The Role of Genomics in Modern Drug Discovery

Michael Steinmetz*

Genomics is a new area of research which will lead to an unprecedented understanding of physiological and pathophysiological processes at the molecular level. As a result, new and more efficacious drugs will be discovered, including small-molecular-weight pharmaceuticals, protein drugs and gene therapeutics. New diagnostics for early detection of diseases and disease susceptibilities will emerge. Also new preventive medicines and preemptive interventions to avoid certain diseases altogether will be developed, based on our increased scientific understanding of molecular mechanisms of pathogenesis.

Human Diseases Have a Genetic Basis

All human diseases, with the exception of some traumatic events, are influenced by genetic factors. The genetic contribution is most evident in monogenic diseases, like sickle-cell anemia, hemophilia and cystic fibrosis, which are caused by mutations of single genes (Fig. 1). Complex or multifactorial diseases, like allergies, cancer, diabetes, obesity and many other common diseases, are caused by unfortunate combinations of genetic, behavioral and environmental factors. Even infectious diseases are influenced in the course they take by host genes.

Of the many genes associated with human diseases, only a few have been identified to date. They are conveniently classified into disease-causing genes (monogenic diseases), susceptibility and modifier genes (complex or multifactorial diseases) and resistance genes (infectious diseases) (Fig. 2).

Understanding Pathophysiological Mechanisms

A rational way to elucidate molecular mechanisms underlying diseases is to first identify the associated genes and then to unravel the biochemistry of pathogenesis. The better we understand molecular pathogenesis, the better we get at identifying the most appropriate molecular targets for disease intervention.

Genomic techniques, in particular genotyping and positional cloning, have significantly improved our ability to identify genes associated with human diseases. About 50 disease genes have been positionally cloned so far. With more genes being sequenced and mapped to their respective locations in the human genome, other approaches, like the candidate-gene approach, will become possible and speed up disease-gene identification. These genes can then be used as tools to characterize biochemical pathways controlled by them and the most appropriate target protein for disease intervention can be chosen. This should facilitate the development of drugs with increased efficacy and reduced side effects. Also, the overall success rate of the R&D process should improve (Fig. 3).

For most cases, however, it is not easy to go from genes to drug targets. Functional studies are required to understand the biochemical pathways and to identify the most appropriate point of disease intervention. They are complex, very labor-intensive and often of unpredictable outcome (Fig. 4).

In the emerging area of functional genomics, academic groups and small biotech companies are trying to make functional studies more efficient. Less complex animal model systems, like Drosophila and nematodes are being investigated (with or without implanted human genes). Because of the ease of genetic manipulation, these animals are well suited to de-
fine basic biochemical pathways and to elucidate the function of human disease genes. Furthermore, new techniques are being developed to automate the identification of proteins that interact in a given pathway and to facilitate the generation of mouse strains with defects in any particular gene of interest. Progress is steady but slow. Functional studies remain to be the bottleneck.

Proven Drug Targets belong to Large Gene Families

Given the current bottleneck in moving from disease genes to drug targets, are there alternative ways to explore genomic information for new and more efficient strategies in drug discovery? We believe there are. One strategy we are particularly interested in, focuses on the development of conventional, small-molecular-weight pharmaceuticals, rather than on proteins or gene therapeutics. It takes into account that the types of targets recognized by conventional pharmaceuticals fall into certain classes of proteins. For instance, if one looks at the target proteins recognized by the top 100 prescription drugs sold in 1995, one finds that 80% of them fall into four different classes of proteins: 36 drugs inhibit 15 different enzymes, 22 bind to 6 different G-protein-coupled (GPC) receptors, 12 interact with 5 different ion channels and 9 bind to 5 different nuclear hormone receptors.

These proven drug targets are members of large gene families. It is currently estimated that there are about 10000 different enzymes, 1000 different GPC receptors, ca. 200 different ion channels and perhaps 100 different nuclear hormone receptors.

From Viral and Bacterial Genomes to Antivirals and Antibacterials

The genomes of viruses, bacteria and parasites are less complex than the human genome, and drug targets can be more readily identified.

The human immunodeficiency virus is a good example. The development of drugs to fight the virus started with the isolation of the virus and the determination of its full genomic sequence in 1983. The sequence revealed three different enzymes encoded by the viral genome: a reverse transcriptase inhibitor, and Invirase®, the first protease inhibitor, to the market in 1992 and 1995, respectively.

Viral genomes in general are small and usually contain only 10–100 genes which makes sequencing and target selection relatively easy. Bacterial genomes contain 1000–10000 genes. With more sophisticated sequencing technology and the help of computers, the entire sequences of several bacterial genomes have been determined. The first, published in July 1995, was the sequence of Hemophilus influenzae. It consists of 1.8 x 10^8 base pairs and contains 1743 genes.

Roche has set up a collaboration with Human Genome Sciences, Inc., to determine the genome sequences of Staphylococcus aureus and Streptococcus pneumoniae, two bacteria associated with severe infections which often cannot be treated adequately with available antibiotics. The goal is to identify new antibacterial targets based on the genome sequences and to develop novel, more efficacious antibiotics.

From the Human Genome to Drugs

The human genome is at least an order of magnitude more complex than bacterial genomes, with an estimated set of 100000 genes. Large-scale experimentation and data handling is required to exploit the rapidly increasing information on the human genome as a tool in drug discovery.
Such large-scale techniques are already available or are being developed now. They address four distinct steps of the process.

First, the Human Genome Project, which started in October 1990, is project-ed to decipher the entire sequence by the year 2003 (Fig. 5). Once finished, all potential pharmaceutical target proteins will be known. However, we don't have to wait so long to get started. Partial human gene sequences have already become available in certain proprietary and public databases for most, if not all human genes based on the sequencing of cDNAs obtained from gene transcripts. These sequences have been annotated and can already be used to identify members of pharmaceutically interesting gene families (Fig. 6).

Second, chip hybridization technologies have come available which can be used to determine expression profiles for several thousand genes simultaneously. Thus, it will be possible to identify genes which are specifically expressed in certain human cells or tissues, and which are up- or downregulated in certain disease states. Also, their responses to physiological and pharmacological stimuli can be analyzed. Gene-expression profiles will help to select from a very large number of potentially interesting drug genes a smaller number of primary candidate genes. These genes would subsequently be expressed and the proteins isolated to develop screening assays for the identification of potent drug leads.

Third, ultrahigh flux, miniaturized screening technologies are being developed using 'biochips' and 'nanoplates' to facilitate the screening of up to a million synthetic compounds in a few days. Screening larger libraries in miniaturized format and in shorter time frames saves money and facilitates the identification of potentially more potent drug leads and the analysis of larger numbers of potential drug targets.

Fourth, combinatorial chemistry or parallel synthesis techniques can be used to generate large compound libraries for screening and for lead optimization (Fig. 7).

Taken together these new technologies will facilitate the exploration of the human genome for the discovery of new drug targets and the identification of potential lead compounds for drug development. The overall goal is to speed up discovery of more and more potent drug leads which can be validated directly in cell-based assays and animal models. Validating a target with a small molecule which is close to the final drug is of course much better than validating a target with a gene or protein which, in the case of small-molecule drug discovery, is very different from the final drug.

To explore the human genome for drug discovery, Roche has established collaborations with Incyte, Synteni and The Jackson Laboratory to gain access to human gene sequences, cDNA clones, gene-expression profiles and expertise in bioinformatics. Additional collaborations are planned to develop automated, miniaturized instruments for ultrahigh throughput screening and to synthesize large compound libraries by combinatorial chemistry.

Summary

In conclusion (Fig. 8), in genome-based drug discovery at Roche two different strategies are used. First, in collaboration with Millennium, positional chemi-
**Roche Pharma Research from the Past to the Present**

Eckart Gwinner* and Bruno Dalle Carbonare

This year, on October 1, the Basel firm *F. Hoffmann-La Roche* celebrated its 100th anniversary. During its life span, *Roche* has grown from a small laboratory to a broadly based multinational health-care company that is driven from the beginning by intensive research. Measured by its annual sales, *Roche* is ranking today in the league of the top ten pharmaceutical companies in the world and holds the number one position in the hospital sector. Last year, the *Roche* group invested on Research and Development (R&D) again more than 20% of its pharmaceutical sales of 9.2 billion CHF (prescription drugs and OTC products). These expenditures of 1.96 billion CHF, among the highest in the industry both in absolute and relative terms, have been spent to generate further growth above average (Table 1).

**Table 1:**

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Research Center</th>
<th>Partner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify drugs based on genes associated with obesity and type II diabetes</td>
<td>Nutley</td>
<td>Millennium: Positional cloning, target identification</td>
</tr>
<tr>
<td>Identify drugs based on genes coding for potential pharmaceutical targets</td>
<td>Basel</td>
<td>Human Genome Sciences: Genome sequencing</td>
</tr>
<tr>
<td>- Novel antivirals for HBV, HCV, HPV infections</td>
<td>Welwyn</td>
<td>None</td>
</tr>
<tr>
<td>- Novel antibiotics for Staph. aureus, Strep. pneumonise infections</td>
<td>Nutley</td>
<td>Incyte: Human gene sequences, transcript maps, cDNA clones</td>
</tr>
<tr>
<td>- All therapeutic categories</td>
<td>Synteni</td>
<td>Gene expression profiling</td>
</tr>
<tr>
<td></td>
<td>Jackson Laboratories</td>
<td>Bioinformatics</td>
</tr>
</tbody>
</table>

The Beginning: A Young Entrepreneur

With financial assistance from his father, the 28-year-old businessman Fritz Hoffmann, just married to Adele La Roche, established 1896 a small pharmaceutical specialties company on the banks of the Rhine River in Basel. At this time, when the Industrial Revolution was changing the face of Europe, migration from rural areas to centers of industry and trade was in full swing, causing city populations to swell and the need of medicines to grow rapidly.

However, a standardized drug therapy in the modern sense did not exist. In those days, most prescriptions had to be compounded individually by pharmacists. The effectiveness and safety of medicines could vary very considerably depending on the quality of the raw materials used, the skill of the pharmacists and the experience of the prescribing physicians. Pharmacies provided the first synthetic compounds from coal-tar and, of course, preparations like morphine or quinine which had been isolated from plants or plant components. What was lacking, however, was a broad-

*Correspondence: Diplomchemiker E. Gwinner*  
*Pharma Research Staff International – Communication*  
*F. Hoffmann-La Roche Ltd.*  
*CH-4070 Basel*