, 1,200 drugs, and 3,500 food components that can be found in the human body. Thus, while genomics and proteomics do not tell the whole story of what might be happening within a cell, metabolic profiling can give an instantaneous snapshot of the physiology of that cell.

High-performance liquid chromatography coupled with sophisticated nuclear magnetic resonance

(NMR) and mass spectrometry (MS) techniques is used to separate and quantify complex metabolite mixtures found in biological fluids to get a picture of the metabolic continuum of an organism influenced by an internal and external environment. The field of metabonomics is the holistic study of the metabolic continuum at the equivalent level to the study of genomics and proteomics. However, unlike genomics and proteomics, microarray technology is little used since the molecules assayed in metabonomics are small molecule end products of gene expression and resulting protein function. The term metabolomics has arisen as the metabolic composition of a cell at a specified time, whereas metabonomics includes both the static metabolite composition and concentrations and the full-time course fluctuations. Coupling the information being collected in biobanks, large collections of patient's biological samples and medical records, with metabonomic and metabolomic studies, will not only detect why a given metabolite level is increasing or decreasing but may reliably predict the onset of disease. Recent research and discoveries in oncology have led to reconsiderations regarding metabolic dysfunctions in cancer cell proliferation and differentiation. Metabolomic studies may be able to interrogate cancer cells for oxidative stress, a leading cause of genetic instability underpinning carcinogenesis, therefore indicative windows during the life of a cancerous cell for optimal therapeutic intervention (D'Alessandro and Zolla 2012). Also, the techniques are finding use in drug safety screening, identification of clinical biomarkers, and systems biology studies (see below).

■ Pharmacogenetics and Pharmacogenomics

It has been noted for decades that patient response to the administration of a drug was highly variable within a diverse patient population. Efficacy as determined in clinical trials is based upon a standard dose range derived from the large population studies. Better understanding of the molecular interactions occurring within the pharmacokinetics phase of a drug's action, coupled with new genetics knowledge and then genomic knowledge of the human have advanced us closer to a rational means to optimize drug therapy. Optimization with respect to the patients' genotype, to ensure maximum efficacy with minimal adverse effects, is the goal. Environment, diet, age, lifestyle, and state of health all can influence a person's response to medicines, but understanding an individual's genetic makeup is thought to be the key to creating personalized drugs with greater efficacy and safety. Approaches such as the related pharmacogenetics and pharmacogenomics promise the advent of "personalized medicine," in which drugs and drug combinations are optimized for each individual's unique genetic makeup. This

chapter will only serve as an introduction, as entire classes are now offered and many books and review articles have been written about pharmacogenetics and pharmacogenomics (Lindpainter 2007; Knoell and Sadee 2009; Grossman and Goldstein 2010; Zdanowicz 2010; Pirmohamed 2011; Brazeau and Brazeau 2011a).

Single-Nucleotide Polymorphisms (SNPs)

While comparing the base sequences in the DNA of two individuals reveals them to be approximately 99.5 % identical, base differences, or polymorphisms, are scattered throughout the genome. The best-characterized human polymorphisms are single-nucleotide polymorphisms (SNPs) occurring approximately once every 1,000 bases in the three billion base pair human genome (Kassam et al. 2005). The DNA sequence variation is a single nucleotide – A, T, C, or G – in the genome difference between members of a species (or between paired chromosomes in an individual). For example, two sequenced DNA fragments from different individuals, AAGTTCCTA to AAGTTCTTA, contain a difference in a single nucleotide. Commonly referred to as "snips," these subtle sequence variations account for most of the genetic differences observed among humans. Thus, they can be utilized to determine inheritance of genes in successive generations. Technologies available from several companies allow for genotyping hundreds of thousands of SNPs for typically under \$1,000 in a couple of days.

Research suggests that, in general, humans tolerate SNPs as a probable survival mechanism. This tolerance may result because most SNPs occur in noncoding regions of the genome. Identifying SNPs occurring in gene coding regions (cSNPs) and/or regulatory sequences may hold the key for elucidating complex, polygenic diseases such as cancer, heart disease, and diabetes and understanding the differences in response to drug therapy observed in individual patients (Grossman and Goldstein 2010; Pirmohamed 2011; US DOE 2012b). Some cSNPs do not result in amino acid substitutions in their gene's protein product(s) due to the degeneracy of the genetic code. These cSNPs are referred to as synonymous cSNPs. Other cSNPs, known as non-synonymous, can produce conservative amino acid changes, such as similarity in side chain charge or size or more significant amino acid substitutions.

While SNPs themselves do not cause disease, their presence can help determine the likelihood that an individual may develop a particular disease or malady. SNPs, when associated with epidemiological and pathological data, can be used to track susceptibilities to common diseases such as cancer, heart disease, and diabetes (Davidson and McNerney 2009). Biomedical researchers have recognized that discovering SNPs linked to diseases will lead potentially to the

identification of new drug targets and diagnostic tests. The identification and mapping of hundreds of thousands of SNPs for use in large-scale association studies may turn the SNPs into biomarkers of disease and/or drug response. Genetic factors such as SNPs are believed to likely influence the etiology of diseases such as hypertension, diabetes, and lipidemias directly and via effects on known risk factors (Davidson and McInerney 2009). For example, in the chronic metabolic disease type 2 diabetes, a strong association with obesity and its pathogenesis includes defects of both secretion and peripheral actions of insulin. The association between type 2 diabetes and SNPs in three genes was detected in addition to a cluster of new variants on chromosome 10q. However, heritability values range only from 30 to 70 % as type 2 diabetes is obviously a heterogeneous disease etiologically and clinically. Thus, SNPs, in the overwhelming majority of cases, will likely not be indicators of disease development by themselves.

The projected impact of SNPs on our understanding of human disease led to the formation of the SNP Consortium in 1999, an international research collaboration involving pharmaceutical companies, academic laboratories, and private support. In the USA, the DOE and the NIH Human Genome programs helped establish goals to identify and map SNPs. The goals included the development of rapid large-scale technologies for SNP identification, the identification of common variants in the coding regions of most identified genes, the creation of an SNP map of at least 100,000 elements that may serve as future biomarkers, the development of knowledge that will aid future studies of sequence variation, and the creation of public resources of DNA samples, cell lines, and databases (US DOE 2012b). SNP databases include a database of the SNP Consortium (TSC), the dbSNP database from the National Center for Biotechnology Information (NCBI), and the Human Genome Variation Database (HGVbase).

Pharmacogenetics Versus Pharmacogenomics

In simplest terms, pharmacogenomics is the whole genome application of pharmacogenetics, which examines the single gene interactions with drugs. Tremendous advances in biotechnology are causing a dramatic shift in the way new pharmaceuticals are discovered, developed, and monitored during patient use. Pharmacists will utilize the knowledge gained from genomics and proteomics to tailor drug therapy to meet the needs of their individual patients employing the fields of pharmacogenetics and pharmacogenomics (Kalow 2009; Knoell and Sadee 2009; Grossman and Goldstein 2010; Zdanowicz 2010; Pirmohamed 2011; Brazeau and Brazeau 2011a).

Pharmacogenetics is the study of how an individual's genetic differences influence drug action, usage, and dosing. A detailed knowledge of a patient's pharmacogenetics in relation to a particular drug therapy may lead to enhanced efficacy and greater safety. Pharmacogenetic analysis may identify the responsive patient population prior to administration, i.e., personalized medicine. The field of pharmacogenetics is over 50 years old, but is undergoing renewed, exponential growth at this time. Of particular interest in the field of pharmacogenetics is our understanding of the genetic influences on drug pharmacokinetic profiles such as genetic variations affecting liver enzymes (i.e., cytochrome P450 group) and drug transporter proteins and the genetic influences on drug pharmacodynamic profiles such as the variation in receptor protein expression (Abla and Kroetz 2009; Frye 2009; Johnson 2009; Kalow 2009; Wang 2009).

In contrast, pharmacogenomics is linked to the whole genome, not an SNP in a single gene. It is the study of the entire genome of an organism (i.e., human patient), both the expressed and the non-expressed genes in any given physiologic state. Pharmacogenomics combines traditional pharmaceutical sciences with annotated knowledge of genes, proteins, and single-nucleotide polymorphisms. It might be viewed as a logical convergence of the stepwise advances in genomics with the growing field of pharmacogenetics. Incorrectly, the definitions of pharmacogenetics and pharmacogenomics are often used interchangeably. Whatever the definitions, they share the challenge of clinical translation, moving from bench top research to bedside application for patient care.

Genome-Wide Association Studies (GWAS)

The methods of genome-wide association studies (GWAS), also known as whole genome association studies, are powerful tools to identify genetic loci that affect, for instance, drug response or susceptibility to adverse drug reactions (Davidson and McInerney 2009; Wu et al. 2011a). These studies are an examination of the many genetic variations found in different individuals to determine any association between a variant (genotype) and a biological trait (phenotype). The majority of GWAS typically study associations between SNPs and drug response or SNPs and major disease. While the first GWAS was published only in 2005, they have emerged as important tools with, as per data from the NHGRI GWAS Catalog, hundreds of thousands of individuals now tested in over 1,200 human GWAS examining over 200 diseases and traits and 6,229 SNPs as of early 2012 (Hindroff et al. 2012). While believed to be a core driver in the vision for personalized medicine, GWAS when coupled to the HapMap Project (an international effort to identify and map regions of DNA

sequence nearly identical within the broad population) to date have been plagued by inconsistencies in genotypes, difficulties in assigning phenotypes, and overall quality of the data (Hong et al. 2010; Miclaus et al. 2010; Wu et al. 2011a, b, c). Challenges have included difficulties identifying the key genetic loci due to two or more genes with small and additive effects on the trait (epistasis), the trait caused by gene mutations at several different chromosomal loci (locus heterogeneity), environmental causes modifying expression of the trait or responsible for the trait, and undetected population structure in the study such as those arising when some study members share a common ancestral heritage (Brazeau and Brazeau 2011b). The practical use of this approach and its introduction into the everyday clinical setting remain a challenge, but will undoubtedly be aided by new next-generation sequencing techniques, enhanced bioinformatics capabilities, and better genomic understanding.

■ On the Path to Personalized Medicine: A Brief Introduction

Much of modern medical care decision-making is based upon observations of successful diagnosis and treatment at the larger population level. There is an expectation, however, that healthcare is starting to undergo a revolutionary change as new genomic and other “omic” technologies become available to the clinic that will better predict, diagnose, monitor, and treat disease at the level of the specific patient. A goal is match individual patients with the most effective and safest drugs and doses. Direct-to-consumer genomic tests became more readily available (such as 23andMe, Navigenics, and deCODE Genetics) (McGuire et al. 2010). Academic medical centers have begun to demonstrate the feasibility of routine clinical genotyping as a means of informing pharmacotherapeutic treatment selection in oncology (Tursz et al. 2011). Likewise, demonstration projects in pharmacogenomics entered pharmacy practice in several settings (Koomer and Ansong 2010; Crews et al. 2011; Padgett et al. 2011). Pharmacy education curricula are evolving to prepare graduates practice in a personalized medicine environment (Lee et al. 2009; Krynetskiy and Calligaro 2009; Koomer 2010; Murphy et al. 2010; Zembles 2010). This approach is entirely consistent with the concept of patient-centered care to improve patient outcomes (Clancy and Collins 2010; Waldman and Terzic 2011; Kaye et al. 2012).

Modern genomics, proteomics, metabolomics, pharmacogenomics, epigenomics (to be discussed later in this chapter), and other technologies, implemented in the clinic in faster and less expensive instrumentation and methodologies, are now being introduced to identify genetic variants, better inform healthcare pro-

viders about their individual patient, tailor evidence-based medical treatment, and suggest rational approaches toward preventative care. The hopes and realities of personalized medicine (sometimes referred to as part of “molecular medicine”), pharmacotherapy informed by a patient’s individual genomic and proteomic information, are global priorities (Knoell and Sadee 2009; Grossman and Goldstein 2010; Rahbar et al. 2011). As a pharmaceutical biotechnology text, our limited discussion here will focus on personalized medicine in a primarily pharmacogenomic and pharmacogenetic context. However, other genomic-type technologies including GWAS, next-generation sequencing, proteomics, and metabolomics will be crucial for the successful implementation of personalized medicine. The hope is that “omic” science will bring predictability to the optimization of drug selection and drug dosage to assure safe and effective pharmacotherapy (Fig. 8.4).

For our discussion, it is again important to recognize that pharmacogenetics and pharmacogenomics are subtly different (Brazeau and Brazeau 2011a). Pharmacogenomics introduces the additional element of our present technical ability to pinpoint patient-specific DNA variation using genomic techniques. The area looks at the genetic composition or genetic variations of an organism and their connection to drug response. Variations in target pathways are studied to understand how the variations are manifested and how they influence response. While overlapping fields of study, pharmacogenomics is a much newer term that correlates an individual patient’s DNA variation (SNP level of variation knowledge rather than gene level of variation knowledge) with his or her response to pharmacotherapy. Personalized medicine will employ both technologies.

Optimized personalized medicine utilizing pharmacogenomic knowledge would not only spot disease before it occurs in a patient or detect a critical variant that will influence treatment but should increase drug efficacy upon pharmacotherapy and reduce drug toxicity. Also, it would facilitate the drug development process (see Fig. 8.1) including improving clinical development outcomes, reducing overall cost of drug development, and leading to development of new diagnostic tests that impact on therapeutic decisions (Grossman and Goldstein 2010; Zineh and Huang 2011). Individualized optimized pharmacotherapy would first require a detailed genetic analysis of a patient, assembling a comprehensive list of SNPs. Pharmacogenomic tests most likely in the form of microarray technology and based upon clinically validated biomarkers would be administered to pre-identify responsive patients before dosing with a specific agent. Examples of such microarray-based

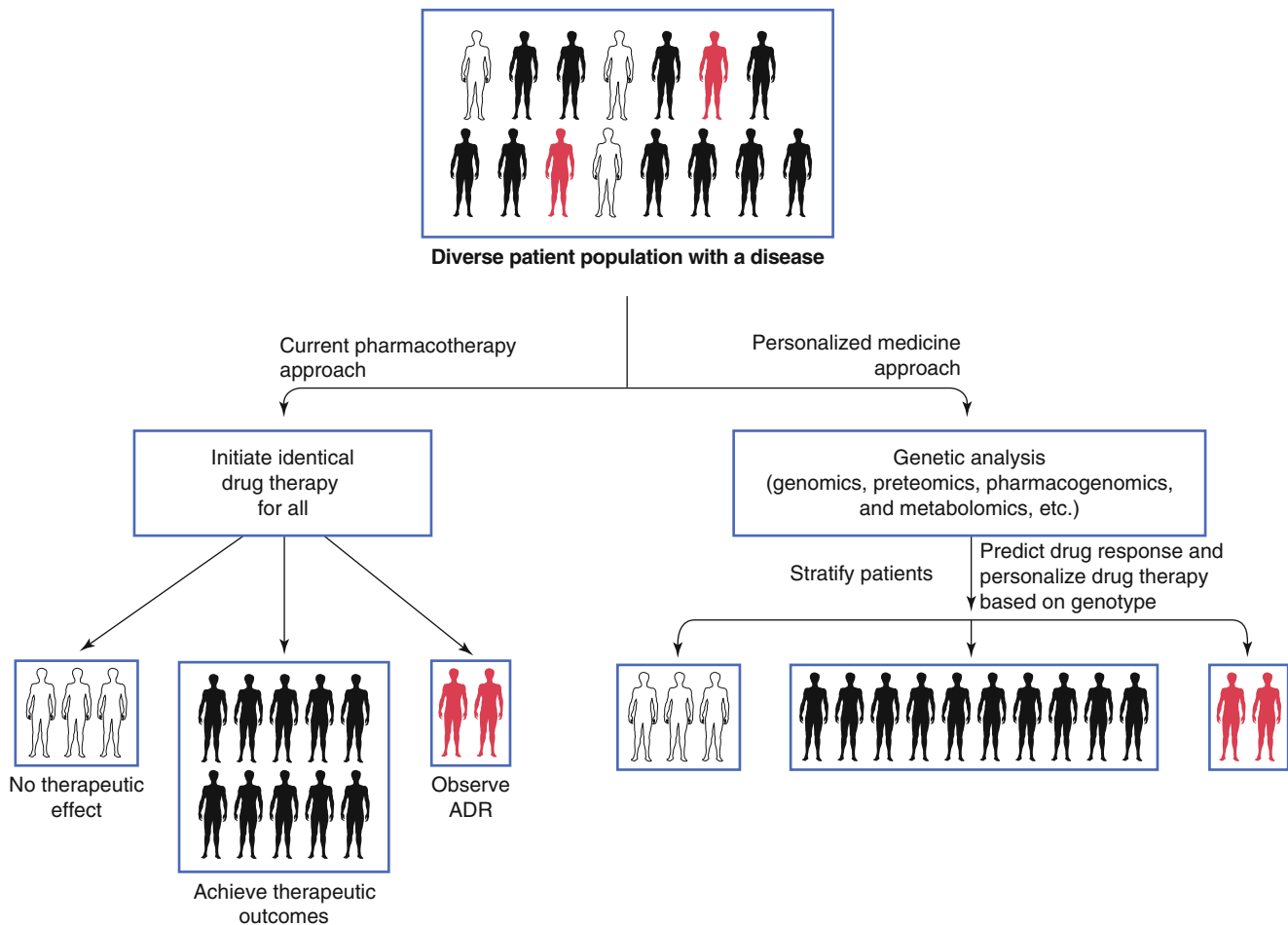


Figure 8.4 ■ The role of "omic" technologies in personalized medicine.

diagnostics are the FDA-approved AmpliChip P450 from Roche to screen a patient for the presence of any of 27 SNPs in *CYP2D6* and *CYP2C19*, the Infiniti 2C9 & VKORC1 Multiplex Assay for warfarin therapy from AutoGenomics, and the Pathwork Tissue of Origin test of 15 common malignant tumor types to better focus treatment options. The impact of the patient's SNPs on the use of new or existing drugs would thus be predicted and individualized drug therapy would be identified that assures maximal efficacy and minimal toxicity (Topol 2010).

Personalized medicine would also require knowledge of an individual patient's genomic profile to help identify potential drug responders and nonresponders. This might be accomplished by testing for the presence or absence of critical biomarkers that may be associated with prediction of response rates. The US FDA provides an online list of all FDA-approved drugs with pharmacogenomic information in their labels (black boxes). Some, but not all, of the labels include specific actions to be taken based on genetic information. The drug labels contain information on genomic biomark-

ers that may be predictive of drug exposure and clinical response rate, risk of adverse reactions, genotype-specific dosing, susceptibility to a specific mechanism of drug action, or polymorphic drug target and disposition genes. Rather than reproducing this table in whole or part in this text, the reader may access it in its constantly updated form at www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm.

It is well understood that beyond genomics and proteomics, a patient's behavioral and environmental factors influence clinical outcomes and susceptibility to disease. Emerging fields of nutrigenomics and envirogenomics are studying these additional layers of complexity. Personalized medicine will become especially important in cases where the cost of testing is less than either the cost of the drug or the cost of correcting adverse drug reactions caused by the drug. Pharmaceutical care would begin by identifying a patient's susceptibility to a disease, then administering the right drug to the right patient at the right time. For example, the monoclonal antibody trastuzumab

(Herceptin) is a personalized breast cancer therapy specifically targeted to the HER2 gene product (25–30 % of human breast cancers overexpress the human epidermal growth factor receptor, HER2 protein) (Kolesar 2009). Exhibiting reduced side effects as compared to standard chemotherapy due to this protein target specificity, trastuzumab is not prescribed to treat a breast cancer patient unless the patient has first tested positive for HER2 overexpression. While currently an immunohistochemical assay, not a sophisticated DNA microarray assay, the example shows the power of such future tests.

The success of targeted therapy for personalized medicine has fostered the concept that the era of the blockbuster drug may be over and will be replaced by the “niche buster” drug, a highly effective medicine individualized for a small group of responding patients identified by genomic and proteomic techniques. Also, while numerous articles predicted that pharmacogenomics would revolutionize medicine, the initial predictions have not been lived up to the hype due to statistical, scientific, and commercial hurdles. With more than 11 million SNP positions believed to be present in the human population, large-scale detection of genetic variation holds the key to successful personalized medicine (Pennisi 2010; Reardon 2011; Baker 2012). Correlation of environmental factors, behavioral factors, genomic and proteomic factors (including pharmacogenomic and metabolomic factors), and phenotypical observables across large populations remains a daunting data-intensive challenge. Yet, pharmacogenetics and pharmacogenomics are having an impact on modern medicine.

Human Genomic Variation Affecting Drug Pharmacokinetics

Genetic variation associated with drug metabolism and drug transport, processes resulting from products of gene expression (metabolic enzymes and transport proteins, respectively) play a critical role in determining the concentration of a drug in its active form at the site of its action and also at the site of its possible toxic action(s). Thus, pharmacogenetic and pharmacogenomic analysis of drug metabolism and drug transport is important to a better clinical understanding of and prediction of the effect of genetic variation on drug effectiveness and safety (Abla and Kroetz 2009; Frye 2009; Wang 2009; Cox 2010; Weston 2010).

It is well recognized that specific drug metabolizer phenotypes may cause adverse drug reactions. For instance, some patients lack an enzymatically active form, have a diminished level, or possess a modified version of CYP2D6 (a cytochrome P450 allele) and will metabolize certain classes of pharmaceutical agents differently to other patients expressing the

Enzyme	Common variant	Potential consequence
CYP1A2	CYP1A2*1F	Increased inducibility
CYP1A2	CYP1A2*1K	Decreased metabolism
CYP2A6	CYP2A6*2	Decreased metabolism
CYP2B6	CYP2B6*5	No effect
CYP2B6	CYP2B6*6	Increased metabolism
CYP2B6	CYP2B6*7	Increased metabolism
CYP2C8	CYP2C8*2	Decreased metabolism
CYP2C9	CYP2C9*2	Altered affinity
CYP2C9	CYP2C9*3	Decreased metabolism
CYP2D6	CYP2D6*10	Decreased metabolism
CYP2D6	CYP2D6*17	Decreased metabolism
CYP2E1	CYP2E1*2	Decreased metabolism
CYP3A7	CYP3A7*1C	Increased metabolism
Flavin-containing monooxygenase 3	FMO3*2	Decreased metabolism
Flavin-containing monooxygenase 3	FMO3*4	Decreased metabolism

Data from references noted in the Pharmacogenomics section of Chap. 8

Table 8.2 ■ Some selected examples of common drug metabolism polymorphisms and their pharmacokinetic consequences.

native active enzyme. All pharmacogenetic polymorphisms examined to date differ in frequency among racial and ethnic groups. For example, CYP2D6 enzyme deficiencies may occur in ≤ 2 % Asian patients, ≤ 5 % black patients, and ≤ 11 % white patients (Frye 2009). A diagnostic test to detect CYP2D6 deficiency could be used to identify patients that should not be administered drugs metabolized predominantly by CYP2D6. Table 8.2 provides some selected examples of common drug metabolism polymorphisms and their pharmacokinetic consequences.

With the burgeoning understanding of the genetics of warfarin metabolism, warfarin anticoagulation therapy is becoming a leader in pharmacogenetic analysis for pharmacokinetic prediction (Limdi and Rettie 2009; Momary and Crouch 2010; Bungard et al. 2011; McDonagh et al. 2011). Adverse drug reactions (ADRs) for warfarin account for 15 % of all ADRs in the USA, second only to digoxin. Warfarin dose is adjusted with the goal of achieving an INR (International Normalized Ratio = ratio of patient’s prothrombin time as compared to that of a normal control) of 2.0–3.0. The clinical challenge is to limit hemorrhage, the primary ADR, while achieving the optimal degree of protection against thromboembolism. Deviation in the INR has been shown to be the strongest risk factor for bleeding

complications. The major routes of metabolism of warfarin are by CYP2C9 and CYP3A4. Some of the compounds, which have been identified to influence positively or negatively warfarin's INR, include cimetidine, clofibrate, propranolol, celecoxib (a competitive inhibition of CYP2C9), fluvoxamine (an inhibitor of several CYP enzymes), various antifungals and antibiotics (e.g., miconazole, fluconazole, erythromycin), omeprazole, alcohol, ginseng, and garlic. Researchers have determined that the majority of individual patient variation observed clinically in response to warfarin therapy is genetic in nature, influenced by the genetic variability of metabolizing enzymes, vitamin K cycle enzymes, and possibly transporter proteins. The CYP2C9 genotype polymorphisms alone explain about 10 % of the variability observed in the warfarin maintenance dose. Figure 8.5 shows the proteins involved in warfarin action and indicates the pharmacogenomic variants that more significantly influence warfarin therapy optimal outcome.

Studies at both the basic research and clinical level involve the effect of drug transport proteins on the pharmacokinetic profile of a drug (Abla and Kroetz 2009). Some areas of active study of the effect of genetic variation on clinical effectiveness include efflux transporter proteins (for bioavailability, CNS exposure, and tumor resistance) and neurotransmitter uptake transporters (as valid drug targets). Novel transporter

proteins are still being identified as a result of the Human Genome Project and subsequent proteomic research. More study is needed on the characterization of expression, regulation, and functional properties of known and new transporter proteins to better assess the potential for prediction of altered drug response based on transporter genotypes.

Human Genomic Variation Affecting Drug Pharmacodynamics

Genomic variation affects not only the pharmacokinetic profile of drugs, it also strongly influences the pharmacodynamic profile of drugs via the drug target. To understand the complexity of most drug responses, factors influencing the expression of the protein target directly with which the drug interacts must be studied. Targets include the drug receptor involved in the response as well as the proteins associated with disease risk, pathogenesis, and/or toxicity including infectious disease (Johnson 2009; Rogers 2009; Webster 2010). There are increasing numbers of prominent examples of inherited polymorphisms influencing drug pharmacodynamics. To follow on the warfarin example above (see Fig. 8.5), the majority of individual patient variation observed clinically in response to anticoagulant therapy is genetic in nature. However, the CYP2C9 genotype polymorphisms alone only explain about 10 % of the variability observed in the warfarin

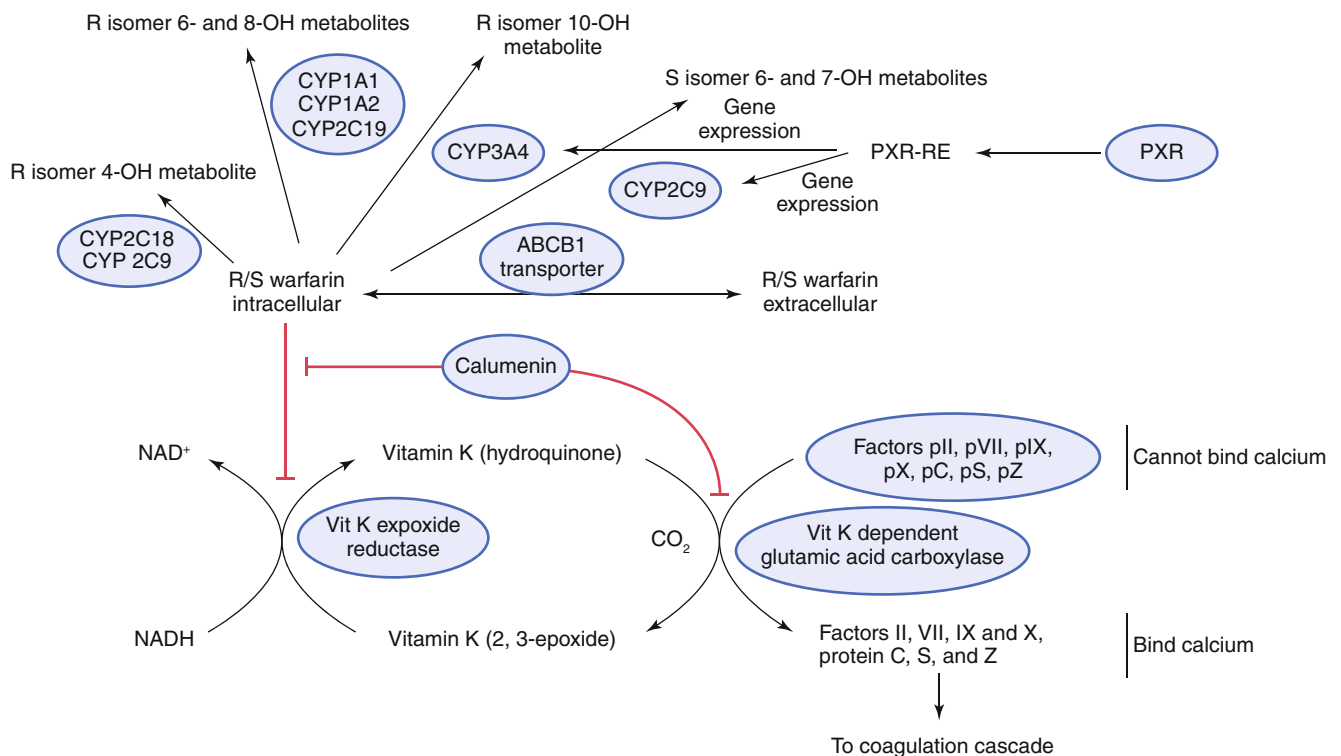


Figure 8.5 ■ Critical pharmacogenomic variants affecting warfarin drug action and ADR.

maintenance dose. Warfarin effectiveness is also influenced by the genetic variability of vitamin K cycle enzymes. The drug receptor for warfarin is generally recognized as vitamin K epoxide reductase, the enzyme that recycles vitamin K in the coagulation cascade. Vitamin K epoxide reductase complex1 (VKORC1) has been determined to be highly variant with as much as 50 % of the clinical variability observed for warfarin resulting from polymorphisms of this enzyme.

Associations have been implicated between drug response and genetic variations in targets for a variety of drugs including antidepressants (G-protein β 3), antipsychotics (dopamine D2, D3, D4; serotonin 5HT2A, 5HT2C), sulfonyleureas (sulfonyleurea receptor protein), and anesthetics (ryanodine receptor) (Johnson 2009). In addition, similar associations have been studied for drug toxicity and disease polymorphisms including abacavir (major histocompatibility proteins; risk of hypersensitivity), cisapride and terfenadine (HERG, KvLQT1, Mink, MiRP1; increased risk of drug-induced torsade de pointes), and oral contraceptives (prothrombin and factor V; increased deep vein thrombosis) (Johnson 2009). Likewise, similar associations for efficacy are known such as statins (apolipoprotein E; enhanced survival prolongation with simvastatin) and tacrine (apolipoprotein E; clinical improvement of Alzheimer's symptoms) (Johnson 2009).

Value of Personalized Medicine in Disease

Due to the intimate role of genetics in carcinogenesis, personalized medicine is rapidly becoming a success story in oncology based on genetic profiling using proteomic analyses of tumor biopsies (Garnett 2012; Kelloff and Sigman 2012; Shaw and Johnson 2012). As described above, targeted cancer therapies such as trastuzumab (Herceptin) are successful and are viewed as the way of the future. Also, clinically important polymorphisms predict increased toxicity in patients with cancer being treated with the chemotherapeutic drugs, for example, 6-mercaptopurine (thiopurine methyltransferase *2, *3A, and *3C variants), 5-fluorouracil (5-FU) (dihydropyrimidine dehydrogenase *2A variant), and irinotecan (UGT1A1*28 allele; FDA-approved Invader UGT1A1 Molecular Assay diagnostic available to screen for presence of this allele associated with irinotecan toxicity) (Kolesar 2009). Likewise, clinically important pharmacogenetics predicts efficacy in oncology patients treated with 5-FU (thymidylate synthase *2 and *3C variants) (Petros and Sharma 2009).

A classic application of pharmacogenetics is our present understanding of the potentially fatal hematopoietic toxicity that occurs in some patients administered standard doses of the antileukemic agents azathioprine, mercaptopurine, and thioguanine (Zhou

2006; Petros and Sharma 2009). These drugs are metabolized by the enzyme thiopurine methyltransferase (TPMT) to the inactive S-methylated products. Gene mutations (polymorphisms) may occur in as many as 11 % of patients resulting in decreased TPMT-mediated metabolism of the thiopurine drugs. A diagnostic test for TPMT is now available and used clinically. Identified patients with poor TPMT metabolism may need their drug dose lowered 10–15-fold. Mechanisms of multidrug resistance to cancer drugs are influenced by genetic differences. A number of polymorphisms in the MDR-1 gene coding for P-glycoprotein, the transmembrane protein drug efflux pump responsible for multidrug resistance, have been identified. One, known as the T/T genotype and correlated with decreased intestinal expression of P-glycoprotein and increased drug bioavailability, has an allele frequency of 88 % in African-American populations, yet only approximately 50 % in Caucasian-American populations (Kolesar 2009).

Pharmacogenetic and pharmacogenomic analysis of patients is being actively studied in many disease states. However, a detailed discussion goes beyond what this introduction may provide. The reader is encouraged to read further in the pharmacogenetic/pharmacogenomic-related references at the end of this chapter. Some examples include infectious disease (genetic predisposition to infection in the host; Rogers 2009), cardiovascular disease (genes linked to heart failure and treating hypertension, warfarin anticoagulant therapy, lipid lowering drugs; Zineh and Pacanowski 2009), psychiatry (the roles of drug metabolism and receptor expression in drug response rates for antidepressants and antipsychotic drugs, weight gain from antipsychotics; Ellingrod et al. 2009), asthma (leukotriene inhibitors and beta-agonists; Blake et al. 2009), and transplantation (cyclosporine metabolism and multidrug resistance efflux mechanisms; Burckert 2009).

Challenges in Personalized Medicine

There are many keys to success for personalized medicine that hinge on continued scientific advancement. While it is great for the advancement of the genomic sciences, some have questioned how good it is for patients at this stage of its development due to exaggerated claims falling short of the predictive and preventative healthcare paradigm promised (Browman et al. 2011; Nature Biotechnology Editorial Staff 2012). The pace of advancement has been slower than promised (Zuckerman and Milne 2012). There are also economic, societal, and ethical issues that must be addressed to successfully implement genetic testing-based individualized pharmacotherapy (Huston 2010). It is fair to state that most drugs will not be effective in

all patients all of the time. Thus, the pressure of payers to move from a “payment for product” to a “payment for clinically significant health outcomes” model is reasonable. The use of omic health technologies and health informatics approaches to stratify patient populations for drug effectiveness and drug safety is a laudable goal. However, the technologies are currently quite expensive and the resulting drug response predictability is now just being validated clinically. Cost-effectiveness and cost-benefit analyses are limited at this date (Chalkidou and Rawlins 2011). Also, the resulting environment created by these technologies in the context of outcomes expectations and new drug access/reimbursement models will give rise to a new pharmaceutical business paradigm that is still evolving and not well understood.

In 2005, the FDA approved what some referred to as the “first racially targeted drug,” BiDil (isosorbide/hydralazine; from NitroMed) (Branca 2005). Omic technologies were not generally involved in the development and approval process. Based on the analysis of health statistics suggesting that the rate of mortality in blacks with heart disease is twice as high in whites in the 45–64 age group, a clinical trial of this older drug combination in 1,050 African-Americans was conducted and the 43 % improvement in survival in the treatment arm resulted in FDA approval of the drug exclusively in African-Americans. Yet, modern anthropology and genetics have shown that while race does exist as a real social construct, there are no genetically distinguishable human racial groups (Ossorio 2004). Thus, attributing observed differences in biomedical outcomes and phenotypical observations to genetic differences among races is problematic and ethically challenging. Race is likely just a surrogate marker for the environmental and genetic causes of disease and response to pharmacotherapy. Now, factor in the introduction of omic technologies broadly into healthcare in a manner to segregate patient populations based on genomic and proteomic characteristics. It is obvious that these modern technologies pose provocative consequences for public policy (including data protection, insurability, and access to care), and these challenges must be addressed by decision makers, scientists, healthcare providers, and the public for personalized medicine to be successful (for further insight into this complex area, please read Brazeau and Brazeau 2011c). In conclusion, even with challenges and questioned progress, personalized medicine is a global concern and an unprecedented opportunity if the science and the clinic can both succeed.

Epigenetics and Epigenomics

DNA is the heritable biomolecule that contains the genetic information resulting in phenotype from

parent to offspring. Modern genomics, GWAS and SNP analyses, confirm this and identify genetic variants that may be associated with a different phenotype. However, genome-level information alone does not generally predict phenotype at an individual level (Daxinger and Whitelaw 2012). For instance, researchers and clinicians have known for some time that an individual’s response to a drug is affected by their genetic makeup (DNA sequence, genotype) and a set of disease and environmental characteristics working alone or in concert to determine that response. Research in animal models has suggested that in addition to DNA sequence, there are a number of other “levels” of information that influence transcription of genomic information. As you are aware, every person’s body contains trillions of cells, all of which have essentially the same genome and, therefore, the same genes. Yet some cells are optimized for development into one or more of the 200+ specialized cells that make up our bodies: muscles, bones, brain, etc. For this to transpire from within the same genome, some genes must be turned on or off at different points of cell development in different cell types to affect gene expression, protein production, and cell differentiation, growth, and function. There is a rapidly evolving field of research known as epigenetics (or epigenomics) that can be viewed as a conduit between genotype and phenotype. Epigenetics literally means “above genetics or over the genetic sequence.” It is the factor or factors that influence cell behavior by means other than via a direct effect on the genetic machinery. Epigenetic regulation includes DNA methylation and covalent histone modifications (Fig. 8.6) and is mitotically and/or meiotically heritable changes in gene expression that result without a change in DNA sequence (Berger et al. 2009). Epigenomics is the merged science of genomics and epigenetics (Raghavachari 2012). Functionally, epigenetics acts to regulate gene expression, gene silencing during genomic imprinting, apoptosis, X-chromosome inactivation, and tissue-specific gene activation (such as maintenance of stem cell pluripotency) (Garske and Denu 2009).

The more we understand epigenetics and epigenomics, the more we are likely to understand those phenotypic traits that are not a result of genetic information alone. Epigenetics/epigenomics may also explain low association predictors found in some pharmacogenetic/pharmacogenomic studies. Etiology of disease, such as cancer, likely involves both genetic variants and epigenetic modifications that could result from environmental effects (Bjornsson et al. 2004; Jirtle and Skinner 2007). Age also likely influences epigenetic modifications as studies of identical twins show greater differences in global DNA methylation in older rather than younger sets of twins (Feinberg et al. 2010).

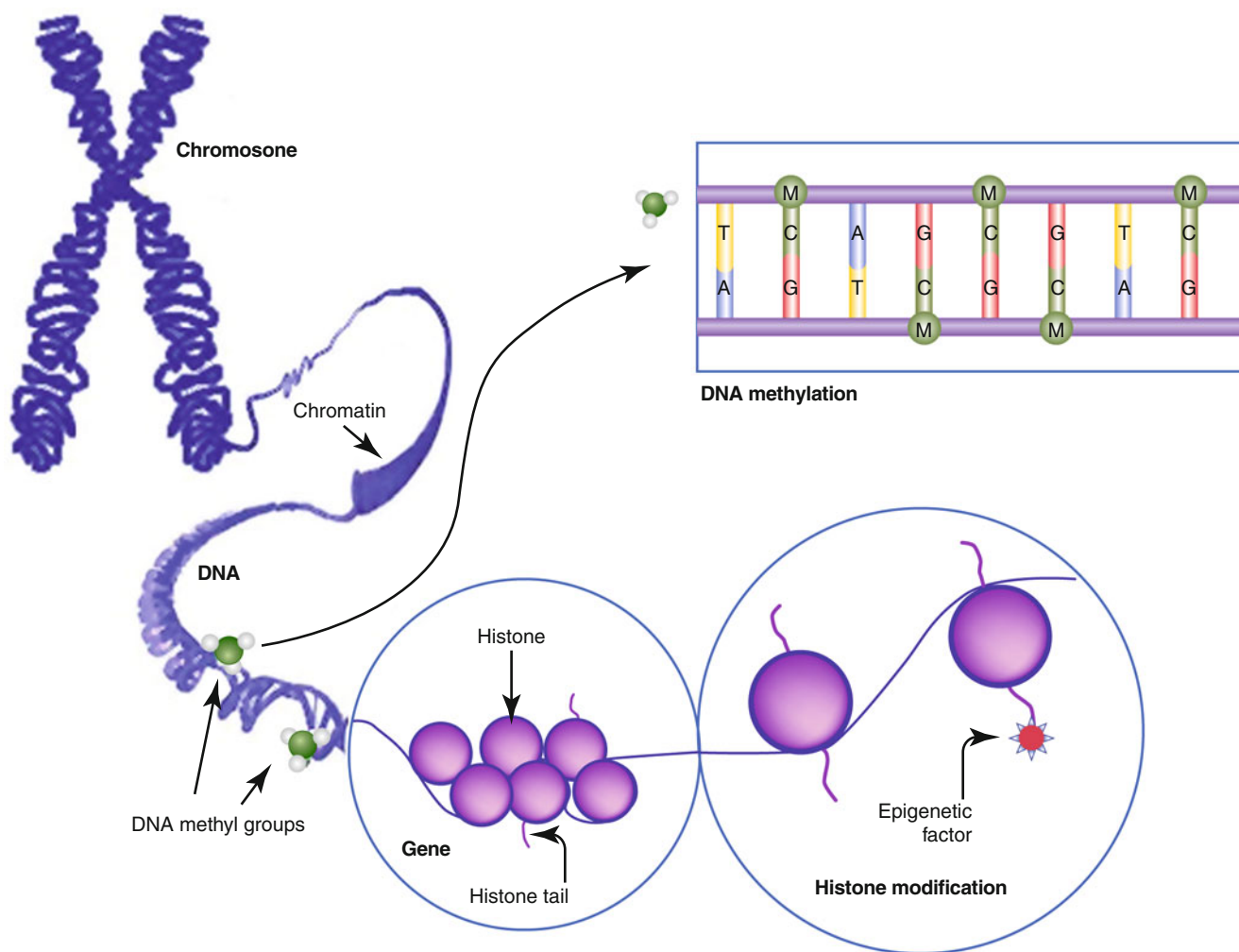


Figure 8.6 ■ Epigenetic regulation via DNA methylation, histone modifications, and chromatin structure.

Abnormal epigenetic regulation is likely a feature of complex diseases such as diabetes, cancer, and heart disease (Chen and Zhang 2011; Hamm and Costa 2011; Rakyan et al. 2011). Therefore, epigenetics targets are being explored for drug design, especially those observed in cancer (Woster 2010). The first generation of FDA-approved epigenetics-based drugs is available with two DNA demethylating agents (5-azacytidine and decitabine) and two histone deacetylase (HDAC) inhibitors (vorinostat and romidepsin). These have been approved mainly for the treatment of blood cancers, in particularly myelodysplastic syndromes (MDS).

One of the most studied and best understood molecular mechanisms of epigenetic regulation is methylation of cytosine residues at specific positions in the DNA molecule (Fig. 8.6) (Portela and Esteller 2010). Another mechanism of epigenomic control appears to occur at the level of chromatin. In the cell, DNA is wrapped around 8 different histone proteins to form

chromatin. Packaging of DNA into chromatin can render large regions of the DNA inaccessible and prevent processes such as DNA transcription from occurring. Epigenetic regulation of histone proteins can be by chemical modification including acetylation, methylation, sumoylation and ubiquitylation (Herceg and Murr 2011). Each can cause structural changes in chromatin affecting DNA accessibility. Non-protein-coding RNAs, known as ncRNAs, have also been shown to contribute to epigenetic regulation as have mRNAs which can be processed and participate in various interference pathways (Collins and Schonfeld 2011).

■ Toxicogenomics

Toxicogenomics, related to pharmacogenomics, combines toxicology, genetics, molecular biology, and environmental health to elucidate the response of living organisms to stressful environments or xenobiotic agents based upon their genetic makeup (Rockett 2003). While toxicogenomic studies how the genome

responds to toxic exposures, pharmacogenetics studies how an individual's genetic makeup affects his/her response to environmental stresses and toxins such as carcinogens, neurotoxins, and reproductive toxins (Smith 2010). Toxicogenomics can be very useful in drug discovery and development as new drug candidates can be screened through a combination of gene expression profiling and toxicology to understand gene response, identify general mechanisms of toxicity, and possibly predict drug safety (Furness 2002; Blomme 2009). There have been suggestions that toxicogenomics may decrease the time needed for toxicological investigations of new drug candidates and reduce both cost and animal usage versus conventional toxicity studies.

Genomic techniques utilized in toxicogenomic studies include gene expression level profiling, SNP analysis of the genetic variation, proteomics, and/or metabolomic methods so that gene expression, protein production, and metabolite production may be studied (Raghavachari 2012). The rapid growth in next-generation DNA sequencing capability may drive a conversion from microarrays now most commonly used for SNP analysis) to NGS technology.

Toxicogenomic studies attempt to discover associations between the development of drug toxicities and genotype. Clinicians and researchers are attempting to correlate genetic variation in one population to the manifestations of toxicity in other populations to identify and then to predict adverse toxicological effects in clinical trials so that suitable biomarkers for these adverse effects can be developed. Using such methods, it would then theoretically possible to test an individual patient for his or her susceptibility to these adverse effects before prescribing a medication. Patients that would show the marker for an adverse effect would be switched to a different drug. Therefore, toxicogenomics will become increasingly more powerful in predicting toxicity as new biomarkers are identified and validated. First described in 1999, the field is in its infancy, yet emerging rapidly. Much of the new toxicogenomic technology is developing in the pharmaceutical industry and other corporate laboratories.

■ Glycomics and Glycobiology

The novel scientific field of glycomics, or glycobiology, may be defined most simply as the study of the structure, synthesis, and biological function of all glycans (may be referred to as oligosaccharides or polysaccharides, depending on size) and glycoconjugates in simple and complex systems (Varkin et al. 1999; Fukuda and Hindsgaul 2000; Raghavachari 2012). The application of glycomics or glycobiology is sometimes called glycotecology to distinguish it from biotechnology

(referring to glycans rather than proteins and nucleic acids). However, many in the biotech arena consider glycobiology one of the research fields encompassed by the term biotechnology. In the postgenomic era, the intricacies of protein glycosylation, the mechanisms of genetic control, and the internal and external factors influencing the extent and patterns of glycosylation are important to understanding protein function and proteomics. Like proteins and nucleic acids, glycans are biopolymers. While once referred to as the last frontier of pharmaceutical discovery, recent advances in the biotechnology of discovering, cloning, and harnessing sugar cleaving and synthesizing enzymes have enabled glycobiologists to analyze and manipulate complex carbohydrates more easily (Walsh and Jefferis 2006).

Many of the proteins produced by animal cells contain attached sugar moieties, making them glycoproteins. The majority of protein-based medicinal agents contain some form of posttranslational modification that can profoundly affect the biological activity of that protein. Bacterial hosts for recombinant DNA could produce the animal proteins with identical or nearly identical amino acid sequences. However, early work in bacteria lacked the ability to attach sugar moieties to proteins (a process called glycosylation). New methodologies may help overcome this issue (cf. Chap. 3). Many of the non-glycosylated proteins differ in their biological activity as compared to the native glycoprotein. The production of animal proteins that lacked glycosylation provided an unexpected opportunity to study the functional role of sugar molecules on glycoproteins. There has been extensive progress in glycoengineering of yeast to humanize N-glycosylation pathways resulting in therapeutic glycoprotein expression in yeasts (Wildt and Gerngross 2005).

The complexity of the field can best be illustrated by reviewing the building blocks of glycans, the simple carbohydrates called saccharides or sugars and their derivatives (i.e., amino sugars). Simple carbohydrates can be attached to other types of biological molecules to form glycoconjugates including glycoproteins (predominantly protein), glycolipids and proteoglycans (about 95 % polysaccharide and 5 % protein). While carbohydrate chemistry and biology have been active areas of research for centuries, advances in biotechnology have provided techniques and added energy to the study of glycans. Oligosaccharides found conjugated to proteins (glycoproteins) and lipids (glycolipids) display a tremendous structural diversity. The linkages of the monomeric units in proteins and in nucleic acids are generally consistent in all such molecules. Glycans, however, exhibit far greater variability in the linkage between monomeric units than that found in the other

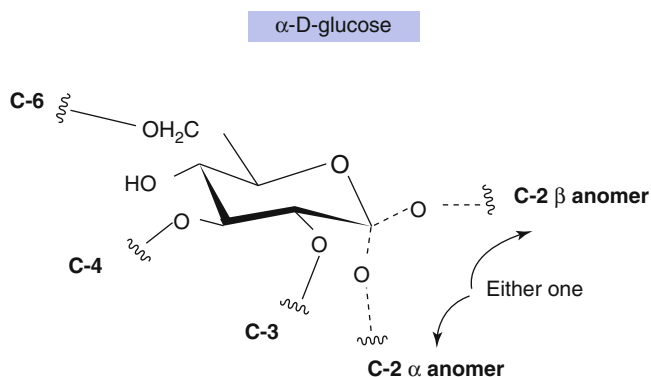


Figure 8.7 ■ Illustration of the common linkage sites to create biopolymers of glucose. Linkages at four positions: C-2, C-3, C-4, and C-6 and also can take one of two possible anomeric configurations at C-2 (α and β).

biopolymers. As an example, Fig. 8.7 illustrates the common linkage sites to create polymers of glucose. Glucose can be linked at four positions: C-2, C-3, C-4, and C-6 and also can take one of two possible anomeric configurations at C-2 (α and β). The effect of multiple linkage arrangements is seen in the estimate of (Kobata 1996). He has estimated that for a 10-mer (oligomer of length 10), the number of structurally distinct linear oligomers is DNA (with 4 possible bases), 1.04×10^6 ; protein (with 20 possible amino acids), 1.28×10^{13} ; and oligosaccharide (with eight monosaccharide types), 1.34×10^{18} .

Glycosylation and Medicine

Patterns of glycosylation significantly affect the biological activity of proteins (Wildt and Gerngross 2005; Walsh and Jefferis 2006). Many of the therapeutically used recombinant DNA-produced proteins are glycosylated including erythropoietin, glucocerebrosidase, and tissue plasminogen activator. Without the appropriate carbohydrates attached, none of these proteins will function therapeutically as does the parent glycoprotein. Glycoforms (variations of the glycosylation pattern of a glycoprotein) of the same protein may differ in physicochemical and biochemical properties. For example, erythropoietin has one O-linked and three N-linked glycosylation sites. The removal of the terminal sugars at each site destroys *in vivo* activity and removing all sugars results in a more rapid clearance of the molecule and a shorter circulatory half-life (Takeuchi et al. 1990). Yet, the opposite effect is observed for the deglycosylation of the hematopoietic cytokine granulocyte-macrophage colony-stimulating factor (GM-CSF) (Cebon et al. 1990). In that case, removing the carbohydrate residues increases the specific activity sixfold. The sugars of glycoproteins are known to play a role in the recognition and binding of

biomolecules to other molecules in disease states such as asthma, rheumatoid arthritis, cancer, HIV infection, the flu, and other infectious diseases.

■ Lipidomics

Lipids, the fundamental components of membranes, play multifaceted roles in cell, tissue, and organ physiology. The relatively new research area of lipidomics may be defined as the large-scale study of pathways and networks of cellular lipids in biological systems (Wenk 2005; Raghavachari 2012). The metabolome would include the major classes of biological molecules: proteins (and amino acids), nucleic acids, and carbohydrates. The “lipidome” would be a subset of the metabolome that describes the complete lipid profile within a cell, tissue, or whole organism. In lipidomic research, a vast amount of information (structures, functions, interactions, and dynamics) quantitatively describing alterations in the content and composition of different lipid molecular species is accrued after perturbation of a cell, tissue, or organism through changes in its physiological or pathological state. The study of lipidomics is important to a better understanding of many metabolic diseases, as lipids are believed to play a role in obesity, atherosclerosis, stroke, hypertension, and diabetes.

Lipid profiling is a targeted metabolomic platform that provides a comprehensive analysis of lipid species within a cell or possibly a tissue. The progress of modern lipidomics has been greatly accelerated by the development of sensitive analytical techniques such as electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI). Currently, the isolation and subsequent analysis of lipid mixtures is hampered by extraction and analytical limitations due to characteristics of lipid chemistry.

■ Nutrigenomics

The well-developed tools and techniques of genomics and bioinformatics have been applied to the examination of the intricate interplay of mammalian diet and genetic makeup. Nutrigenomics or nutritional genomics has been defined as the influence of genetic variation on nutrition. This appears to result from gene expression and/or gene variation (e.g., SNP analysis) on a nutrient’s absorption, distribution, metabolism, elimination, or biological effects (Laursen 2010). This includes how nutrients impact on the production and action of specific gene products and how the expressed proteins in turn affect the response to nutrients. Nutrigenomic studies aim to develop predictive means to optimize nutrition, with respect to an individual’s genotype. Areas of study include dietary supplements, common foods and beverages, mother’s milk, as well as diseases such as cardiovascular, obesity, and diabetes.

While still in its infancy, nutrigenomics is thought to be a critical science for personalized health and public health over the next decade (Kaput et al. 2007).

■ Other “Omic” Technologies

Pharmaceutical scientists and pharmacists may hear about other “omic” technologies in which the “omic” terms derive from the application of modern genomic techniques to the study of various biological properties and processes. For example, interactomics is the data-intensive broad system study of the interactome, which is the interaction among proteins and other molecules within a cell. Proteogenomics has been used as a broadly encompassing term to describe the merging of genomics, proteomics, small molecules, and informatics. Cellomics has been defined as the study of gene function and the proteins they encode in living cells utilizing light microscopy and especially digital imaging fluorescence microscopy.

■ “Omics” Integrating Technology: Systems Biology

The Human Genome Project and the development of bioinformatics technologies have catalyzed fundamental changes in the practice of modern biology and helped unveil a remarkable amount of information about many organisms (Aderem and Hood 2001; Price et al. 2010). Biology has become an information science defining all the elements in a complex biological system and placing them in a database for comparative interpretation. As seen in Fig. 8.2, the hierarchy of information collection goes well beyond the biodata contained in the genetic code that is transcribed and translated. Systems biology involves a generally complex interactive system. This research area is often described as a noncompetitive or precompetitive technology by the pharmaceutical industry because it is believed to be a foundational technology that must be better developed to be successful at the competitive technology of drug discovery and development. It is the study of the interactions between the components of a biological system and how these interactions give rise to the function and behavior of that system. Systems biology is essential for our understanding of how all the individual parts of intricate biological networks are organized and function in living cells. The biological system may involve enzymes and metabolites in a metabolic pathway or other interacting biological molecules affecting a biological process. Molecular biologists have spent the past 50+ years teasing apart cellular pathways down to the molecular level. Characterized by a cycle of theory, computational modeling, and experiment to quantitatively describe cells or cell processes, systems biology is a data-intensive endeavor that results in a conceptual framework for the analysis and understanding of complex

biological systems in varying contexts (Rothberg et al. 2005; Klipp et al. 2005; Meyer et al. 2011). Statistical mining, data alignment, probabilistic and mathematical modeling, and data visualization into networks are among the mathematical models employed to integrate the data and assemble the systems network (Gehlenborg et al. 2010). New measurements are stored with existing data, including extensive functional annotations, in molecular databases, and model assembly provides libraries of network models (Schrattenholz et al. 2010).

As the biological interaction networks are extremely complex, so are graphical representations of these networks. After years of research, a set of guidelines known as the Systems Biology Graphical Notation (SBGN) has been generally accepted to be the standard for graphical representation by all researchers. These standards are designed to facilitate the interchange of systems biology information and storage. Due to the complexity of these diagrams depending on the interactions examined and the level of understanding, a figure related to systems biology has not been included in this chapter. However, the reader is referred to the following website authored by the SBGN organization for several excellent examples of complex systems biology-derived protein interaction networks: <http://www.sbgn.org/Documents/Examples>

The inability to visualize the complexity of biological systems has in the past impeded the identification and validation of new and novel drug targets. The accepted SBGN standards should facilitate the efforts of pharmaceutical scientists to validate new and novel targets for drug design (Hood and Perlmutter 2004).

Since the objective is a model of all the interactions in a system, the experimental techniques that most suit systems biology are those that are system-wide and attempt to be as complete as possible. High-throughput “omic” technologies such as genomics, epigenomics, proteomics, pharmacogenomics, transcriptomics, metabolomics, and toxicogenomics are used to collect quantitative data for the construction and validation of systems models. Pharmaceutical and clinical end points include systems level biomarkers, genetic risk factors, aspects of personalized medicine, and drug target identification (Price et al. 2010; Yuryev 2012). In the future, application of systems biology approaches to drug discovery promises to have a profound impact on patient-centered medical practice, permitting a comprehensive evaluation of underlying predisposition to disease, disease diagnosis, and disease progression. Also, realization of personalized medicine and systems medicine will require new analytical approaches such as systems biology to decipher extraordinarily large, and extraordinarily noisy, data sets (Price et al. 2010).

TRANSGENIC ANIMALS AND PLANTS IN DRUG DISCOVERY, DEVELOPMENT, AND PRODUCTION

For thousands of years, man has selectively bred animals and plants either to enhance or to create desirable traits in numerous species. The explosive development of recombinant DNA technology and other molecular biology techniques have made it possible to engineer species possessing particular unique and distinguishing genetic characteristics. As described in Chap. 3, the genetic material of an animal or plant can be manipulated so that extra genes may be inserted (transgenes), replaced (i.e., human gene homologs coding for related human proteins), or deleted (knockout). Theoretically, these approaches enable the introduction of virtually any gene into any organism. A greater understanding of specific gene regulation and expression will contribute to important new discoveries made in relevant animal models. Such genetically altered species have found utility in a myriad of research and potential commercial applications including the generation of models of human disease, protein drug production, creation of organs and tissues for xenotransplantation, a host of agricultural uses, and drug discovery (Dunn et al. 2005; Clark and Pazdernik 2012a, b).

■ Transgenic Animals

As describe in Chap. 3, the term transgenic animal describes an animal in which a foreign DNA segment (a transgene) is incorporated into their genome. Later, the term was extended to also include animals in which their endogenous genomic DNA has had its molecular structure manipulated. While there are some similarities between transgenic technology and gene therapy, it is important to distinguish clearly between them. Technically speaking, the introduction of foreign DNA sequences into a living cell is called gene transfer. Thus, one method to create a transgenic animal involves gene transfer (transgene incorporated into the genome). Gene therapy (cf. Chap. 24) is also a gene transfer procedure and, in a sense, produces a transgenic human. In transgenic animals, however, the foreign gene is transferred indiscriminately into all cells, including germ line cells. The process of gene therapy differs generally from transgenesis since it involves a transfer of the desired gene in such a way that involves only specific somatic and hematopoietic cells, and not germ cells. Thus unlike in gene therapy, the genetic changes in transgenic organisms are conserved in any offspring according to the general rules of Mendelian inheritance.

The production of transgenic animals is not a new technology (Dunn et al. 2005; Clark and Pazdernik 2012b; Khan 2012a). They have been produced since

Year	Cloned animals
1996	Sheep (Dolly)
1997	Mouse
1998	Cattle
1999	Goat
2000	Pig
2001	Cat
2002	Rabbit
2003	Rat
2003	Mule
2003	Horse
2003	Deer
2005	Dog

Table 8.3 ■ Some cloned animals.

the 1970s. However, modern biotechnology has greatly improved the methods of inducing the genetic transformation. While the mouse has been the most studied animal species, transgenic technology has been applied to cattle, fish (especially zebra fish), goats, poultry, rabbits, rats, sheep, swine, cats, dogs, horses, mules, deer, and various lower animal forms such as mosquitoes (Table 8.3). Transgenic animals have already made valuable research contributions to studies involving regulation of gene expression, the function of the immune system, genetic diseases, viral diseases, cardiovascular disease, and the genes responsible for the development of cancer. Transgenic animals have proven to be indispensable in drug lead identification, lead optimization, preclinical drug development, and disease modeling.

Production of Transgenic Animals by DNA Microinjection and Random Gene Addition

The production of transgenic animals has most commonly involved the microinjection (also called gene transfer) of 100–200 copies of exogenous transgene DNA into the larger, more visible male pronucleus (as compared to the female pronucleus) of a recipient fertilized embryo (see Fig. 8.8) (Clark and Pazdernik 2012b; Khan 2012a). The transgene contains both the DNA encoding the desired target amino acid sequence along with regulatory sequences that will mediate the expression of the added gene. The microinjected eggs are then implanted into the reproductive tract of a female and allowed to develop into embryos. The foreign DNA generally becomes randomly inserted at a single site on just one of the host chromosomes (i.e., the

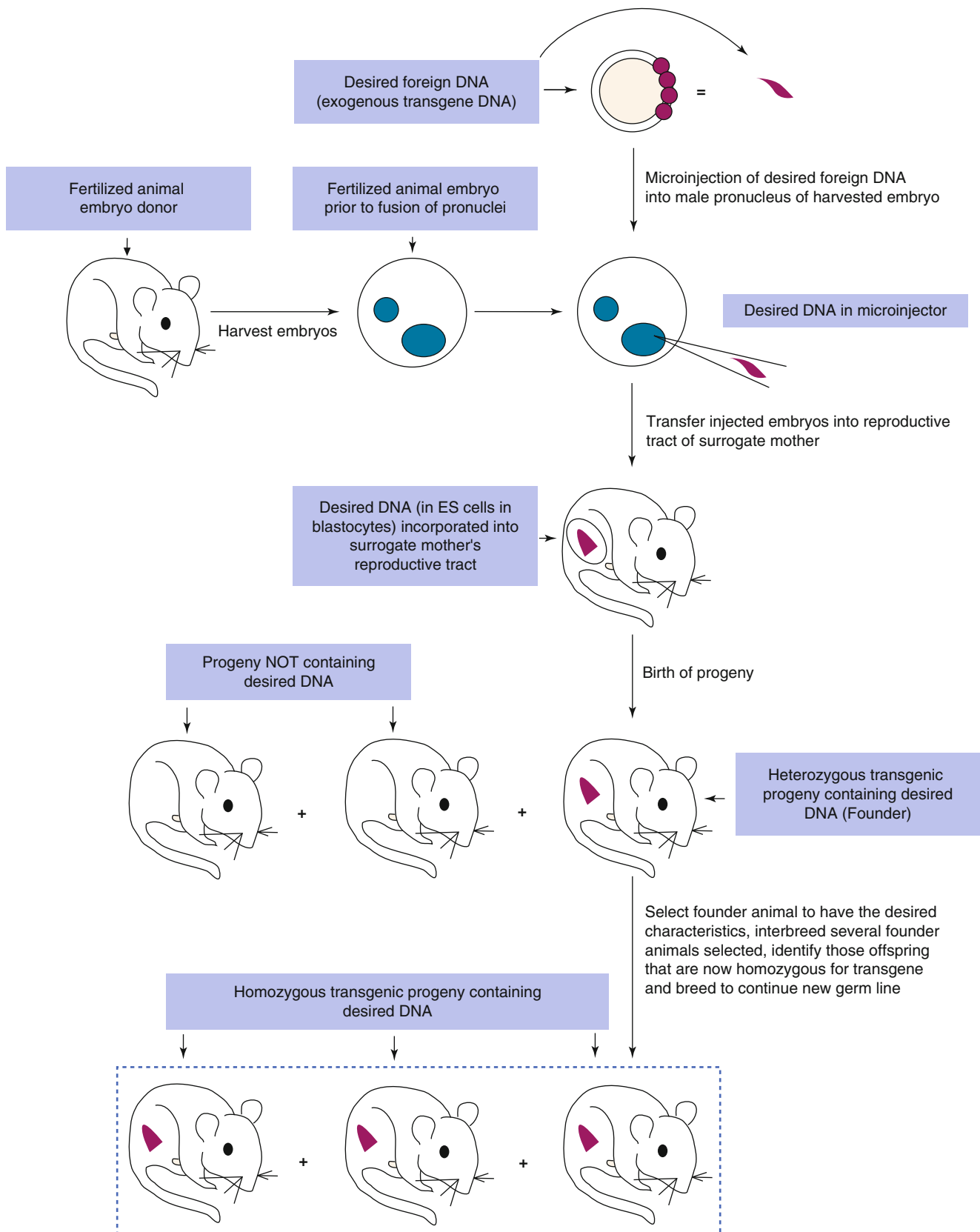


Figure 8.8 ■ Schematic representation of the production of transgenic animals by DNA microinjection.

founder transgenic animal is heterozygous). Thus, each transgenic founder animal (positive transgene incorporated animals) is a unique species. Interbreeding of founder transgenic animals where the transgene has been incorporated into germ cells may result in the birth of a homozygous progeny provided the transgene incorporation did not induce a mutation of an essential endogenous gene. All cells of the transgenic animal will contain the transgene if DNA insertion occurs prior to the first cell division. However, usually only 20–25 % of the offspring contain detectable levels of the transgene. Selection of neonatal animals possessing an incorporated transgene can readily be accomplished either by the direct identification of specific DNA or mRNA sequences or by the observation of gross phenotypic characteristics.

Production of Transgenic Animals by Retroviral Infection

The production of the first genetically altered laboratory mouse embryos was by insertion of a transgene via a modified retroviral vector (Clark and Pazdernik 2012b; Khan 2012a) (see Chap. 24 for more detailed description of retroviral vectors in gene therapy). The non-replicating viral vector binds to the embryonic host cells, allowing subsequent transfer and insertion of the transgene into the host genome. Many of the experimental human gene therapy trials employ the same viral vectors. Advantages of this method of transgene production are the ease with which genes can be introduced into embryos at various stages of development, and the characteristic that only a single copy of the transgene is usually integrated into the genome. Disadvantages include possible genetic recombination of the viral vector with other viruses present, the size limitation of the introduced DNA (up to 7 kb of DNA, less than the size of some genes), and the difficulty in preparing certain viral vectors.

Production of Transgenic Animals by Homologous Recombination in Embryonic Stem Cells Following Microinjection of DNA

Transgenic animals can also be produced by the *in vitro* genetic alteration of pluripotent embryonic stem cells (ES cells) (see Fig. 8.9; cf. Chap. 25) (Clark and Pazdernik 2012b; Khan 2012a). ES cell technology is more efficient at creating transgenics than microinjection protocols. ES cells, a cultured cell line derived from the inner cell mass (blastocyst) of a blastocyte (early preimplantation embryo), are capable of having their genomic DNA modified while retaining their ability to contribute to both somatic and germ cell lineages. The desired gene is incorporated into ES cells by one of several methods such as microinjection. This

is followed by introduction of the genetically modified ES cells into the blastocyst of an early preimplantation embryo, selection, and culturing of targeted ES cells which are transferred subsequently to the reproductive tract of the surrogate host animal. The resulting progeny is screened for evidence that the desired genetic modification is present and selected appropriately. In mice, the process results in approximately 30 % of the progeny containing tissue genetically derived from the incorporated ES cells. Interbreeding of selected founder animals can produce species homozygous for the mutation.

While transforming embryonic stem cells is more efficient than the microinjection technique described first, the desired gene must still be inserted into the cultured stem cell's genome to ultimately produce the transgenic animal. The gene insertion could occur in a random or in a targeted process. Nonhomologous recombination, a random process, readily occurs if the desired DNA is introduced into the ES cell genome by a gene recombination process that does not require any sequence homology between genomic DNA and the foreign DNA. While most ES cells fail to insert the foreign DNA, some do. Those that do are selected and injected into the inner cell mass of the animal blastocyst and thus eventually lead to a transgenic species. In still far fewer ES cells, homologous recombination occurs by chance. Segments of DNA base sequence in the vector find homologous sequences in the host genome, and the region between these homologous sequences replaces the matching region in the host DNA. A significant advance in the production of transgenic animals in ES cells is the advent of targeted homologous recombination techniques.

Homologous recombination, while much more rare to this point in transgenic research than nonhomologous recombination, can be favored when the researcher carefully designs (engineers) the transferred DNA to have specific sequence homology to the endogenous DNA at the desired integration site and also carefully selects the transfer vector conditions. This targeted homologous recombination at a precise chromosomal position provides an approach to very subtle genetic modification of an animal or can be used to produce knockout mice (to be discussed later).

A modification of the procedure involves the use of hematopoietic bone marrow stem cells rather than pluripotent embryonic stem cells. The use of ES cells results in changes to the whole germ line, while hematopoietic stem cells modified appropriately are expected to repopulate a specific somatic cell line or lines (more similar to gene therapy).

The science of cloning and the resulting ethical debate surrounding it is well beyond the scope of this

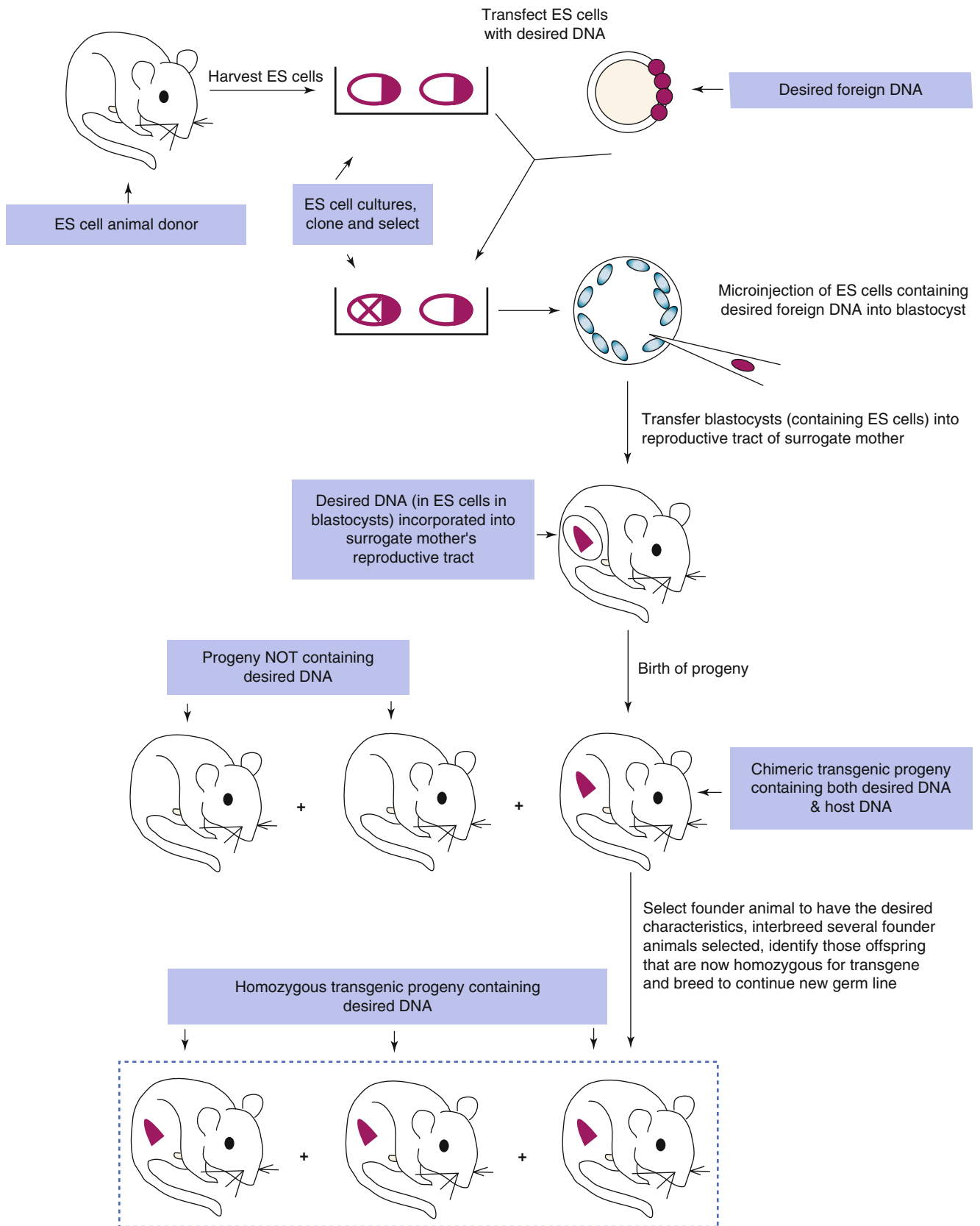


Figure 8.9 ■ Schematic representation of the production of transgenic animals by pluripotent embryonic stem cell methodology.

chapter. Yet it is important to place the concept of animal cloning within the pharmaceutically important context of transgenic animal production. The technique of microinjection (and its variations) has formed the basis for commercial transgenic animal production. While successful, the microinjection process is limited to the creation of only a small number of transgenic animals in a given birth. The slow process of conventional breeding of the resulting transgenic progeny must follow to produce a larger number of transgenic animals with the same transgene as the original organism. To generate a herd (or a flock, etc.), an alternative approach would be advantageous. The technique of nuclear transfer, the replacement of the nuclear genomic DNA of an oocyte (immature egg) or a single-cell fertilized embryo with that from a donor cell, is such an alternative breeding methodology. Animal “cloning” can result from this nuclear transfer technology. Judged the journal *Science’s* most important breakthrough of 1997 creating the sheep Dolly, the first cloned mammal, from a single cell of a 6-year old ewe was a feat many had thought impossible. Dolly was born after nuclear transfer of the genome from an adult mammary gland cell (Khan 2012a). Since this announcement, commercial and exploratory development of nuclear transfer technology has progressed rapidly with various species cloned. It is important to note that the cloned sheep Dolly was NOT a transgenic animal. While Dolly was a clone of an adult ewe, she did not possess a transgene. However, cloning could be used to breed clones of transgenic animals or to directly produce transgenic animals (if prior to nuclear transfer, a transgene was inserted into the genome of the cloning donor). For example, human factor IX (a blood factor protein) transgenic sheep was generated by nuclear transfer from transfected fetal fibroblasts (Schniecke et al. 1997). Several of the resulting progeny were shown to be transgenic (i.e., possessing the human factor IX gene), and one was named Polly. Thus, animal cloning can be utilized not only for breeding but also for the production of potential human therapeutic proteins and other useful pharmaceutical products (Shultz et al. 2007).

■ Transgenic Plants

A variety of biotechnology genetic engineering techniques have been employed to create a wealth of transgenic plant species as mentioned in Chap. 3: cotton, maize, soybean, potato, petunia, tobacco, papaya, rose, and others (Clark and Pazdernik 2012a; Khan 2012b). Agricultural enhancements have resulted by engineering plants to be more herbicide tolerant, insect resistant, fungus resistant, virus resistant, and stress tolerant (Baez 2005). Of importance for human health and pharmaceutical biotechnology, gene transfer technology is routinely used to manipulate bulk human protein production in a wide variety of transgenic plant species. It

is significant to note that transgenic plants are attractive bulk bioreactors because their posttranslational modification processes often result in plant-derived recombinant human proteins with greater glycosylation pattern similarity to that found in the corresponding native human proteins than would be observed in a corresponding mammalian production system. The transgenic seeds would result in seedlings capable of expressing the human protein. Transplantation to the field followed by normal growth, harvest of the biomass, and downstream isolation and protein purification results in a valuable alternative crop for farming. Tobacco fields producing pharmaceutical grade human antibodies (sometimes referred to as “plantibodies”) and edible vaccines contained in transgenic potatoes and tomatoes are not futuristic visions, but current research projects in academic and corporate laboratories. Farmers’ fields are not the sole sites for human protein production from flora. For example, cultured eukaryotic microalgae is also being developed into a useful expression system, especially for human and humanized antibodies (Mayfield and Franklin 2005). With antibody-based targeted therapeutics becoming increasingly important, the use of transgenic plants will likely continue to expand once research helps solve problems related to the isolation of the active protein drug and issues concerning cross fertilization with non-genetically modified organisms (non-GMOs).

Biopharmaceutical Protein Production in Transgenic Animals and Plants: “Biopharming”

The use of transgenic animals and plants as bioreactors for the production of pharmaceutically important proteins may become one of the most important uses of engineered species once numerous practical challenges are addressed (Baez 2005; Klimyuk et al. 2005). Table 8.4 provides a list of some selected examples of biopharmaceuticals from transgenic animals and plants. Utilizing conventional agronomic and farming techniques, transgenic animals and plants offer the opportunity to produce practically unlimited quantities of biopharmaceuticals.

The techniques to produce transgenic animals have been used to develop animal strains that secrete high levels of important proteins in various end organs such as milk, blood, urine, and other tissues. During such large animal “gene farming,” the transgenic animals serve as bioreactors to synthesize recoverable quantities of therapeutically useful proteins. Among the advantages of expressing protein in animal milk is that the protein is generally produced in sizable quantities and can be harvested manually or mechanically by simply milking the animal. Protein purification from the milk requires the usual separation techniques for proteins. In general, recombinant genes coding for the desired protein product are fused to the regulatory

Species	Protein product	Potential indication(s)
Cow	Collagen	Burns, bone fracture
Cow	Human fertility hormones	Infertility
Cow	Human serum albumin	Surgery, burns, shock, trauma
Cow	Lactoferrin	Bacterial GI infection
Goat	α -1-antiprotease inhibitor	Inherited deficiency
Goat	α -1-antitrypsin	Anti-inflammatory
Goat	Antithrombin III (ATryn)	Associated complications from genetic or acquired deficiency
Goat	Growth hormone	Pituitary dwarfism
Goat	Human fertility hormones	Infertility
Goat	Human serum albumin	Surgery, burns, shock, trauma
Goat	LAtPA2	Venous status ulcers
Goat	Monoclonal antibodies	Colon cancer
Goat	tPA2	Myocardial infarct, pulmonary embolism
Pig	Factor IX	Hemophilia
Pig	Factor VIII	Hemophilia
Pig	Fibrinogen	Burns, surgery
Pig	Human hemoglobin	Blood replacement for transfusion
Pig	Protein C	Deficiency, adjunct to tPA
Rabbit	Insulin-like growth factor	Wound healing
Rabbit	Interleukin-2	Renal cell carcinoma
Rabbit	Protein C	Deficiency, adjunct to tPA
Sheep	α -1-antitrypsin	Anti-inflammatory
Sheep	Factor VIII	Hemophilia
Sheep	Factor IX	Hemophilia
Sheep	Fibrinogen	Burns, surgery
Sheep	Protein C	Deficiency, adjunct to tPA
Tobacco	IgG	Systemic therapy (rabies virus, hepatitis B virus)
Tobacco	TGF- β 2	Ovarian cancer
Tobacco	Vitronectin	Protease
Tobacco	RhinoR	Fusion of human adhesion protein and human IgA for common cold
Safflower	Insulin	Diabetes
Corn	Meripase	Cystic fibrosis
Duckweed	Lacteron	Controlled release of α -interferon for hepatitis B and C
Potato	Poultry vaccine	Avian influenza (H5N1)

Data from references noted in this section of Chap. 8

Abbreviations: tPA tissue plasminogen activator, LAtPA long acting tissue plasminogen activator, TGF- β 3 tissue growth factor-beta

Table 8.4 ■ Some examples of human proteins under development in transgenic animals and plants.

sequences of the animal's milk-producing genes. The animals are not endangered by the insertion of the recombinant gene. The logical fusion of the protein product gene to the milk-producing gene targets the

transcription and translation of the protein product exclusively in mammary tissues normally involved in milk production and does not permit gene activation in other, non-milk-producing tissues in the animal.

Transgenic strains are established and perpetuated by breeding the animals since the progeny of the original transgenic animal (founder animal) usually also produce the desired recombinant protein.

Yields of protein pharmaceuticals produced transgenically are expected to be 10–100 times greater than those achieved in recombinant cell culture. Protein yields from transgenic animals are generally good [conservative estimates of 1 g/l (g/L) with a 30 % purification efficiency] with milk yield from various species per annum estimated at: cow=10,000 L; sheep=500 L; goat=400 L; and pig=250 L (Rudolph 1995). PPL Therapeutics has estimated that the cost to produce human therapeutic proteins in large animal bioreactors could be as much as 75 % less expensive than cell culture. In addition, should the desired target protein require posttranslational modification, the large mammals used in milk production of pharmaceuticals would be a bioreactor capable of adding those groups (unlike a recombinant bacterial culture).

Some examples of human long peptides and proteins under development in the milk of transgenic animals include growth hormone, interleukin-2, calcitonin, insulin-like growth factor, alpha 1 antitrypsin (AT), clotting factor VIII, clotting factor IX, tissue plasminogen activator (tPA), lactoferrin, gastric lipase, vaccine derived from *Escherichia coli* LtB toxin subunit, protein C, and various human monoclonal antibodies (such as those from the Xenomouse) (Garner and Colman 1998; Rudolph 2000). The first approved biopharmaceutical from transgenic animals is recombinant human antithrombin which remains on the market today (trade name ATryn). Produced in a herd of transgenic dairy goats, rhAT is expressed in high level in the milk. The human AT transgene was assembled by linking the AT cDNA to a normal milk protein sequence (XhoI site of the goat beta casein vector). See Table 8.4 for additional examples.

Using genetic engineering techniques to create transgenic plants, “pharming” for pharmaceuticals is producing an ever-expanding list of drugs and diagnostic agents derived from human genes (Baez 2005; Opar 2011). Some examples of human peptides and proteins under development in transgenic plants include TGF-beta, vitronectin, thyroid-stimulating hormone receptor, insulin, glucocerebrosidase, apolipoprotein A-1, and taliglucerase alfa. See Table 8.4 for additional examples.

■ Xenotransplantation: Transplantable Transgenic Animal Organs

An innovative use of transgenics for the production of useful proteins is the generation of clinically transplantable transgenic animal organs, the controversial cross-species transplant (Khan 2012c). The success of human-to-human transplantation of heart, kidney,

liver, and other vascularized organs (allotransplantation) created the significant expectation and need for donor organs. Primate-to-human transplantation (xenotransplantation) was successful, but ethical issues and limited number of donor animals were significant barriers. Transplant surgeons recognized early on that organs from the pig were a rational choice for xenotransplantation (due to physiological, anatomical, ethical, and supply reasons) if the serious hyperacute rejection could be overcome. Several research groups in academia and industry have pioneered the transgenic engineering of pigs expressing both human complement inhibitory proteins and key human blood group proteins (antigens) (McCurry et al. 1995; Dunn et al. 2005; Van Eyck et al. 2010). Cloning has now produced transgenic pigs for xenotransplantation. Cells, tissues, and organs from these double transgenic animals appear to be very resistant to the humoral immune system-mediated reactions of both primates and likely humans. These findings begin to pave the way for potential xenograft transplantation of animal components into humans with a lessened chance of acute rejection. A continuing concern is that many animals, such as pigs, have shorter life spans than humans, meaning that their tissues age at a quicker rate.

■ Knockout Mice

While many species including mice, zebra fish, and nematodes have been transformed to lose genetic function for the study of drug discovery and disease modeling, mice have proven to be the most useful. Mice are the laboratory animal species most closely related to humans in which the knockout technique can be easily performed, so they are a favorite subject for knockout experiments. While a mouse carrying an introduced transgene is called a transgenic mouse, transgenic technologies can also produce a knockout animal. A knockout mouse, also called a gene knockout mouse or a gene-targeted knockout mouse, is an animal in which an endogenous gene (genomic wild-type allele) has been specifically inactivated by replacing it with a null allele (Sharpless and DePinho 2006; Wu et al. 2011b; Clark and Pazdernik 2012b). A null allele is a nonfunctional allele of a gene generated by either deletion of the entire gene or mutation of the gene resulting in the synthesis of an inactive protein. Recent advances in intranuclear gene targeting and embryonic stem cell technologies as described above are expanding the capabilities to produce knockout mice routinely for studying certain human genetic diseases or elucidating the function of a specific gene product.

The procedure for producing knockout mice basically involves a four-step process. A null allele (i.e., knockout allele) is incorporated into one allele of murine ES cells. Incorporation is generally quite low;

approximately one cell in a million has the required gene replacement. However, the process is designed to impart neomycin and ganciclovir resistance only to those ES cells in which homologous gene integration has resulted. This facilitates the selection and propagation of the correctly engineered ES cells. The resulting ES cells are then injected into early mouse embryos creating chimeric mice (heterozygous for the knockout allele) containing tissues derived from both host cells and ES cells. The chimeric mice are mated to confirm that the null allele is incorporated into the germ line. The confirmed heterozygous chimeric mice are bred to homozygosity producing progeny that are homozygous knockout mice. Worldwide, three major mouse knockout programs are proceeding in collaboration to create a mutation in each of the approximately 20,000 protein-coding genes in the mouse genome using a combination of gene trapping and gene targeting in mouse embryonic stem (ES) cells (Staff 2007). These include (1) KOMP (KnockOut Mouse Project, <http://www.knockoutmouse.org>), funded by the NIH; (2) EUCOMM (EUropean Conditional Mouse Mutagenesis Program, <http://www.eucomm.org>), funded by the FP6 program of the EC; and (3) NorCOMM (North American Conditional Mouse Mutagenesis Project, <http://norcomm.phenogenomics.ca/index.htm>), a Canadian project funded by Genome Canada and partners. To date, over 4,000 targeted knockouts of genes have been accomplished. This comprehensive and publicly available resource will aid researchers examining the role of each gene in normal physiology and development and shed light on the pathogenesis of abnormal physiology and disease.

The continuing discoveries of the three worldwide mouse knockout consortia and independent research laboratories around the world will further create better models of human monogenic and polygenic diseases such as cancer, diabetes, obesity, cardiovascular disease, and psychiatric and neurodegenerative diseases. For example, knockout mice have been engineered that have extremely elevated cholesterol levels while being maintained on normal chow diets due to their inability to produce apolipoprotein E (apoprotein E). Apoprotein E is the major lipoprotein component of very low-density lipoprotein (VLDL) responsible for liver clearance of VLDL. These engineered mice are being examined as animal models of atherosclerosis useful in cardiovascular drug discovery and development. Table 8.5 provides a list of some additional selected examples of knockout mouse disease models.

The knockout mouse is becoming the basic tool for researchers to determine gene function in vivo in numerous biological systems. For example, knockout mouse technology has helped transform our understanding of the immune response. The study of single and multiple gene knockout animals has provided new perspectives

Genetic engineering	Gene	Disease model
Knockout	<i>BRCA1</i> , <i>BRCA2</i>	Breast cancer
Knockout	Apolipoprotein E	Atherosclerosis
Knockout	Glucocerebrosidase	Gaucher's disease
Knockout	HPRT	Lesch-Nyhan syndrome
Knockout	Hexokinase A	Tay-Sachs disease
Knockout	Human CFTR	Cystic fibrosis
Knockout	p53	Cancer suppressor gene deletion
Knockout	P-glycoprotein	Multidrug resistance (MDR)
Knockout	α -globin and β -globin	Sickle cell anemia
Knockout	Urate oxidase	Gout
Knockout	Retinoblastoma-1	Familial retinoblastoma
Transgene	<i>c-neu</i> oncogene	Cancer
Transgene	<i>c-myc</i> oncogene	Cancer
Transgene	growth hormone	Dwarfism
Transgene	<i>H-ras</i> oncogene	Cancer
Transgene	Histocompatibility antigens	Autoimmunity
Transgene	HIV <i>tat</i>	Kaposi sarcoma
Transgene	Human APP	Alzheimer's disease
Transgene	Human β -globin	Thalassemia
Transgene	Human CD4 expression	HIV infection
Transgene	Human β -globin mutant	Sickle cell anemia
Transgene	Human CETP	Atherosclerosis
Transgene	LDL receptor	Hypercholesterolemia

Abbreviations: *BRCA1*, *BRCA2* suspected breast cancer genes, *CETP* cholesterol (cholesteryl) ester transfer protein, *CFTR* cystic fibrosis transport regulator, *HIV* human immunodeficiency virus, *HPRT* hypoxanthine phosphoribosyl transferase, *LDL* low-density lipoprotein

Table 8.5 ■ Some selected examples of genetically engineered animal disease models.

on T-cell development, costimulation, and activation. "Humanized mice," transgenic severe combined immunodeficiency (SCID) mice grafted with human cells and tissues, enable research in regenerative medicine, infectious disease, cancer, and human hematopoiesis. In addition, high-throughput DNA sequencing efforts, positional cloning programs, and novel embryonic stem cell-based gene discovery research areas all exploit the knockout mouse as their laboratory.

Engineered animal models are proving invaluable to pharmaceutical research since small animal models of disease may be created and validated to mimic a disease in human patients. Mouse, rat, and zebra fish are the most common models explored and used. Genetic engineering can predispose an animal to a particular disease under scrutiny, and the insertion of human genes into the animal can initiate the development of a more clinically relevant disease condition. In human clinical studies, assessments of efficacy and safety often rely on measured effects for surrogate biomarkers and adverse event reporting. Validated transgenic animal models of human disease allow for parallel study and possible predictability prior to entering clinical trials. Also, it is possible to screen potential drug candidates *in vivo* against a human receptor target inserted into an animal model. The number of examples of transgenic animal models of human disease useful in drug discovery and development efforts is growing rapidly (Sharpless and DePinho 2006; Schultz et al. 2007; Wu et al. 2011b; Clark and Pazdernik 2012b). Such models have the potential to increase the efficiency and decrease the cost of drug discovery and development by reducing the time it

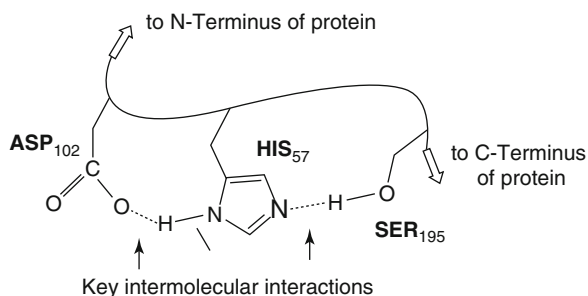
takes to move a candidate medicinal agent from discovery into clinical trials. Table 8.5 provides a list of some selected examples of genetically engineered animal models of human disease.

SITE-DIRECTED MUTAGENESIS

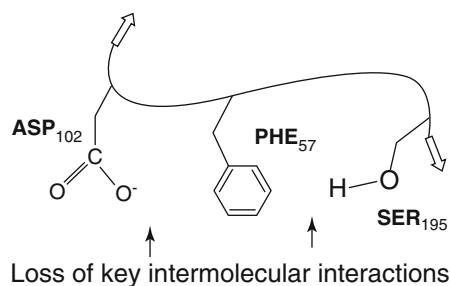
Site-directed mutagenesis, also called site-specific mutagenesis or oligonucleotide-directed mutagenesis, is a protein engineering technique allowing specific amino acid residue (site-directed) alteration (mutation) to create new protein entities (Johnson and Reitz 1998). Mutagenesis at a single amino acid position in an engineered protein is called a point mutation. Therefore, site-directed mutagenesis techniques can aid in the examination at the molecular level of the relationship between 3D structure and function of interesting proteins. The technique is commonly used in protein engineering. This technique resulted in a Nobel Prize for one of the early researchers in this field, Dr. Michael Smith (Hutchison et al. 1978).

Figure 8.10 suggests an excellent example of possible theoretical mutations of the active site of a model serine protease enzyme that could be engineered to

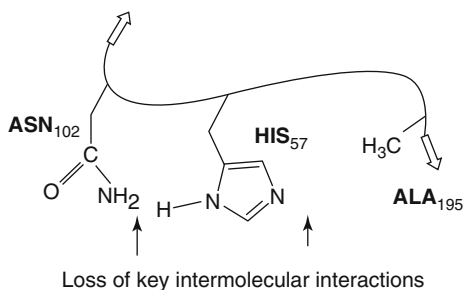
a. Catalytic triad (the catalytic machinery) at active site of a wild type, parent serine protease



b. Theoretical site-directed mutagenesis studies HIS₅₇ to PHE₅₇ mutant



c. Theoretical site-directed mutagenesis studies ASP₁₀₂ to ASN₁₀₂ and SER₁₉₅ to ALA₁₉₅ mutant



d. ASP₁₀₂ to ASN₁₀₂ mutant from site-directed mutagenesis (from Craik et al. 1987)

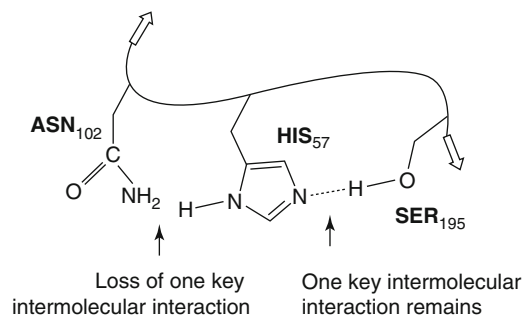


Figure 8.10 ■ Some possible site-directed mutations of the amino acids composing the catalytic triad of a serine protease: influence on key hydrogen bonding. (a) Catalytic triad (the catalytic machinery) at active site of a wild type, parent serine protease. (b) Theoretical site-directed mutagenesis studies HIS₅₇ to PHE₅₇ mutant. (c) Theoretical site-directed mutagenesis studies ASP₁₀₂ to ASN₁₀₂ and SER₁₉₅ to ALA₁₉₅ mutant. (d) ASP₁₀₂ to ASN₁₀₂ mutant from site-directed mutagenesis (From Craik et al. 1987).

probe the mechanism of action of the enzyme. Structures B and C of Fig. 8.10 represent a theoretical mutation to illustrate the technique. Craik and coworkers have actually tested the role of the aspartic acid residue in the serine protease catalytic triad Asp, His, and Ser. They replaced Asp¹⁰² (carboxylate anion side chain) of trypsin with Asn (neutral amide side chain) by site-directed mutagenesis and observed a pH-dependent change in the catalytic activity compared to the wild-type parent serine protease (see Fig. 8.10, structure D) (Craik et al. 1987). Site-directed mutagenesis studies also provide invaluable insight into the nature of intermolecular interactions of ligands with their receptors. For example, studies of the effect of the site-directed mutagenesis of various key amino acid residues on the binding of neurotransmitters to G-protein-coupled receptors have helped in defining more accurate models for alpha-adrenergic, D2-dopaminergic, 5HT2a-serotonergic, and both M1 and M3 muscarinic receptors (Bikker et al. 1998).

SYNTHETIC BIOLOGY

Modern biotechnology tools have allowed for a number of ways to study very complex biological systems. For example, as described above systems biology examines complex biological systems as interacting and integrated complex networks. The new and developing field of study known as synthetic biology explores how to build artificial complex biological systems employing many of the same tools and experimental techniques favored by system biologists. Synthetic biology looks at both the strategic redesign and/or the fabrication of existing biological systems and the design and construction of biological components and systems that do not already exist in nature (Khalil and Collins 2010). Synthetic genomics is a subset of synthetic biology that focuses on the redesign and fabrication of new genetic material constructed from raw chemicals. The focus of synthetic biology is often on ways of taking parts of natural biological systems, characterizing and simplifying them, and using them as a component of a highly unnatural, engineered, biological system. Synthetic biology studies may provide a more detailed understanding of complex biological systems down to the molecular level. Being able to design and construct a complex system is also one very practical approach to understanding that system under various conditions. The levels a synthetic biologist may work at include the organism, tissue and organ, intercellular, intracellular, biological pathway, and down to the molecular level.

There are many exciting applications for synthetic biology that have been explored or hypothesized across various fields of scientific study including designed

and optimized biological pathways, natural product manufacturing, new drug molecule synthesis, and biosensing (Ruder et al. 2011). From an engineering perspective, synthetic biology could lead to the design and building of engineered biological systems that process information, modify existing chemicals, fabricate new molecules and materials, and maintain and enhance human health and our environment. Because of the obvious societal concerns that synthetic biology experiments raise, the broader science community has engaged in considerable efforts at developing guidelines and regulations and addressing the issues of intellectual property and governance and the ethical, societal, and legal implications. Several bioethics research institutes published reports on ethical concerns and the public perception of synthetic biology. A report from the United States Presidential Commission for the Study of Bioethical Issues called for enhanced federal oversight on this emerging technology.

In 2006, a research team, at the J. Craig Venter Institute constructed and filed a patent application for a synthetic genome of a novel synthetic minimal bacterium named *Mycoplasma laboratorium* (Smith et al. 2003; Glass et al. 2007). The team was able to construct an artificial chromosome of 381 genes, and the DNA sequence they have pieced together is based upon the bacterium *Mycoplasma genitalium*. The original bacterium had a fifth of its DNA removed and was able to live successfully with the synthetic chromosome in place. Venter's goal is to make cells that might take carbon dioxide out of the atmosphere and produce methane, used as a feedstock for other fuels.

Chang and coworkers published pioneering work utilizing a synthetic biology approach to assemble two heterologous pathways for the biosynthesis of plant-derived terpenoid natural products (Chang 2007). Terpenoids are a highly diverse class of lipophilic natural products that have historically provided a rich source for discovery of pharmacologically active small molecules, such as the anticancer agent paclitaxel (Taxol) and the antimalarial artemisinin. Unfortunately, these secondary metabolites are typically produced in low abundance in their host organism, and their isolation consequently suffers from low yields and high consumption of natural resources. A key step is developing methods to carry out cytochrome P450 (P450)-based oxidation chemistry in vivo. Their work suggests that potentially, entire metabolic pathways can be designed in silico and constructed in bacterial hosts.

BIOTECHNOLOGY AND DRUG DISCOVERY

Pharmaceutical scientists have taken advantage of every opportunity or technique available to aid in the long, costly, and unpredictable drug discovery process.

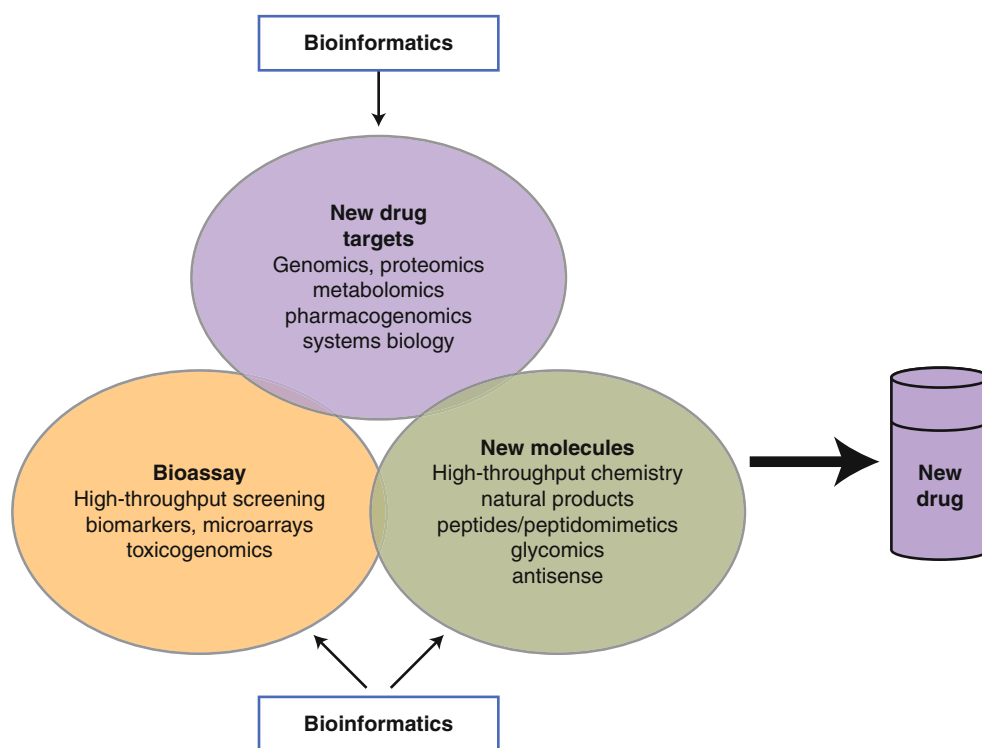


Figure 8.11 ■ Elements of modern drug discovery: impact of biotechnology.

In essence, Chap. 8 is an overview of some of the many applications of biotechnology and related techniques useful in drug discovery or design, lead optimization, and development. In addition to recombinant DNA and hybridoma technology, the techniques described throughout Chap. 8 have changed the way drug research is conducted, refining the process that optimizes the useful pharmacological properties of an identified novel chemical lead and minimizing the unwanted properties. The promise of genomics, proteomics, metabolomics, pharmacogenomics/pharmacogenetics, epigenomics, toxicogenomics, systems biology, and bioinformatics to radically change the drug discovery paradigm is eagerly anticipated (Beeley et al. 2000; Basu and Oyelere 2003; Loging et al. 2007). Figure 8.11 shows schematically the interaction of three key elements that are essential for modern drug discovery: new targets identified by genomics, proteomics, and related technologies; validation of the identified targets; rapid, sensitive bioassays utilizing high-throughput screening methods; and new molecule creation and optimization employing a host of approaches. The key elements are underpinned at each point by bioinformatics. Several of the technologies, methods, and approaches listed in Fig. 8.11 have been described previously in this chapter. Others will be described below.

■ Screening and Synthesis

Traditionally, drug discovery programs relied heavily upon random screening followed by analog synthesis

and lead optimization via structure-activity relationship studies. Discovery of novel, efficacious, and safer small molecule medicinal agents with appropriate “drug-like characteristics is an increasingly costly and complex process” (Williams 2007). Therefore, any method allowing for a reduction in time and money is extremely valuable. Advances in biotechnology have contributed to a greater understanding of the cause and progression of disease and have identified new therapeutic targets forming the basis of novel drug screens. New technical discoveries in the fields of proteomics for target discovery and validation and systems biology are expected to facilitate the discovery of new agents with novel mechanisms of action for diseases that were previously difficult or impossible to treat. In an effort to decrease the cost of identifying and optimizing useful, quality drug leads against a pharmaceutically important target, researchers have developed newer approaches including high-throughput screening and high-throughput synthesis methods.

Advances in Screening: High-Throughput Screening (HTS)

Recombinant DNA technology has provided the ability to clone, express, isolate, and purify receptor enzymes, membrane-bound proteins, and other binding proteins in larger quantities than ever before. Instead of using receptors present in animal tissues or partially purified enzymes for screening, in vitro bioassays now utilize the exact human protein target. Applications of biotechnology to in vitro screening include the improved

preparation of (1) cloned membrane-bound receptors expressed in cell lines carrying few endogenous receptors; (2) immobilized preparations of receptors, antibodies, and other ligand-binding proteins; and (3) soluble enzymes and extracellular cell-surface expressed protein receptors. In most cases today, biotechnology contributes directly to the understanding, identification, and/or the generation of the drug target being screened (e.g., radioligand binding displacement from a cloned protein receptor).

Previously, libraries of synthetic compounds along with natural products from microbial fermentation, plant extracts, marine organisms, and invertebrates provide a diversity of molecular structures that were screened randomly. Screening can be made more directed if the compounds to be investigated are selected on the basis of structural information about the receptor or natural ligand. The development of sensitive radioligand binding assays and the access to fully automated, robotic screening techniques have accelerated the screening process.

High-throughput screening (HTS) provides for the bioassay of thousands of compounds in multiple assays at the same time (Cik and Jurzak 2007; Rankovic et al. 2010). The process is automated with robots and utilizes multi-well microtiter plates. While 96-well microtiter plates are a versatile standard in HTS, the development of 1,536- and 3,456-well nanoplate formats and enhanced robotics brings greater miniaturization and speed to cell-based and biochemical assays. Now, companies can conduct 100,000 bioassays a day. In addition, modern drug discovery and lead optimization with DNA microarrays allows researchers to track hundreds to thousands of genes.

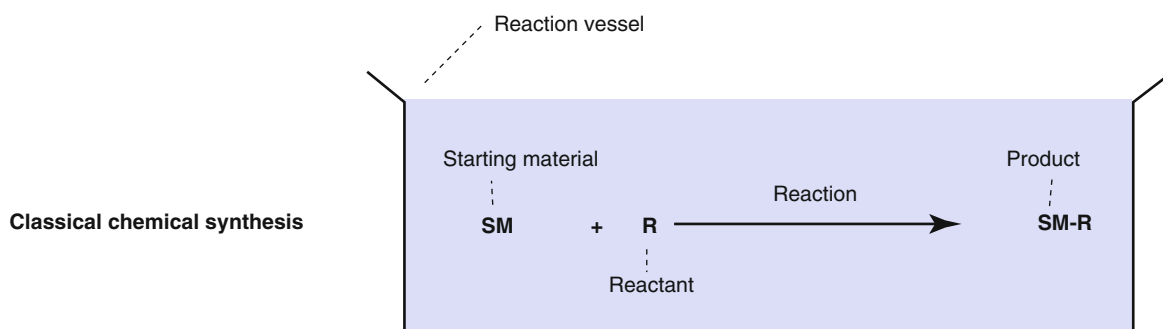
Enzyme inhibition assays and radioligand binding assays are the most common biochemical tests employed. The technology has become so sophisticated and the interactive nature of biochemical events so much better understood (through approaches such as systems biology) that HT whole cell assays have become commonplace. Reporter gene assays are routinely utilized in HTS (Ullmann 2007). Typically, a reporter gene, that is a reporter that indicates the presence or absence of a particular gene product that in turn reflects the changes in a biological process or pathway, is transfected into a desired cell. When the gene product is expressed in the living cell, the reporter gene is transcribed and the reporter is translated to yield a protein that is measured biochemically. A common reporter gene codes for the enzyme luciferase, and the intensity of the resulting green fluorescent protein (i.e., a quantitative measure of concentration) is a direct function of the assayed molecule's ability to stimulate or inhibit the biologic process or signaling pathway under study. A further advance in HT screening technologies for lead optimization is rapid, high-content pharmacology.

This HT screening approach can be used to evaluate solubility, adsorption, toxicity, metabolism, and other important drug characteristics.

High-Throughput Chemistry: Combinatorial Chemistry and Multiple Parallel Synthesis

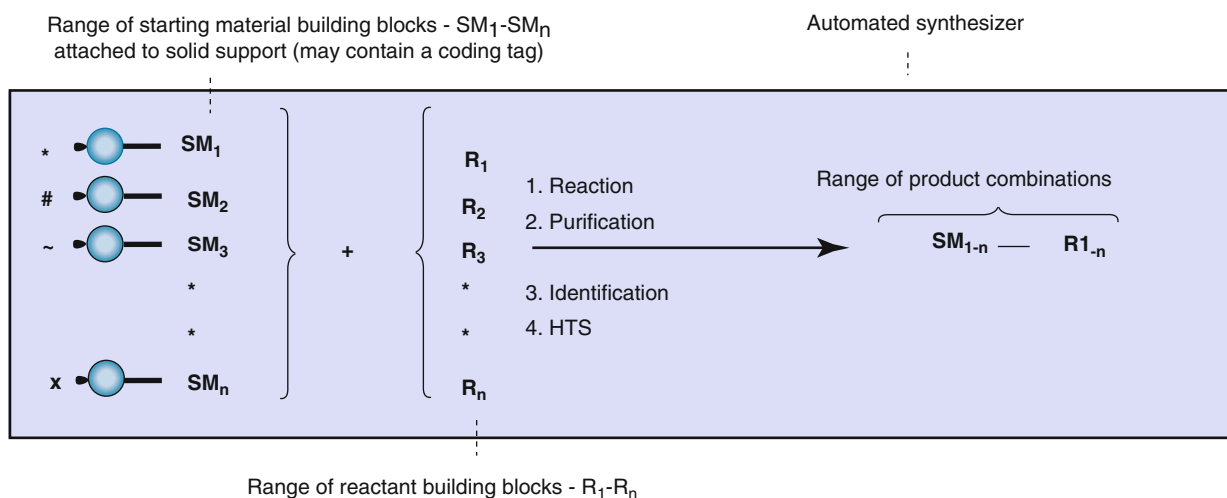
Traditionally, small drug molecules were synthesized by joining together structural pieces in a set sequence to prepare one product. One of the most powerful tools to optimize drug discovery is automated high-throughput synthesis. When conducted in a combinatorial approach, high-throughput synthesis provides for the simultaneous preparation of hundreds or thousands of related drug candidates (Fenniri 2000; Sucholeiki 2001; Mason and Pickett 2003; Pirrung 2004). The molecular libraries generated are screened in high-throughput screening assays for the desired activity, and the most active molecules are identified and isolated for further development.

There are two overall approaches to high-throughput synthesis: combinatorial chemistry that randomly mixes various reagents (such as many variations of reagent A with many variations of reagent B to give random mixtures of all products in a reaction vessel) and parallel synthesis that selectively conducts many reactions parallel to each other (such as many variations of reagent A in separate multiple reaction vessels with many variations of reagent B to give many single products in separate vessels) (Mitscher and Dutta 2003; Seneci 2007; Ashton and Maloney 2007). True combinatorial chemistry or sometimes referred to as combiChem, applies methods to substantially reduce the number of synthetic operations or steps needed to synthesize large numbers of compounds. CombiChem is conducted on solid supports (resins) to facilitate the manipulations required to reduce labor during purification of multiple products in the same vessel. Differing from combinatorial chemistry, multiple parallel synthesis procedures apply automation to the synthetic process to address the many separate reaction vessels needed, but the number of operations needed to carry out a synthesis is practically the same as the conventional approach. Thus, the potential productivity of multiple parallel methods is not as high as combinatorial chemistries. Parallel chemistries can be conducted on solid-phase supports or in solution to facilitate purification. Figure 8.12 provides an illustration of a combinatorial mix-and-match process in which a simple building block (a starting material such as an amino acid, peptide, heterocycle, other small molecule, etc.) is joined to one or more other simple building blocks in every possible combination. Assigning the task to automated synthesizing equipment results in the rapid creation of large collections or libraries (as large as 10,000 compounds) of diverse molecules. Ingenious methods have been devised to direct the molecules to be



In classical chemical synthesis, a coupling reaction of one starting material (SM) with one reactant (R) would yield just one product, SM-R. One or several reactions may be run simultaneously in separate reaction vessels

Combinatorial chemical synthesis: Example = Coupling reaction



In a combinatorial chemical synthesis such as a coupling reaction, a range of starting material building blocks are reacted with a range of reactant building blocks yielding any or all possible product combinations, SM_{1-n} R_{1-n}. The automated reactions may occur in the same reaction vessel (and coding tag used to separate/identify) or may each occur in small, separate reaction vessels (parallel synthesis)

Figure 8.12 ■ A schematic representation of a coupling reaction: difference between classical chemical synthesis and combinatorial chemistry.

synthesized, to identify the structure of the products, to purify the products via automation, and to isolate compounds. When coupled with high-throughput screening, thousands of compounds can be generated, screened, and evaluated for further development in a matter of weeks.

Building blocks include amino acids, peptides, nucleotides, carbohydrates, lipids, and a diversity of small molecule scaffolds or templates (Mason and Pickett 2003). A selection of reaction types used in combinatorial chemistry to produce compound libraries is found in Table 8.6.

■ Chemical Genomics

Chemical genomics is an emerging “omic” technology not discussed above. The development of high-throughput screening and high-throughput chemistry coupled with “omic” technologies has changed the drug discovery paradigm and the approach for the investigation of target pharmacology (Kubinyi and Müller 2004; Kubinyi 2007; Flaumenhaft 2007; Rankovic et al. 2010) (see Fig. 8.11). In modern drug discovery, chemical genomics (sometimes called chemogenomics or more generally included as a subset of chemical biology) involves the screening of large chemical

Compound types
α,β -Unsaturated ketones
α -Hydroxy acids
Acyl piperidines
Azoles
β -Mercaptoketones
β -Turn mimetics
Benzisothiazolones
Benzodiazepines
Biaryls
Cyclopentenones
Dihydropyridines
γ -Butyropyridines
Glycosylamines
Hydantoins
Isoxazoles
Modified oligonucleotides
Peptoids
Piperazinediones
Porphyrins
Porphyrins
1,3-Propanediols
Protease inhibitors
Pyrrolidines
Sulfamoylbenzamides
Tetrahydrofurans
Thiazole
Thiazolidinones

Table 8.6 ■ A sample of the diversity of compounds capable of being synthesized by combinatorial chemistry methods.

libraries (typically combinatorially derived “druggable” small molecule libraries covering a broad expanse of “diversity space”) against all genes or gene products, such as proteins or other targets (i.e., chemical universe screened against target universe). In Fig. 8.11, basically chemical genomics would occur when the “new molecules” to be tested in the “bioassay” developed from the “new drug targets” “omic” approaches came from large chemical libraries typically of small molecules synthesized by high-throughput chemistry. As part of the National Institutes of Health (NIH) Roadmap for Biomedical Research, the National Human Genome Research Institute (NHGRI) will lead an effort to offer public sector biomedical researchers access to libraries of small organic molecules that can be used as chemical probes to study cellular pathways in greater depth. It remains difficult to predict which small molecule compounds will be most effective in a given situation. Researchers can maximize the likelihood of a successful match between a chemical compound, its usefulness as a research tool, and its desired therapeutic effect by systematically screening libraries containing thousands of small molecules. Drug

candidates are expected from the correlations observed during functional analysis of the molecule – gene product interactions. Genomic profiling by the chemical library may also yield relevant new targets and mechanisms. Chemical genomics is expected to be a critical component of drug lead identification and proof of principle determination for selective modulators of complex enzyme systems including proteases, kinases, G-protein-coupled receptors, and nuclear receptors.

CONCLUSION

Tremendous advances have occurred in biotechnology since Watson and Crick determined the structure of DNA. Improved pharmaceuticals, novel therapeutic agents, unique diagnostic products, and new drug design tools have resulted from the escalating achievements of pharmaceutical biotechnology. While recombinant DNA technology and hybridoma techniques received most of the press in the late 1980s and early 1990s, a wealth of additional and innovative biotechnologies and approaches have been, and will continue to be, developed in order to enhance pharmaceutical research. Genomics, proteomics, transcriptomics, microarrays, pharmacogenomics/genetics, epigenomics, personalized medicine, metabonomics/metabolomics, toxicogenomics, glycomics, systems biology, genetically engineered animals, high-throughput screening, and high-speed combinatorial synthesis are directly influencing the pharmaceutical sciences and are well positioned to significantly impact modern pharmaceutical care. Application of these and yet to be discovered biotechnologies will continue to reshape effective drug therapy as well as improve the competitive, challenging process of drug discovery and development of new medicinal agents and diagnostics. Pharmacists, pharmaceutical scientists, and pharmacy students should be poised to take advantage of the products and techniques made available by the unprecedented scope and pace of discovery in biotechnology in the twenty-first century.

SELF-ASSESSMENT QUESTIONS

■ Questions

1. What were the increasing levels of genetic resolution of the human genome planned for study as part of the HGP?
2. What is functional genomics?
3. What is proteomics?
4. What are SNPs?
5. What is the difference between pharmacogenetics and pharmacogenomics?
6. Define metabonomics.
7. What is a DNA microarray?

8. What phase(s) of drug action is affected by genetic variation?
9. Define personalized medicine.
10. What is a biomarker?
11. Define systems biology.
12. Why are engineered animal models valuable to pharmaceutical research?
13. What two techniques are commonly used to produce transgenic animals?
14. What is a knockout mouse?
15. What are two approaches to high-throughput synthesis of drug leads and how do they differ?

■ Answers

1. HGP structural genomics was envisioned to proceed through increasing levels of genetic resolution: detailed human genetic linkage maps [approx. 2 megabase pairs (Mb = million base pairs) resolution], complete physical maps (0.1 Mb resolution), and ultimately complete DNA sequencing of the approximately 3.5 billion base pairs (23 pairs of chromosomes) in a human cell nucleus [1 base pair (bp) resolution].
2. Functional genomics is a new approach to genetic analysis that focuses on genome-wide patterns of gene expression, the mechanisms by which gene expression is coordinated, and the interrelationships of gene expression when a cellular environmental change occurs.
3. A new research area called proteomics seeks to define the function and correlate that with expression profiles of all proteins encoded within a genome.
4. While comparing the base sequences in the DNA of two individuals reveals them to be approximately 99.9 % identical, base differences, or polymorphisms, are scattered throughout the genome. The best-characterized human polymorphisms are single-nucleotide polymorphisms (SNPs) occurring approximately once every 1,000 bases in the 3.5 billion base pair human genome.
5. Pharmacogenetics is the study of how an individual's genetic differences influence drug action, usage, and dosing. A detailed knowledge of a patient's pharmacogenetics in relation to a particular drug therapy may lead to enhanced efficacy and greater safety. While sometimes used interchangeably (especially in pharmacy practice literature), pharmacogenetics and pharmacogenomics are subtly different. Pharmacogenomics introduces the additional element of our present technical ability to pinpoint patient-specific DNA variation using genomic techniques. While overlapping fields of study, pharmacogenomics is a much newer term that correlates an individual patient's DNA variation (SNP level of variation knowledge rather than gene level of variation knowledge) with his or her response to pharmacotherapy.
6. The field of metabonomics is the holistic study of the metabolic continuum at the equivalent level to the study of genomics and proteomics.
7. The biochips known as DNA microarrays and oligonucleotide microarrays are a surface collection of hundreds to thousands of immobilized DNA sequences or oligonucleotides in a grid created with specialized equipment that can be simultaneously examined to conduct expression analysis.
8. Genomic variation affects not only the pharmacokinetic profile of drugs (via drug metabolizing enzymes and drug transporter proteins), it also strongly influences the pharmacodynamic profile of drugs via the drug target.
9. Pharmacotherapy informed by a patient's individual genomic and proteomic information. Sometimes referred to as giving the right drug to the right patient in the right dose at the right time.
10. Biomarkers are clinically relevant substances used as indicators of a biologic state. Detection or concentration change of a biomarker may indicate a particular disease state physiology or toxicity. A change in expression or state of a protein biomarker may correlate with the risk or progression of a disease, with the susceptibility of the disease to a given treatment or the drug's safety profile.
11. Systems biology is the study of the interactions between the components of a biological system, and how these interactions give rise to the function and behavior of that system.
12. Engineered animal models are proving invaluable since small animal models of disease are often poor mimics of that disease in human patients. Genetic engineering can predispose an animal to a particular disease under scrutiny, and the insertion of human genes into the animal can initiate the development of a more clinically relevant disease condition.
13. (a) DNA microinjection and random gene addition and (b) homologous recombination in embryonic stem cells.
14. A knockout mouse, also called a gene knockout mouse or a gene-targeted knockout mouse, is an animal in which an endogenous gene (genomic wild-type allele) has been specifically inactivated by replacing it with a null allele.
15. There are two overall approaches to high-throughput synthesis. True combinatorial chemistry applies methods to substantially reduce the number of synthetic operations or steps needed to synthesize large numbers of compounds. CombiChem, as it is sometimes referred to, is conducted on solid supports (resins) to facilitate the needed manipulations that reduce labor. Differing from combinatorial chemistry, parallel procedures

apply automation to the synthetic process, but the number of operations needed to carry out a synthesis is practically the same as the conventional approach. Thus, the potential productivity of parallel methods is not as high as combinatorial chemistries. Parallel chemistries can be conducted on solid-phase supports or in solution.

Acknowledgment I wish to acknowledge the tremendous contribution of Dr. Arlene Marie Sindelar, my wife, to some of the graphics found in figures in all four editions of this textbook.

REFERENCES

- Abla N, Kroetz DL (2009) Drug transporter pharmacogenetics. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 70–89
- Aderem A, Hood L (2001) Immunology in the post-genomic era. *Nat Immunol* 2:373–375
- Alkan C, Coe BP, Eichler EE (2011) Genome structural variation discovery and genotyping. *Nat Rev Genet* 12:363–375
- Amaratunga D, Gohlmann H, Peeters PJ (2007) Microarrays. In: Kubinyi H (ed) *Comprehensive medicinal chemistry II*, vol 3. Elsevier, Amsterdam, pp 87–106
- Anderson PO, McGuinness SM, Bourne PE (2010) What is pharmacy informatics? In: Anderson PO, McGuinness SM, Bourne PE (eds) *Pharmacy informatics*. CRC Press, Boca Raton, pp 3–5
- Ashton M, Maloney B (2007) Solution phase parallel chemistry. In: Kubinyi H (ed) *Comprehensive medicinal chemistry II*, vol 3. Elsevier, Amsterdam, pp 761–790
- Baez J (2005) Biopharmaceuticals derived from transgenic plants and animals. In: Knablein J (ed) *Modern biopharmaceuticals*, vol 3. Wiley-VCH, Weinheim, pp 833–892
- Baker M (2012) Structural variation: the genome's hidden architecture. *Nat Methods* 9:133–137
- Basu S, Oyelere AK (2003) Application of recombinant DNA technology in medicinal chemistry and drug discovery. In: Abraham DJ (ed) *Burger's medicinal chemistry and drug discovery*, vol 2, 6th edn. John Wiley & Sons, Inc, New York, pp 81–114
- Beeley LJ, Duckworth DM, Southan C (2000) The impact of genomics on drug discovery. *Prog Med Chem* 37:1–43
- Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A (2009) An operational definition of epigenetics. *Genes Dev* 23:781–783
- Bikker JA, Trumpp-Kallmeyer S, Humblet C (1998) G-protein coupled receptors: models, mutagenesis, and drug design. *J Med Chem* 41:2911–2927
- Bjornsson HT, Fallin MD, Feinberg AP (2004) An integrated epigenetic and genetic approach to common human disease. *Trends Genet* 20:350–358
- Blake K, Wang J, Lima J (2009) Respiratory diseases. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 230–255
- Blomme E (2009) Fundamental principles of toxicogenomics. In: Semizarov D, Blomme E (eds) *Genomics in drug discovery and development*. John Wiley & Sons, Inc, Hoboken, pp 167–217
- Branca MA (2005) BiDil raises questions about race as a marker. *Nat Rev Drug Discov* 4:615–616
- Brazeau DA, Brazeau GA (2011a) Introduction: pharmacogenomics and pharmacogenetics – a historical look. In: *Principles of the human genome and pharmacogenomics*. American Pharmacists Association, Washington, D.C., pp 1–10
- Brazeau DA, Brazeau GA (2011b) Information flow in biological systems. In: *Principles of the human genome and pharmacogenomics*. American Pharmacists Association, Washington, D.C., pp 11–34
- Brazeau DA, Brazeau GA (2011c) Ethical challenges and opportunities in pharmacogenetics and pharmacogenomics. In: *Principles of the human genome and pharmacogenomics*. American Pharmacists Association, Washington, D.C., pp 81–95
- Browman G, Hebert PC, Coutts J, Stanbrook MB, Flegel K, MacDonald NE (2011) Personalized medicine: a windfall for science, but what about patients? *Can Med Assoc J* 183:E1277
- Bungard TJ, Yakiwchuk E, Foisy M, Brocklebank C (2011) Drug interactions involving warfarin: practice tool and practical management tips. *Can Pharm J* 144:21–25
- Burckert GJ (2009) Transplantation. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 256–270
- BusinessWeek (2012) Illumina and life technologies add genome-in-a-day machines. Available at: <http://www.businessweek.com/news/2012-01-18/illumina-life-technologies-add-genome-in-a-day-machines.html>. Accessed 19 Jan 2012
- Cebon J, Nicola N, Ward M, Gardner I, Dempsey P, Layton J, Duhrsen U, Burgess AW, Nice E, Morstyn G (1990) Granulocyte-macrophage colony stimulating factor from human lymphocytes. *J Biol Chem* 265:4483–4491
- Chalkidou K, Rawlins M (2011) Pharmacogenetics and the cost-effectiveness analysis: a two-way street. *Drug Discov Today* 16:873–883
- Chang MCY, Eachus RA, Trieu W, Ro DK, Keasling JD (2007) Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. *Nat Chem Biol* 3:274–277
- Chen M, Zhang L (2011) Epigenetic mechanisms in developmental programming of adult disease. *Drug Discov Today* 16:1007–1018
- Cherf GM, Lieberman KR, Rashid H, Lam CE, Karplus K, Akeson M (2012) *Nat Biotechnol* 30:344–348
- Cik M, Jurzak M (2007) High-throughput and high-content screening. In: Kubinyi H (ed) *Comprehensive medicinal chemistry II*, vol 3. Elsevier, Amsterdam, pp 679–696
- Clancy C, Collins FS (2010) Patient-centered outcomes research institute: the intersection of science and health care. *Sci Transl Med* 2:11–13, Collect

- Clark DP, Pazdernik NJ (2012a) Transgenic plants and plant biotechnology. In: Clark DP, Pazdernik NJ (eds) *Biotechnology: academic cell update*. Elsevier, Burlington, pp 397–423
- Clark DP, Pazdernik NJ (2012b) Transgenic animals. In: Clark DP, Pazdernik NJ (eds) *Biotechnology: academic cell update*. Elsevier, Burlington, pp 425–455
- Collins LJ, Barker AD (2007) Mapping the cancer genome. *Sci Am* 296:50–57
- Collins LJ, Schonfeld B (2011) The epigenetics of non-coding RNA. In: Tollesbol T (ed) *Handbook of epigenetics: the new molecular and medical genetics*. Elsevier, London, pp 49–61
- Cox AG (2010) Pharmacogenomics and drug transport/efflux. In: Zdanowicz MM (ed) *Concepts in pharmacogenomics*. American Society of Health-systems Pharmacists, Bethesda, pp 129–153
- Craik CS, Rocznik S, Largman C, Rutter WJ (1987) The catalytic role of the active site aspartic acid in serine proteases. *Science* 237:909–913
- Crews KR, Cross SJ, McCormick JN, Baker DK, Molinelli AR, Mullins R, Relling MV, Hoffman JM (2011) Development and implementation of a pharmacist-managed clinical pharmacogenetics service. *Am J Health Syst Pharm* 68:143–150
- D'Alessandro A, Zolla L (2010) Pharmacoproteomics: a chess game on a protein field. *Drug Discov Today* 15:1015–1023
- D'Alessandro A, Zolla L (2012) Metabolomics and cancer discovery: let the cells do the talking. *Drug Discov Today* 17:3–9
- Davidson RG, McNerney JD (2009) Principles of genetic medicine. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 341–379
- Davies K (2010) *The \$1,000 genome*. Free Press, a division of Simon and Schuster, Inc., New York
- Daxinger L, Whitelaw E (2012) Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nat Rev Genet* 13:153–162
- DePrimo SE (2007) Biomarkers. In: Kubinyi H (ed) *Comprehensive medicinal chemistry II*, vol 3. Elsevier, Amsterdam, pp 69–85
- Dieterle F, Sistare FD (2010) Biomarkers of acute kidney injury. In: Vaidya VS, Bonventre JV (eds) *Biomarkers in medicine, drug discovery, and environmental health*. John Wiley & Sons, Inc, Hoboken, pp 237–279
- Dietert RR (2010) Biomarkers of immunotoxicity. In: Vaidya VS, Bonventre JV (eds) *Biomarkers in medicine, drug discovery, and environmental health*. John Wiley & Sons, Inc, Hoboken, pp 307–322
- Dunn DA, Kooyman DL, Pinkert CA (2005) Transgenic animals and their impact on the drug discovery industry. *Drug Discov Today* 10:757–767
- Ellingrod VL, Bishop JR, Thomas K (2009) Psychiatry. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 204–229
- Feinberg AP, Irizarry RA, Fradin D, Aryee MJ, Murakami P, Aspelund T, Eirksdottir G, Harris TB, Launer L, Gudnason V, Fallin MD (2010) Personalized epigenomic signatures that are stable over time and covary with body mass index. *Sci Transl Med* 2:45–51, Collect
- Fenniri H (2000) *Combinatorial chemistry: a practical approach*. Oxford University Press, Oxford
- Flaumenhaft R (2007) Chemical biology. In: Kubinyi H (ed) *Comprehensive medicinal chemistry II*, vol 3. Elsevier, Amsterdam, pp 129–149
- Fox BI (2010) Informatics and the medication use process. In: Fox BI, Thrower MR, Felkey BG (eds) *Building core competencies in pharmacy informatics*. American Pharmacists Association, Washington, D.C., pp 11–31
- Frank R, Hargreaves R (2003) Clinical biomarkers in drug discovery and development. *Nat Rev Drug Discov* 2:564–565
- Frye RF (2009) Pharmacogenetics of oxidative drug metabolism and its clinical applications. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 32–53
- Fukuda M, Hindsgaul O (2000) *Molecular and cellular glycolysis*. Oxford University Press, Oxford
- Furness LM (2002) Genomics applications that facilitate the understanding of drug action and toxicity. In: Licinio J, Wong M-L (eds) *Pharmacogenomics, the search for individualized therapies*. Wiley-BCH Verlag GmbH & Co. KGaA, Weinheim, pp 83–125
- Garner I, Colman A (1998) Therapeutic proteins from livestock. In: Clark AJ (ed) *Animal breeding-technology for the 21st century*. Harwood Academic Publishers, Amsterdam, pp 215–227
- Garnett MJ (2012) Exploiting genetic complexity in cancer to improve therapeutic strategies. *Drug Discov Today* 17:188–193
- Garske AL, Denu JM (2009) New frontiers in epigenetic modifications. In: Sippl W, Jung M (eds) *Epigenetic targets in drug discovery*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, pp 3–22
- Gehlenborg N, O'Donoghue SI, Baliga NS, Goesmann A, Hibbs MA, Kitano H, Kohlbacher O, Neuweger H, Schneider R, Tenenbaum D, Gavin A-C (2010) Visualization of omics data for systems biology. *Nat Methods* 7:S56–S58
- Glass JI, Smith HO, Hutchison III CA, Alperovich NY, Assad-Garcia N (2007) Minimal bacterial genome. United States Patent Application 20070122826, May 31, 2007
- Goldman J, Becker ML, Jones B, Clements M, Leeder JS (2011) Development of biomarkers to optimize pediatric patient management: what makes children different? *Biomark Med* 5:781–794
- Grossman I, Goldstein DB (2010) Pharmacogenetics and pharmacogenomics. In: Ginsburg GS, Willard HF (eds) *Essentials of genomic and personalized medicine*. Academic Press, San Diego, pp 175–190
- Hamm CA, Costa FF (2011) The impact of epigenomics on future drug design and new therapies. *Drug Discov Today* 16:626–635

- Herceg Z, Murr R (2011) Mechanisms of histone modifications. In: Tollesbol T (ed) *Handbook of epigenetics: the new molecular and medical genetics*. Elsevier, London, pp 25–45
- Hindroff LA, MacArthur J (European Bioinformatics Institute), Wise A, Junkins HA, Hall PN, Klemm AK, Manolio TA (2012) A catalog of published genome-wide association studies. Available at: <http://www.genome.gov/gwastudies>. Accessed 14 April 2012
- Hong H, Shi L, Su Z, Ge W, Jones WD, Czika W, Miclauss K, Lambert CG, Vega SC, Zhang J, Ning B, Liu J, Green B, Xu L, Fang H, Perkins R, Lin SM, Jafari N, Park K, Ahn T, Chierici M, Furlanello C, Zhang L, Wolfinger RD, Goodsaid F, Tong W (2010) *Pharmacogenom J* 10: 364–374
- Hood L, Perlmutter RM (2004) The impact of systems approaches on biological problems in drug discovery. *Nat Biotechnol* 22:1215–1217
- Huston SA (2010) Ethical consideration in pharmacogenomics. In: Zdanowicz MM (ed) *Concepts in pharmacogenomics*. American Society of Health-systems Pharmacists, Bethesda, pp 375–397
- Hutchison CA, Phillips S, Edgell MH, Gillham S, Jahnke P, Smith M (1978) Mutagenesis at a specific position in a DNA sequence. *J Biol Chem* 253:651–656
- Jirtle RL, Skinner MK (2007) Environmental epigenomics and disease susceptibility. *Nat Rev Genet* 8: 253–262
- Johnson JA (2009) Drug target pharmacogenetics. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 90–106
- Johnson AC, Reitz M (1998) Site-directed mutagenesis. In: Greene JJ, Rao VB (eds) *Recombinant DNA principles and methodologies*. Marcel Dekker, Inc, New York, pp 699–719
- Kalow W (2009) Pharmacogenetics: a historical perspective. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 22–31
- Kaput J, Perlina A, Hatipoglu B, Bartholomew A, Nikolsky Y (2007) Nutrigenomics: concepts and applications to pharmacogenomics and clinical medicine. *Pharmacogenomics* 8:369–390
- Kassam S, Meyer P, Corfield A, Mikuz G, Sergi C (2005) Single nucleotide polymorphisms (SNPs): history, biotechnological outlook and practical applications. *Curr Pharmacogenomics* 3:237–245
- Kaye J, Curren L, Anderson N, Edwards K, Fullerton SM, Kanellopoulou N, Lund D, MacArthur DG, Mascalonzi D, Shepherd J, Taylor PL, Terry SF, Winter SF (2012) From patients to partners: participant-centric initiatives in biomedical research. *Nat Rev Genet* 13:371–376
- Kelloff GJ, Sigman CC (2012) Cancer biomarkers: selecting the right drug for the right patient. *Nat Rev Drug Discov* 11:201–214
- Khalil AS, Collins JJ (2010) Synthetic biology: applications come of age. *Nat Rev Genet* 11:367–379
- Khan FA (2012a) Animal biotechnology. In: Khan FA (ed) *Biotechnology fundamentals*. CRC Press, Boca Raton, pp 201–235
- Khan FA (2012b) Agricultural biotechnology. In: Khan FA (ed) *Biotechnology fundamentals*. CRC Press, Boca Raton, pp 163–199
- Khan FA (2012c) Ethics in biotechnology. In: Khan FA (ed) *Biotechnology fundamentals*. CRC Press, Boca Raton, pp 397–415
- Klimyuk V, Marillonnet S, Knablein J, McCaman M, Gleba Y (2005) Biopharmaceuticals derived from transgenic plants and animals. In: Knablein J (ed) *Modern biopharmaceuticals*, vol 3. Wiley-VCH, Weinheim, pp 893–917
- Klipp E, Herwig R, Kowald A, Wierling C, Lehrach H (2005) Basic principles. In: Klipp E, Herwig R, Kowald A, Wierling C, Lehrach H (eds) *Systems biology in practice*. Wiley-BCH Verlag GmbH & Co. KGaA, Weinheim, pp 3–17
- Knoell DL, Sadee W (2009) Applications of genomics in human health and complex disease. In: AACP Editorial Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 1–19
- Kobata A (1996) Function and pathology of the sugar chains of human immunoglobulin G. *Glycobiology* 1:5–8
- Kodavanti UP (2010) Lung injury biomarkers. In: Vaidya VS, Bonventre JV (eds) *Biomarkers in medicine, drug discovery, and environmental health*. John Wiley & Sons, Inc, Hoboken, pp 157–201
- Kolesar JM (2009) Drug target pharmacogenomics of solid tumors. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 123–134
- Koomer A (2010) Pharmacogenomics in pharmacy education. In: Zdanowicz MM (ed) *Concepts in pharmacogenomics*. American Society of Health-systems Pharmacists, Bethesda, pp 363–374
- Koomer A, Ansong MA (2010) The role of pharmacists in pharmacogenomics. In: Zdanowicz MM (ed) *Concepts in pharmacogenomics*. American Society of Health-systems Pharmacists, Bethesda, pp 341–350
- Krynetskiy E, Calligaro IL (2009) Introducing pharmacy students to pharmacogenomic analysis. *Am J Pharm Educ* 73:1–7
- Kubinyi H (2007) Chemogenomics. In: Kubinyi H (ed) *Comprehensive medicinal chemistry II*, vol 3. Elsevier, Amsterdam, pp 921–937
- Kubinyi H, Müller H (2004) Chemogenomics in drug discovery a medicinal chemistry perspective. In: Kubinyi H, Müller G, Mannhold R, Folkers G. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA
- Laursen L (2010) Big science at the table. *Nature* 468:S2–S4
- Lee LJ, Hughes TR, Frey BJ (2006) How many new genes are there? *Science* 311:1709
- Lee KC, Ma JD, Kuo GM (2009) Pharmacogenomics: bridging the gap between science and practice. *Pharm Today* 15(12):36–48

- Lengauer T, Hartmann C (2007) Bioinformatics. In: Kubinyi H (ed) *Comprehensive medicinal chemistry II*, vol 3. Elsevier, Amsterdam, pp 315–348
- Limdi NA, Rettie AE (2009) Warfarin pharmacogenetics. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 143–164
- Lindpainter K (2007) Pharmacogenomics. In: Kubinyi H (ed) *Comprehensive medicinal chemistry II*, vol 3. Elsevier, Amsterdam, pp 51–68
- Loging W, Harland L, Williams-Jones B (2007) High-throughput electronic biology: mining information for drug discovery. *Nat Rev Drug Discov* 6:220–230
- Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordaro A, Elbert BL, Mak RH, Ferando AA (2005) MicroRNA expression profiles classify human cancers. *Nature* 435:834–838
- Maetzler W, Berg D (2010) Biomarkers of Alzheimer's and Parkinson's disease. In: Vaidya VS, Bonventre JV (eds) *Biomarkers in medicine, drug discovery, and environmental health*. John Wiley & Sons, Inc, Hoboken, pp 91–117
- Mason JS, Pickett SD (2003) Combinatorial library design, molecular similarity, and diversity applications. In: Abraham DJ (ed) *Burger's Medicinal chemistry and drug discovery*, vol 1, 6th edn. John Wiley & Sons, Inc, New York, pp 187–242
- Mayfield SP, Franklin SE (2005) Expression of human antibodies in eukaryotic micro-algae. *Vaccine* 23(15):1828–1832
- McCurry KR, Kooyman DL, Alvarado CG, Cotterell AH, Martin MJ, Logan JS, Platt JL (1995) Human complement regulatory proteins protect swine-to-primate cardiac xenografts from humoral injury. *Nat Med* 1:423–427
- McDonagh EM, Whirl-carillo M, Garten Y, Altman RB, Klein TE (2011) From pharmacogenomics knowledge acquisition to clinical applications: the PharmGKB as a clinical pharmacogenomics biomarker resource. *Biomark Med* 5:795–806
- McGuire AL, Evans BJ, Caulfield T, Burke W (2010) Regulating direct-to-consumer personal genome testing. *Science* 330:181–182
- McLean AS, Huang SJ (2010) Biomarkers of cardiac injury. In: Vaidya VS, Bonventre JV (eds) *Biomarkers in medicine, drug discovery, and environmental health*. John Wiley & Sons, Inc, Hoboken, pp 119–155
- Meyer P, Alexopoulos LG, Bonk T, Califano A, Cho CR, de la Fuente A, de Graaf D, Hartemink AJ, Hoeng J, Ivanov NV, Koepl H, Linding R, Marbach D, Norel R, Peitsch MC, Rice JM, Royyuru A, Schacherer F, Sprengel J, Stolle K, Vitkup D, Stolovitzky G (2011) Verification of systems biology research in the age of collaborative competition. *Nat Biotechnol* 29:811–813
- Miclaus K, Chierici M, Lambert C, Zhang L, Vega S, Hong H, Yin S, Furlanello C, Wolfinger R, Goodsaid F (2010) Variability in GWAS analysis: the impact of genotype calling algorithm inconsistencies. *Pharmacogenomics J* 10:324–335
- Mitscher LA, Dutta A (2003) Combinatorial chemistry and multiple parallel synthesis. In: Abraham DJ (ed) *Burger's medicinal chemistry and drug discovery*, vol 2, 6th edn. John Wiley & Sons, Inc, New York, pp 1–35
- Momary KM, Crouch MA (2010) Cardiovascular disease. In: Zdanowicz MM (ed) *Concepts in pharmacogenomics*. American Society of Health-systems Pharmacists, Bethesda, pp 183–211
- Murphy JE, Green JS, Adams LA, Squire RB, Kuo GM, McKay A (2010) Pharmacogenomics in the curricula of colleges and schools of pharmacy in the United States. *Am J Pharm Educ* 74:1–10
- Nature Biotechnology Editorial Staff (2012) What happened to personalized medicine? *Nat Biotechnol* 30:1
- Nicholson JK, Wilson ID (2003) Understanding 'global' systems biology: metabonomics and the continuum of metabolism. *Nat Rev Drug Discov* 2:668–676
- Opar A (2011) 'Pharmers' hope for first plant drug harvest. *Nat Rev Drug Discov* 10:81–82
- Ossorio P (2004) Societal and ethical issues in pharmacogenomics. In: Allen WL, Johnson JA, Knoell DL, Kolesar JM, McInerney JD, McLeod HL, Spencer HT, Tami JA (eds) *Pharmacogenomics: applications to patient care*. American College of Clinical Pharmacy, Kansas City, pp 399–442
- Ozer JS, Reagan WJ, Schomaker S, Palandra J, Baratta M, Ramaiah S (2010) Translational biomarkers of acute drug-induced liver injury: the current state, gaps, and future opportunities. In: Vaidya VS, Bonventre JV (eds) *Biomarkers in medicine, drug discovery, and environmental health*. John Wiley & Sons, Inc, Hoboken, pp 203–236
- Padgett L, O'Connor S, Roederer M, McLeod H, Ferreri S (2011) Pharmacogenomics in a community pharmacy: ACT now. *J Am Pharm Assoc* 51:189–193
- Paik Y-K, Jeong S-K, Omenn GS, Uhlen M, Hanash S, Cho SY, Lee H-J, Na K, Choi E-Y, Yan F, Zhang F, Zhang Y, Snyder M, Cheng Y, Chen R, Marko-Varga G, Deutsch EW, Kim H, Kwon J-Y, Aebersold R, Bairoch A, Taylor AD, Kim KY, Lee E-Y, Hochstrasser D, Legrain P, Hancock WS (2012) A chromosome-centric human proteome project to characterize the sets of proteins encoded in the genome. *Nature Biotechnol* 30:221–223
- Patti GJ, Yanes O, Siuzdak G (2012) Metabolomics: the apogee of the omics trilogy. *Nat Rev Mol Cell Biol* 13:263–269
- Pennisi E (2010) 1000 Genomes Project gives new map of genetic diversity. *Science* 330:574–575
- Petros WP, Sharma M (2009) Pharmacogenomics of hematologic malignancies. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 135–142
- Petrucelli N, Daly MB, Feldman GL (2011) *BRCA1* and *BRCA2* hereditary breast and ovarian cancer. *Gene Reviews* [Internet]. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK1247/>. Accessed 12 Feb 2012
- PhRMA (2012) 2011 report: medicines in development – biotechnology. Available at: <http://www.phrma.org/research/new-medicines>. Accessed 5 Feb 2012
- Pirmohamed M (2011) Pharmacogenetics: past, present and future. *Drug Discov Today* 16:852–861

- Pirrung MC (2004) Molecular diversity and combinatorial chemistry. Elsevier, Amsterdam
- Portela A, Esteller M (2010) Epigenetic modifications and human disease. *Nat Biotechnol* 28:1057–1068
- Price ND, Edelman LB, Lee I, Yoo H, Hwang D, Carlson G, Galas DJ, Heath JR, Hood L (2010) Systems biology and systems medicine. In: Ginsburg GS, Willard HF (eds) *Essentials of genomic and personalized medicine*. Academic Press, San Diego, pp 131–141
- Pritchard JF, Jurima-Romet M (2010) Enabling go/no go decisions. In: Bleavins MR, Carini C, Jurima-Romet M, Rahbari R (eds) *Biomarkers in drug development – a handbook of practice, application, and strategy*. John Wiley & Sons, Inc, Hoboken, pp 31–39
- Raghavachari N (2012) Overview of omics. In: Barh D, Blum K, Madigan MA (eds) *OMICS-biomedical perspectives and applications*. CRC Press, Boca Raton, pp 1–19
- Rahbar A, Rivers R, Boja E, Kinsinger C, Mesri M, Hiltke T, Rodriguez H (2011) Realizing individualized medicine: the road to translating proteomics from the laboratory to the clinic. *Personal Med* 8:45–57
- Rakyan VK, Down TA, Balding DJ, Beck S (2011) Epigenome-wide association studies for common human diseases. *Nat Rev Genet* 12:529–541
- Rankovic Z, Jamieson C, Morphy R (2010) High throughput screening approach to lead discovery. In: Rankovic Z, Morphy R (eds) *Lead generation approaches in drug discovery*. John Wiley & Sons, Inc, Hoboken, pp 21–71
- Reardon J (2011) The ‘persons’ and ‘genomics’ of personal genomics. *Personal Med* 8:95–107
- Rockett JC (2003) The future of toxicogenomics. In: Burczynski ME (ed) *An introduction to toxicogenomics*. CRC Press LLC, Boca Raton, pp 299–317
- Rogers PD (2009) Infectious diseases. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 165–175
- Rothberg BEG, Pena CEA, Rothberg JM (2005) A systems biology approach to target identification and validation for human chronic disease drug discovery. In: Knablein J (ed) *Modern biopharmaceuticals, design, development and optimization*, vol 1. Wiley-BCH Verlag GmbH & Co. KgaA, Weinheim, pp 99–125
- Rudolph NS (1995) Advances continue in production of proteins in transgenic animal milk. *Genet Eng News*, October 15, 8–9
- Rudolph NS (2000) Biopharmaceutical production in transgenic livestock. *Trends Biotechnol* 17:367–374
- Sanger Institute (2012) The cancer genome project. Available at: <http://www.sanger.ac.uk/genetics/CGP/>. Accessed 6 Mar 2012
- Schniecke AE, Kind AJ, Ritchie WA, Mycock K, Scott AR, Ritchie M, Wilmut I, Colman A, Campbell KHS (1997) Human factor IX transgenic sheep produced by transfer of nuclei from transfected fetal fibroblasts. *Science* 278:2130–2133
- Schrattenholz A, Groebe K, Soskic V (2010) Systems biology approaches and tools for analysis of interactomes and multi-target drugs. In: Yan Q (ed) *Systems biology in drug discovery and development: methods and protocols*. Springer, New York, pp 29–58
- Semizarov D (2009a) Genomics technologies as tools in drug discovery. In: Semizarov D, Blomme E (eds) *Genomics in drug discovery and development*. John Wiley & Sons, Inc, Hoboken, pp 25–103
- Semizarov D (2009b) Genomic biomarkers. In: Semizarov D, Blomme E (eds) *Genomics in drug discovery and development*. John Wiley & Sons, Inc, Hoboken, pp 105–166
- Seneci P (2007) Combinatorial chemistry. In: Kubinyi H (ed) *Comprehensive medicinal chemistry II*, vol 3. Elsevier, Amsterdam, pp 315–348
- Sharpless NE, DePinho RA (2006) The mighty mouse: genetically engineered mouse models in cancer drug development. *Nat Rev Drug Discov* 5:741–754
- Shaw EC, Johnson PWM (2012) Stratified medicine for cancer therapy. *Drug Discov Today* 17:261–268
- Shultz LD, Ishikawa F, Greiner DL (2007) Humanized mice in translational biomedical research. *Nat Rev Immunol* 7:118–130
- Singh AA, Somvanshi P (2012) Bioinformatics – a brief introduction to changing trends in modern biology. In: Barh D, Blum K, Madigan MA (eds) *OMICS-biomedical perspectives and applications*. CRC Press, Boca Raton, pp 23–39
- Smith HE (2010) Toxicogenomics. In: Zdanowicz MM (ed) *Concepts in pharmacogenomics*. American Society of Health-systems Pharmacists, Bethesda, pp 321–337
- Smith HO, Hutchison CA III, Pfannkoch C, Venter JC (2003) Generating a synthetic genome by whole genome assembly: ϕ X174 bacteriophage from synthetic oligonucleotides. *Proc Natl Acad Sci U S A* 100:15440–15445
- Staff (2007) A mouse for all reasons – The International Mouse Knockout Consortium. *Cell* 128:9–13
- States DJ, Omenn GS, Blackwell TW, Fermin D, Eng J, Speicher DW, Hanash SM (2006) Challenges in deriving high-confidence protein identifications from data gathered by HUPO plasma proteome collaborative study. *Nat Biotechnol* 24:333–338
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102:15545–15550
- Sucholeiki I (2001) *High-throughput synthesis: principles and practices*. Marcel Dekker, Inc, New York
- Takeuchi M, Takasaki S, Shimada M, Kobata A (1990) Role of sugar chains in the in vitro biological activity in human erythropoietin produced in recombinant Chinese hamster ovary cells. *J Biol Chem* 265:12127–12130
- The International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921
- Thornton M, Gladwin A, Payne R, Moore R, Cresswell C, McKechnie D, Kelly S, March R (2005) Automation and validation of DNA-banking systems. *Drug Discov Today* 10:1369–1375
- Topol EJ (2010) Pharmacy benefit managers, pharmacies, and pharmacogenomic testing: prescription for progress? *Sci Transl Med* 2:14–16, collect
- Treangen TJ, Salzberg SL (2011) Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat Rev Genet* 13:36–46

- Tursz T, Andre F, Lazar V, Lacroix L, Soria J-C (2011) Implications of personalized medicine – perspective from a cancer center. *Nat Clin Coll, Personalized cancer medicine* 8(3):S46–S52
- U.S. DOE (2012a) Human genome project information. Available at: http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml. Accessed 3 Jan 2012
- U.S. DOE (2012b) Human genome project information. Available at: http://www.ornl.gov/sci/techresources/Human_Genome/faq/snps.shtml. Accessed 27 Mar 2012
- U.S. National Academies (2011) Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease. The National Academies Press, Washington, D.C., pp 1–4, http://www.nap.edu/catalog.php?record_id=1328
- U.S. NIH (2012) The cancer genome atlas. Available at: <http://cancergenome.nih.gov/>. Accessed 7 Mar 2012
- Ullmann D (2007) Fluorescence screening assays. In: Kubinyi H (ed) *Comprehensive medicinal chemistry II*, vol 3. Elsevier, Amsterdam, pp 599–615
- Van Eyck AS, Bouzin C, Feron O (2010) Both host and graft vessels contribute to revascularization of xenografted human ovarian tissue in a murine model. *Fertil Steril* 93:1676–1685
- Varkin A, Cummings R, Esko J, Freeze H, Hart G, Marth J (1999) *Essentials of glycobiology*. Cold Spring Harbor Laboratory Press, LaJolla
- Veenstra TD (2010) Proteomics for biomarker discovery. In: Vaidya VS, Bonventre JV (eds) *Biomarkers in medicine, drug discovery, and environmental health*. John Wiley & Sons, Inc, Hoboken, pp 25–46
- Venter JC et al (2001) The sequence of the human genome. *Science* 291:1304–1351
- Waldman SA, Terzic A (2011) Patient-centric clinical pharmacology advances the path to personalized medicine. *Biomark Med* 5:697–700
- Walsh G, Jefferis R (2006) Post-translational modifications in the context of therapeutic proteins. *Nat Biotechnol* 24:1241–1252
- Wang L (2009) Phase II drug-metabolizing enzymes. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 54–69
- Warner S (2004) Diagnostics + therapy = theranostics. *Scientist* 18:38–39
- Webster KD (2010) Pharmacogenomics: pharmacodynamics and pharmacogenomics. In: Zdanowicz MM (ed) *Concepts in pharmacogenomics*. American Society of Health-systems Pharmacists, Bethesda, pp 155–180
- Wenk MR (2005) The emerging field of lipidomics. *Nat Rev Drug Discov* 4:594–610
- Weston GS (2010) The pharmacogenetics of drug metabolism. In: Zdanowicz MM (ed) *Concepts in pharmacogenomics*. American Society of Health-systems Pharmacists, Bethesda, pp 85–127
- Wildt S, Gerngross TU (2005) The humanization of N-glycosylation pathways in yeast. *Nat Rev Microbiol* 3:119–126
- Williams SA, Slavin DE, Wagner JA, Webster CJ (2006) A cost effectiveness approach to the qualification and acceptance of biomarkers. *Nat Rev Drug Discov* 5:897–902
- Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, Cheng D, Jewell K, Arndt D, Sawhney S, Fung C, Nikolai L, Lewis M, Coutouly M-A, Forsythe I, Tang P, Shrivastava S, Jeroncic K, Stothard P, Amegbey G, Block D, Hau DD, Wagner J, Miniaci J, Clements M, Gebremedhin M, Guo N, Zhang Y, Duggan GE, MacInnis GD, Weljie AM, Dowlatabadi R, Bamforth F, Clive D, Greiner R, Li L, Marrie T, Sykes BD, Vogel HJ, Querengesser L (2007) HMDB: the human metabolome database. *Nucleic Acids Res* 35(Database Issue):D521–D526
- Woollard PM, Mehta NAL, Vamathevan JJ, Van Horn S, Bonde BK, Dow DJ (2011) The application of next-generation sequencing technologies to drug discovery and development. *Drug Discov Today* 16:512–519
- Woster PM (2010) Epigenetic targets and cancer drug discovery. *Annu Rep Med Chem* 45:245–260
- Wright FA et al (2001) A draft annotation and overview of the human genome. *Genome Biol* 2:1–18
- Wu R, Tong C, Wang Z, Mauger D, Tantisira K, Szeffler SJ, Cinchilli VM, Isreal E (2011a) A conceptual framework for pharmacodynamics genome-wide association studies in pharmacogenomics. *Drug Discov Today* 16:884–890
- Wu W, Zhang HH, Welsh MJ, Kaufman PB (2011b) Bioinformatics superhighway and computer databases of nucleic acids and proteins. In: *Gene technology*, 3rd edn. CRC Press, Boca Raton, pp 133–153
- Wu W, Zhang HH, Welsh MJ, Kaufman PB (2011c) Strategies for gene double knockout. In: *Gene technology*, 3rd edn. CRC Press, Boca Raton, pp 351–380
- Yang Y, Adelstein SJ, Kassis AI (2009) Target discovery from data mining approaches. *Drug Discov Today* 14:147–154
- Yuryev A (2012) Pathways analysis in drug discovery. In: Young DL, Michelson S (eds) *Systems biology in drug discovery and development*. John Wiley & Sons, Inc, Hoboken, pp 289–302
- Zdanowicz MM (2010) Pharmacogenomics: past, present, and future. In: Zdanowicz MM (ed) *Concepts in pharmacogenomics*. American Society of Health-systems Pharmacists, Bethesda, pp 3–17
- Zemles T (2010) An inservice program on pharmacogenetics to individualize drug therapy. *Am J Pharm Educ* 74:1–5
- Zhou S (2006) Clinical pharmacogenomics of thiopurine S-methyltransferase. *Curr Clin Pharmacol* 1:119–128
- Zineh I, Huang S-M (2011) Biomarkers in drug development and regulation: a paradigm for clinical implementation of personalized medicine. *Biomark Med* 5:705–713

- Zineh I, Pacanowski M (2009) Cardiovascular diseases. In: American College of Clinical Pharmacy Board (ed) *Pharmacogenomics: applications to patient care*, 2nd edn. American College of Clinical Pharmacy, Kansas City, pp 176–203
- Zuckerman R, Milne C-P (2012) Market watch: industry perspectives on personalized medicine. *Nat Rev Drug Discov* 11:178–179