Personalized medicine: revolutionizing drug discovery and patient care

Geoffrey S. Ginsburg and Jeanette J. McCarthy

Advances in human genome research are opening the door to a new paradigm for practising medicine that promises to transform healthcare. Personalized medicine, the use of marker-assisted diagnosis and targeted therapies derived from an individual's molecular profile, will impact the way drugs are developed and medicine is practiced. Knowledge of the molecular basis of disease will lead to novel target identification, toxicogenomic markers to screen compounds and improved selection of clinical trial patients, which will fundamentally change the pharmaceutical industry.The traditional linear process of drug discovery and development will be replaced by an integrated and heuristic approach. In addition, patient care will be revolutionized through the use of novel molecular predisposition, screening, diagnostic, prognostic, pharmacogenomic and monitoring markers. Although numerous challenges will need to be met to make personalized medicine a reality, with time, this approach will replace the traditional trial-and-error practice of medicine.

The post-genome era has begun, and with it the promise of tailoring the practice of medicine to the individual. This emerging field of personalized medicine encompasses the use of risk algorithms, molecular diagnostics, targeted therapies and pharmacogenomics to improve health care. Personalized medicine will provide the link between an individual's molecular and clinical profiles, allowing physicians to make the right patient-care decisions and allowing patients the opportunity to make informed and directed lifestyle decisions for their future well-being. Molecular diagnostics, the use of DNA-, protein- or mRNA-based biological markers to predict the risk of developing disease or the molecular phenotype of an existing one, will change the way we currently define disease. Genomic analysis of diseases with homogeneous clinical phenotypes will unveil distinct molecular entities that require different treatment strategies for optimal outcomes. Clinical diseases as we know them will be replaced by molecular classification. Therapies directed at the root cause of disease will replace those that simply treat the symptoms of disease. Finally, a pharmacogenomic test that predicts therapy response based on a patient's genomic profile will accompany many drugs. Personalized medicine will involve radical changes in the pharmaceutical industry and medical practice and is likely to affect many aspects of society. Most importantly, the individual whose health is at stake will benefit enormously.

Geoffrey S. Ginsburg*
Millennium Pharmaceuticals Inc., 45 Sidney Street, Cambridge, MA 02139, USA.
*e-mail: ginsburg@mpi.com

Jeanette J. McCarthy
Millennium Pharmaceuticals Inc., One Kendall Square, Bldg 700, Cambridge, MA 02139, USA.

Innovation in pharmaceutical industry R&D strategy

Knowledge of the molecular basis of disease is already transforming pharmaceutical development. Drug discovery and development has traditionally been a linear process (Fig. 1) with little feedback from later clinical development stages on the overall process. The adoption of a personalized medicine strategy in drug discovery and development necessitates a paradigm shift from a linear process to an integrated and heuristic one (Fig. 2). This new approach will involve a series of research feedback loops. The early stages of discovery, including selection and validation of drug targets, small-molecule screening and chemistry, and preclinical assessment of compounds, will be linked with later stages of clinical development. Molecular, pharmacological and patient clinical data will be captured at various phases and integrated in a 'knowledge management system' that will be used to facilitate rational drug design around molecular diseases.

Genomic technologies have already taken hold and are impacting the pharmaceutical industry. High-throughput sequencing and transcript profiling have been applied to cell-based and animal models of disease or directly to human tissues to identify rapidly gene targets that initiate the drug discovery process. Bioinformatics, proteomics and animal models are used to further validate genes as targets before proceeding to high-throughput screening of vast compound libraries for the development of small-molecule drugs. The impact of genomics on drug development can already be seen: earlier this year, Millennium Pharmaceuticals (Cambridge, MA) and Bayer AG (Leverkeusen, Germany) announced what is believed to be the first small molecule drug candidate discovered against a genomics-derived target in the field of cancer[1]. In the near future, a flood of new drug therapies targeted at the molecular basis of disease will become available.

Genomic technologies applied to target identification can simultaneously identify genes that are co-regulated with drug targets. Both targets and co-regulated genes could be potential surrogate biomarkers for use in preclinical and clinical studies, an example of integration of the early and late stages of drug discovery and development. Ideal surrogate markers include cell-surface proteins and secreted proteins, which are amenable to sensitive mass-spectroscopic or antibody-based detection in the blood. The gene encoding leptin, a regulator of body fat discovered using genomic technologies[2], is not only proven to be a valuable drug target but blood leptin levels might be of use as a monitoring marker of drug-associated weight gain[3] or as a response to growth-hormone treatment in children[4].

Further down the discovery process, toxicogenomic markers predictive for adverse drug reactions (ADRs) might influence selection and optimization of lead compounds before human studies. Microarray analytical tools to define molecular profiles that predict ADRs in humans are being investigated using existing drugs that are known to produce unwarranted

http://tibtech.trends.com 0167-7799/01/$ – see front matter PII: S0167-7799(01)01814-5

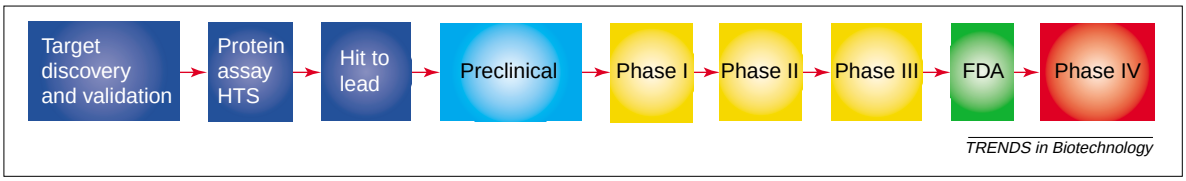

Fig. 1. Traditional drug discovery – a linear process. The historical process of drug discovery has been linear, with little opportunity for feedback or improvement on the upstream components of the process from downstream results. Abbreviations: FDA, Food and Drug Administration; HTS, high-throughput screening.

hepatotoxicity, nephrotoxicity, cardiotoxicity or bone marrow suppression. Companies such as Affymetrix (Santa Clara, CA), GeneLogic (Gaithersburg, MD) and Curagen (New Haven, CT) are developing gene expression-based assays that can be used to test preclinical compounds for their propensity to induce ADRs based on these studies. Predictive toxicogenomic screening of preclinical compounds should become as routine in preclinical drug development as it is today to examine an association between a compound and the HERG (human ether-related a-GoGo) potassium channel[5] – a test that predicts whether the compound will produce prolongation of electrocardiographic QT intervals, and possibly sudden death, in patients.

Pharmacodynamic and pharmacogenomic markers predictive of drug toxicity in humans can be introduced into Phase I, II or III clinical trials where, in principle, patient selection and/or stratification within studies can be guided on the basis of markers correlating with safety and efficacy. Recent studies of human genetic variation in the cytochrome P450 (CYP) enzymes that are largely responsible for drug metabolism, for example, have suggested that using individual genetic variation at these loci to select patients for clinical trials might reduce ADRs by 10%–20% (Ref. 6). Pharmacogenomics will be an important component of personalized medicine and is already being embraced by pharmaceutical companies as a means of improving efficiency in the drug development process. An individual's response to a drug is the complex combination of both genetic and non-genetic factors. Genetic variants in the drug target itself, disease pathway genes or drug-metabolizing enzymes, might all be used as predictors of drug efficacy or toxicity. The pharmaceutical industry has recognized the *a priori* need for tools to enable pharmacogenomic research. In 1999, ten companies and the Wellcome Trust formed a consortium to discover and map the most common type of genetic variation, single nucleotide polymorphisms (SNPs). To date, >800 000 SNPs have been deposited into the SNP Consortium's public database (http://www.snp.cshl.org). A high-resolution SNP map will expedite the identification of genes for complex diseases, such as asthma, diabetes mellitus, atherosclerosis and psychiatric disorders. The SNP database will also be a tool for pharmacogenomic investigations during clinical development. Today, many pharmaceutical companies are designing their clinical trials to enable the routine collection and storage of DNA and other biological specimens that will be used in future pharmacogenomic studies.

Careful biological monitoring during clinical development will not only lead to pharmacogenomic markers that accompany the drug on the market but also will afford opportunities to apply human biological information to earlier phases of discovery and development. Molecular profiles of patients identified in Phase I and II clinical studies as likely non-responders (potentially indicating complex molecular taxonomy of the disease being treated) might represent an opportunity for pharmaceutical companies to initiate discovery programs. Novel therapies could be developed around the non-responders' specific molecular subclass of disease.

The personalized medicine strategy for drug discovery and development should yield a spectrum of product opportunities for the pharmaceutical industry. Diagnostic risk assessment and disease-monitoring tools that accurately quantify disease burden in patients will be a direct outcome of research during the early discovery process. Pharmacogenomic markers of efficacy and side effects will be used in conjunction with specific drugs to target drug therapy to those patients who will have an optimal response. The business rationale for targeted therapies, which some argue will decrease market share, is that such products will eventually expand the market by recruiting patients from less effective therapies or by identifying less symptomatic individuals who might benefit from prophylactic therapy.

The clinical phases of drug development afford the opportunity to capture patient clinical data, imaging and *in vitro* molecular response data simultaneously. Academic medical centers and clinical research organizations are now conducting clinical trials with future research in mind. Archiving biological specimens along with traditional clinical covariates is becoming routine. Some centers are also actively engaged in pharmacogenomic marker research. In the near future, clinical trials might be conducted in specialized units where detailed clinical, biological and genomic data are collected and integrated. Genome- and proteome-wide profiles together with biological pathway databases, imaging and clinical data on every patient will be used to analyze an individual's disease and drug response. The understanding of the biology of disease and drug action gleaned from these sophisticated new paradigms will dramatically accelerate the realization of truly personalized medicine (Box 1).

Molecular diagnosis will determine prognosis and therapy

Personalized medicine is rooted in the hypothesis that diseases are heterogeneous, from their causes to rates of progression to their response to drugs. Each person's disease might be unique and therefore that person needs to be treated as an individual. With limited

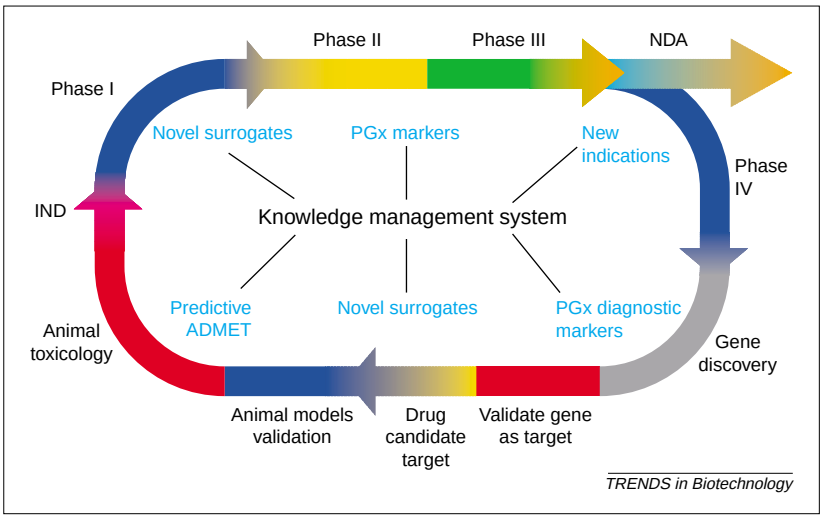

Fig. 2. Future drug discovery – an integrated process. Genomic information and markers emerging at each stage of the discovery process will be used as tools both upstream and downstream, resulting in better pharmaceuticals and personalized medicine products. A knowledge warehouse will store information enabling continued process and product improvements. Abbreviations: ADMET, absorption, distribution, metabolism, excretion and toxicity; IND, investigational new drug; NDA, new drug application; PGx, pharmacogenomics.

understanding of the molecular basis of disease, we have relied on non-specific clinical signs. As genomic tools are sharpened, so will be our ability to dissect disease into its component parts. Clinical phenotypes thought to be one disease will be subclassified by a new genomic taxonomy. Recent discoveries in the molecular pathology of cancer have highlighted important and clinically significant differences in the gene expression patterns of a variety of tumors, including leukemias[7] and breast cancer[8]. In cardiovascular disease, genetic heterogeneity has been identified in the Long QT syndrome, a disorder of ventricular depolarization where clinical manifestations range from no visible signs to sudden death. The etiology of the Long QT syndrome is attributed to mutations in one of at least four different ion channels (HERG, KVLQT1, SCN5A or KCNE1)[9]. The clinical course of the disease, level of aggressive therapy and choice of therapy (Na^+-channel blocker versus K^+-channel blocker versus beta blocker) are now determined by the genetic etiology of the syndrome[10–12].

Familial hypertrophic cardiomyopathy is another example of a genetically heterogeneous disease with a clinical phenotype of ventricular hypertrophy. Familial cardiomyopathy results from >80 different mutations, each affecting the expression of a cardiac muscle sarcomeric protein. Mutation-specific prognoses have been established that mandate screening to determine who requires more frequent clinical monitoring, therapeutic intervention and family screening[13]. As the underlying molecular architecture of other diseases is determined, medical practice will be tailored to properly diagnose and treat them.

Innovation in patient care

The ultimate goal of personalized medicine is to define disease at the molecular level so that preventive resources and therapeutic agents can be directed at the right population of people while they are still well (Box 2). The application of new technologies and the integration of data from an individual will lead to a new paradigm in patient care that will emerge from strategies employed in pharmaceutical research and development – a paradigm that will, for the first time, allow physicians to take a global molecular view of an individual patient's disease. During the course of a chronic disease with a long clinical prodrome, the research and product development strategies for personalized medicine aims to impact the course of the disease at six major points (Fig. 3). Genetic variants can be used to predict the predisposition of an individual for future disease development. Genetic variants associated with increased or decreased risk of disease will be the basis of genotype-directed treatment recommendations. Individuals deemed at high risk of disease can be targeted for preventive therapy or lifestyle modifications. Preventive therapies have been fully embraced by the medical community, as evidenced by the use of selective estrogen-receptor modulators for patients at risk of breast cancer[14] and osteoporosis[15], and the use of statins in patients at risk of developing coronary artery disease[16]. High-risk individuals should be periodically screened (using protein-based markers, serum analytes and/or molecular imaging) for preclinical disease detection. The molecular equivalent of the pap smear, mammogram or blood-pressure measurement will define more precisely the predilection for disease development. In patients with preclinical or symptomatic disease, molecular diagnosis based on gene- or protein-expression fingerprints might differentiate diverse diseases with similar clinical phenotypes. A set of different molecular markers could determine prognosis (the slope of the curve), distinguishing those with an aggressive form and rapid progression of disease from individuals with slower disease progression, tailoring therapy accordingly. In choosing a therapeutic, the decision is guided by molecular markers (pharmacogenomics) that correlate with safety and efficacy of specific compounds. Finally, monitoring the disease progression following therapy will utilize many of the molecular markers developed for screening and diagnosis.

Box 1. Personalized medicine: advantages to the pharmaceutical industry

Increased efficiency and reduced costs of target and lead discovery
Reduced timelines and costs of clinical trials
Emergence of new gene targets for drug discovery
Product differentiation in the market place

Box 2. Personalized medicine: advantages to patients and clinicians

Higher probability of desired outcome with a drug
Low probability of untoward side effects
Preventive strategies
Focused therapies
Reduced costs
Better health and better healthcare

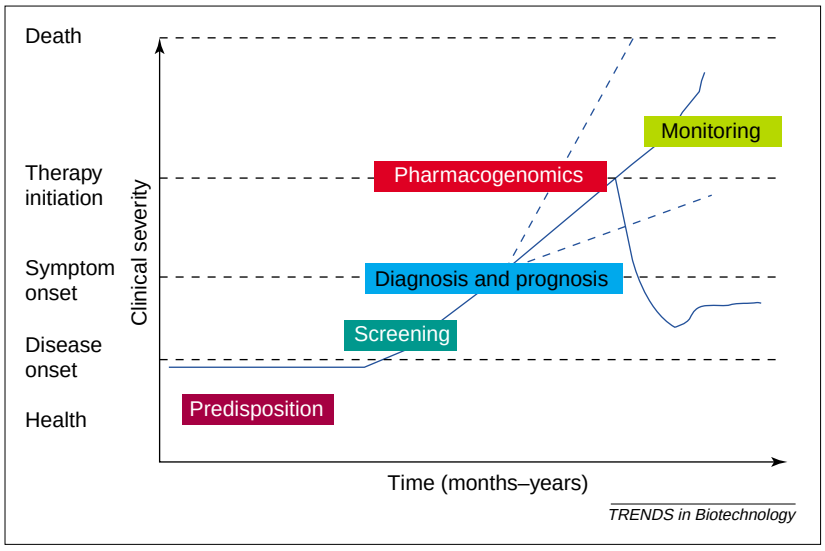

Fig. 3. Research, intervention and personalized medicine opportunities at stages of a hypothetical disease. Six areas where personalized medicine strategies will benefit the individual are predisposition, screening, diagnosis, prognosis, pharmacogenomics and monitoring. At each stage the types of markers and their commercial applications will vary, as they will for specific diseases.

Personalized medicine: it's happening now

The past few years have seen major advances in technology and the growth of genomic information as a byproduct of the human genome project. As a result, new and innovative markers of disease are being uncovered at an unprecedented rate. At the DNA level, >350 genetic tests are currently available (http://www.genetests.org). Although most tests are for rare, monogenic disorders, some are becoming available for more common, complex diseases. Examples include APOE testing among dementia patients for differential diagnosis of Alzheimer disease, and Factor V Leiden testing for predisposition to venous thrombosis. Furthermore, for most common diseases, a large number of genetic markers suggesting association with disease are reported in the literature. Advances in SNP-discovery technologies are providing opportunities for large-scale candidate gene studies[17,18]. Indeed, whole-genome association studies are being contemplated for finding genetic predisposition markers for common, complex diseases[18]. The next three to five years will see an explosion of new information in this area and the development of new predictive tests for complex diseases.

Identification of RNA and protein markers for screening, diagnosis, prognosis and monitoring is also under way, facilitated through advances in transcript profiling and proteomics. The basic research methods used in the discovery of these markers require access to relevant disease tissues. Because tumor samples are routinely biopsied or removed, the first disease area likely to benefit from these technologies is cancer. Advances in our ability to classify disease are best illustrated in the work of Golub *et al.*[7] This landmark paper illustrates how gene-expression profiling can be used to classify two related cancer types. Expression patterns of 50 genes were determined to distinguish accurately between acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML). Similar approaches have been taken to identify candidate prognostic markers for melanoma. Clark *et al.*[20] and Bittner *et al.*[21] used transcript profiling to compare metastatic with nonmetastatic human melanoma cell lines. As a result, they identified several genes that are selectively unregulated in the metastatic lines that could have use in patient management.

The acknowledgement that post-transcriptional modification of proteins might be an important determinant of disease is one factor driving the use of proteomic technologies for discovering molecular disease markers. These technologies include traditional 2D gel electrophoresis in addition to more advanced mass-spectrometry methods. Proteomic analysis might be used in medical microbiology in which the entire proteome of an organism can be studied or in diseases such as rheumatoid arthritis or central nervous system disorders in which protein-rich fluids at the site of injury, such as synovial joint and cerebral spinal fluid, respectively, are available for analysis. Cell line supernatants or explants from tumor tissues have already been used in large-scale expression profiling experiments to identify cancer markers. Page *et al.*[22] compared the proteome of purified normal human luminal (from which most breast cancer is derived) with that of myoepithelial breast cells. They detected 170 proteins that differed between the two cell types. These experiments might shed light onto the process of cancer development and ultimately find use as cancer diagnosis or monitoring markers.

Pharmacogenomic tests are finding their way into practice in several disease areas. Genotype resistance testing of HIV isolates has demonstrable clinical utility and provides a way to assist therapeutic decision-making in patients whose levels of HIV RNA are rising[23]. In addition, assays are available to detect the HER2 protein receptor or copies of the HER2 gene sequence to determine eligibility for herceptin treatment or adriamycin treatment, respectively, in node-positive breast cancer patients[24]. The Food and Drug Administration (FDA) approval of the pharmacogenomic marker HER2 linked to herceptin represents an important precedent for regulatory approval of personalized medicine products.

Challenges of realizing the promise of personalized medicine

In spite of the achievement of a complete human genome sequence, there are numerous challenges in realizing the personalized medicine vision (Box 3). Identifying genetic variants that are markers of disease or drug response requires sifting through several million SNPs in the human genome to find those that contribute to the disease and then demonstrating that the SNPs are clinically valid markers and are useful for managing patients. To uncover DNA variants that predict common, complex diseases that result from a combination of genes and environmental factors will require cost-effective, high-throughput genotyping; large, well-characterized patient populations; sophisticated computational methodologies; and a detailed understanding of the biological pathways of disease. Uncovering mRNA and protein markers for

use in screening, diagnosis, prognosis and monitoring of disease will have its own set of challenges. Access to optimal relevant tissues might not be possible for many diseases. Proteomic technologies require further development, as do computational approaches for analyzing massive amounts of gene and protein profile data. To realize the vision of personalized medicine, the agenda for medical and pharmaceutical research must include the assembly and integration data from many sources on large numbers of patients. Clinical investigations should incorporate genotyping and molecular profiling technologies along with traditional clinical data collection and should establish a repository of patient samples where possible.

New molecular markers might face many hurdles before they can be implemented in patient care. The issues range from FDA regulation and acceptance of these new markers, to developing tractable assay platforms, to resolving issues around the ethical, legal and social implications of obtaining highly sensitive genetic information. Foremost among these, in our opinion, is the education and engagement of physicians and patients in the paradigm shift to objective, quantitative marker-based clinical care. If appropriate patient management systems, integrated databases, educational tools and genetic counseling are not in place, then it will be difficult to realize the significant benefits forecast from this approach. Fortunately, we have already learned valuable lessons from past efforts to implement genetic screening for sickle cell anemia, and from more recent efforts to screen for BRCA1 mutations in breast cancer families. The ethical, legal and social implications (ELSI) of human genetic research are the subject of a government-funded program (http://www.nhgri.nih.gov/ELSI/). The US government is playing an active role in addressing public concern over genetic information by, among other things, drafting legislation to protect patients from discrimination by employers and insurance companies[25]. Furthermore, health professionals are rising to the challenge of educating both their members and the public. The American Medical Association has co-founded the National Coalition for Health Professional Education in Genetics (http://www.nchpeg.org) to promote health professional education and access to genetic information.

Box 3. Challenges of realizing the promise of personalized medicine

SNP or DNA-based marker discovery
Access to patient populations
Genotyping costs
Computational methodologies

mRNA and protein marker discovery
Access to clinical tissue samples
Technology development (proteomics)
Computational methodologies

Marker utilization in practice
Assay platform development
Large scale data and knowledge management
Ethical, legal, and social considerations
Physician and patient education

A vision for the practice of medicine in the 21st century

In the next decade, medical care will undergo revolutionary changes. No longer will medical practice be limited to the empirical extrapolation of a patient's care from generalized clinical-trial results. Traditional medical practice, based on trial-and-error, results in both under-treatment and over-treatment, multiple office visits, the need for drug monitoring, and frequent regimen changes. More than 100 000 deaths per year are attributed to adverse drug reactions[26]. A personalized approach of tailored care for every individual based on their specific, molecular disease will become the standard of care. In the prototypical office visit of 2015, the physician will examine a patient's genetic profile (stored on CD ROMs or equivalent), lifestyle, and results from objective molecular screening and monitoring tests. Algorithms, derived from previous research efforts, will be used to compute the likelihood that a patient develops a host of chronic diseases. The focus of medicine at this juncture will be entirely preventive. Lifestyle modifications and the use of prophylactic therapy will be recommended based on what is best for that patient to avoid chronic disease to which they could be susceptible. The 'office' of the future might itself be virtual; Internet office visits might supplant some of the direct patient–physician contact. Patients will be more knowledgeable of their own health and risk profiles and more active in directing their own healthcare.

Summary

Personalized medicine promises to offer the right treatment for the right patient at the right time. Although that promise might seem far off, there is clear evidence that the traditional trial-and-error practice of medicine is eroding in favor of more precise marker-assisted diagnosis and treatment. For the patient, the benefits are clear: safer and more effective treatment of disease. For industry there appears an equally desirable outcome of this approach: increased efficiency, productivity and better product lines. Society as a whole will also realize a benefit: more focused application of precious healthcare resources to those in need of them most. The realization of personalized medicine is not without challenges, yet many of these challenges are being addressed. By encouraging public dialogue and debate, we expect that there will be continued progress forward. Lastly, as we take on more and more of the burden of our own health and well-being, educational forums must be developed for patients and physicians alike to understand the complex nature of the genomic information that is being used for decision making. Then we will have truly fulfilled the promise of the future.

References

1 Malakoff, D. and Service, R.F. (2001) Genomania meets the bottom line. *Science* 291, 1193–1203
2 Zhang, Y. *et al.* (1994) Positional cloning of the mouse obese gene and its human homologue. *Nature* 372, 425–432
3 Melkersson, K.L. *et al.* (2000) Elevated levels of insulin, leptin and blood lipids in olanzapine-treated patients with schizophrenia or related psychoses. *J. Clin. Psychiatry* 61, 742–749
4 Tillmann, V. *et al.* (2000) Monitoring serum insulin-like growth factor-I (IGF-I), IGF binding protein-3 (IGFBP-3), IGF-1/IGFBP-3 molar ratio and leptin during growth hormone treatment for disordered growth. *Clin. Endocrinol.* 53, 329–336
5 Mitcheson, J.S. *et al.* (2000) A structural basis for drug-induced long QT syndrome. *Proc. Natl. Acad. Sci. U. S. A.* 97, 12329–12333
6 Miller, M.C. *et al.* (2001) Genetic variability in susceptibility and response to toxicants. *Toxicol. Lett.* 269–280
7 Waring, J.F. *et al.* (2001) Microarray analysis of hepatotoxins *in vitro* reveals a correlation between gene expression profiles and mechanisms of toxicity. *Toxicol. Lett.* 120, 359–368
8 Golub, T.R. *et al.* (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286, 531–537
9 Hedenfalk, I. *et al.* (2001) Gene-expression profiles in hereditary breast cancer. *New Engl. J. Med.* 344, 539–544
10 Priori, S.G. *et al.* (1999) Genetic and molecular basis of cardiac arrhythmias: Impact on clinical management. Part III. *Circulation* 99, 674–681
11 Zareba, W. *et al.* (1998) Influence of genotype on clinical course of the long-QT syndrome. *New Engl. J. Med.* 339, 960–965
12 Shimizu, W. and Antzelevitch, C. (1997) Sodium channel block with mexiletine is effective in reducing dispersion of repolarization and preventing torsade de pointes in LQT2 as well as LQT3 models of the long-QT syndrome. *Circulation* 96, 2038–2047
13 Schwartz, P.J. *et al.* (1995) Long QT syndrome patients with mutations of the SCN5A and HERG genes have differential responses to Na^+ channel blockade and to increases in heart rate: implications for gene-specific therapy. *Circulation* 92, 3381–3386
14 Fananapazir, L. (1999) Advances in molecular genetics and management of hypertrophic cardiomyopathy. *J. Am. Med. Assoc.* 281, 1746–1752
15 Jordan, V.C. (2000) Progress in the prevention of breast cancer: concept to reality. *J. Steroid Biochem. Mol. Biol.* 74, 269–277
16 Doren, M. and Samsioe, G. (2000) Prevention of postmenopausal osteoporosis with oestrogen replacement therapy and associated compounds: Update on clinical trials since 1995. *Hum. Reprod.* 6, 419–426
17 Crouch, M.A. (2001) Effective use of statins to prevent coronary heart disease. *Am. Fam. Physician* 63, 309–320
18 Cargill, M. *et al.* (1999) Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat. Genet.* 22, 231–238
19 Halushka, M.K. *et al.* (1999) Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat. Genet.* 22, 239–247
20 Kruglyak, L. (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat. Genet.* 22, 139–144
21 Clark, E.A. *et al.* (2000) Genomic analysis of metastasis reveals an essential role for RhoC. *Nature* 406, 532–535
22 Bittner, M. *et al.* (2000) Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* 406, 536–540
23 Page, M.J. *et al.* (1999) Proteomic definition of normal human luminal and myoepithelial breast cells purified from reduction mammoplasties. *Proc. Natl. Acad. Sci. U. S. A.* 96, 12589–12594
24 Durant, J. *et al.* (1999) Drug-resistance genotyping in HIV-1 therapy: the VIRADAPT randomised controlled trial. *Lancet* 353, 2195–2199
25 Tsongalis, G.J. *et al.* (2000) Gene amplification as means for determining therapeutic strategies in human cancers. *Clin. Chem. Lab. Med.* 38, 837–839
26 Jeffords, J.M. and Daschle, T. (2001) Political issues in the genome era. *Science* 291, 1249–1251
27 Lazarou, J. *et al.* (1998) Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *J. Am. Med. Assoc.* 279, 1200–1205

Drug discovery of the future: the implications of the human genome project

Thomas Reiss

The elucidation of the 3.2-gigabase human genome will have various impacts on drug discovery. The number of drug targets will increase by at least one order of magnitude and target validation will become a high-throughput process. To benefit from these opportunities, a theory-based integration of the vast amount of new biological data into models of biological systems is called for. The skills and knowledge required for genome-based drug discovery of the future go beyond the traditional competencies of the pharmaceutical industry. Cooperation with biotechnology firms and research institutions during drug discovery and development will become even more important.

Thomas Reiss
Fraunhofer Institute for Systems and Innovation Research, Breslauer Str. 48, 76139 Karlsruhe, Germany.
e-mail: rt@isi.fhg.de

Until the late 20th century, drug discovery was mainly a linear process based on the screening and testing of thousands of chemical and natural substances for potential therapeutic activity. Screening was time consuming and more or less random because drug targets and drug functions were in most cases not known. The set of drug targets for the pharmaceutical industry was rather small and according to a recent estimate[1], only 483 targets account for all drugs on the market. Approximately 45% of these are cell membrane receptors, 28% are enzymes and the remaining classes comprise hormones (11%), ion channels (5%), nuclear receptors (2%) and DNA (2%). About 7% of the targets are not known biochemically. The surprisingly low number of drug targets illustrates that the identification of interesting targets was a main bottleneck of the drug discovery process. In addition, the technology portfolio of the pharmaceutical industry was rather narrow and based mainly on chemistry and pharmacology. With the advent of genome research during the past decade, this traditional, although successful, concept of drug development started to change considerably. The publication of the human genome sequence in February 2001 (Refs 2,3) marks a new area of biological research that will have significant implications for drug discovery.

The human genome

As a basis for the following discussion, Table 1 provides a brief summary of important features of the human genome[2–4]. The total size of the genome is estimated to 3.2 gigabases (Gb). Of this, 2.95 Gb is euchromatin, which represents the gene-regions of a genome. Approximately 28% of the sequence is transcribed into RNA. Only 5% of the transcribed sequence (equivalent to 1.1–1.4% of the total sequence) encodes protein. The total number of

  PII: S0167-7799(01)01811-X