CLINICAL GENOMICS

Genomics, Genetic Epidemiology, and Genomic Medicine

KONSTANTINOS N. LAZARIDIS* and GLORIA M. PETERSEN*,‡
*Division of Gastroenterology and Hepatology, Center for Basic Research in Digestive Diseases, and ‡Division of Epidemiology, Mayo Clinic College of Medicine, Rochester, Minnesota

Medical science is on the threshold of unparalleled progress as a result of the advent of genomics and related disciplines. Human genomics, the study of structure, function, and interactions of all genes in the human genome, promises to improve the diagnosis, treatment, and prevention of disease. This opportunity is the result of the recent completion of the Human Genome Project. It is anticipated that genomics will bring to physicians a powerful means to discover hereditary elements that interact with environmental factors leading to disease. However, the expected transformation toward genomics-based medicine will occur over decades. It will require efforts of many scientists and physicians to begin now to sort out the vast amounts of information in the human genome and translate it to meaningful applications in clinical practice. Meanwhile, practicing physicians and health professionals need to be trained in the principles, applications, and limitations of genomics and genomic medicine. Only then will we be in a position to benefit patients, which is the ultimate goal of accelerating scientific progress in medicine. In this inaugural article, we introduce and discuss concepts, facts, and methods of genomics and genetic epidemiology that will be drawn on in the forthcoming topics of the clinical genomics series.

In April 1953, the seminal discovery of the double helical structure of DNA by James Watson and Francis Crick[1] revolutionized the biologic sciences. Exactly 50 years later, the complete sequence of the human genome became a reality,[2,3] a scientific landmark achieved during a period of 13 years by an international effort known as the Human Genome Project (HGP). If the pre-genomic era was ended by the complete sequencing of the genome of *Homo sapiens,* then the genome (or post-genome) era has already started.[4,5] Given these extraordinary scientific achievements, it is timely to assess the current status and future influence of genomics science in gastrointestinal and liver diseases. In this issue of the journal, precisely 52 years after the description of the double helix, we launch the first installment of a series on clinical genomics in *Clinical Gastroenterology and Hepatology*. Topics will be included in the series to cover both single-gene (ie, Mendelian) and complex (ie, multifactorial) gastrointestinal and liver diseases.

The knowledge gained by completion of the HGP, coupled with the rise of the discipline of genomics and other related scientific fields, will positively promote basic and translational studies to better understand the interplay of genetic predisposition and environmental factors in causing disease. These 2 elements have to be dissected to shed light on disease pathogenesis and devise novel treatments before we can prevent illnesses. Amidst all this exciting scientific progress, one important question comes to mind. Will genomics alter the means we currently use to diagnose, treat, and prevent gastrointestinal and liver diseases? The answer to this question is not simple, but the promise is enormous. We need first to understand where we currently stand and to recognize the challenges and opportunities that lie ahead.

In this article, we present principles and approaches of genomics and genetic epidemiology that will be discussed in the future themes of the clinical genomics series. Genomics defines a scientific field that aspires to investigate the structure, function, and interaction of all genes in the entire human genome.[6] Genetic epidemiology is the discipline that investigates the basis for susceptibility to disease by using family and population studies.[7] To this end, 3 remarks are in order. First, although we are at the beginning of the genomic era, this period is more than simply discovering human genes. Indeed, what should define the genomic era in gastroenterology and hepatology is understanding the functions of thousands of genes that are involved in regulating the cellular and molecular pathways of the digestive system and liver. Second, the human genome is in perpetual interaction with the environmental factors that operate long before birth and have a significant con-

Abbreviations used in this paper: HGP, Human Genome Project; SNP, single nucleotide polymorphism; TDT, transmission disequilibrium test.

1542-3565/05/$30.00
PII: 10.1053/S1542-3565(05)00085-6

tribution to gastrointestinal and liver disease biology. Third, clinicians will have a significant influence on the direction and application of genomics in gastroenterology and hepatology, as they effectively assess and classify pertinent disease phenotypes and traits.

In the next sections, we discuss the differences between single-gene and complex diseases, particularly as these relate to discovering disease-causing genes; the structure and elements of human genome; the variation of human genome; the relation of genetic variation to disease phenotypes; study designs to dissect disease-causing genetic variants; the human haplotype map; and the ethical, legal, and social implications of human genomics. A basic glossary of genomics and genetic epidemiology terms is included in the Appendix to familiarize the readers with the terminology used in the present and upcoming articles of the clinical genomics series that will follow.

Single-Gene Diseases Versus Complex Diseases

From a "genetics" perspective, all diseases, aside from most cases of trauma, have a genetic component.[8] In broad terms, there are 3 categories of genetic disorders: chromosomal, single-gene (Mendelian), and complex (multifactorial).

Chromosomal diseases are the outcome of deletion or addition of intact chromosomes or segments that affect approximately 1% of live-born deliveries. Many chromosomal disorders lead to spontaneous abortions or miscarriages because lack of or aberrant chromosomes are usually incompatible with life. Thus, the peak incidence of chromosomal disorders occurs before birth. In case of a live-born delivery, the deficit or gain of a chromosome results in profound physical characteristics of the newborn (eg, Down syndrome – trisomy 21). This is because there is loss or addition of hundreds of genes that are normally expressed on a chromosome. The advent of techniques like cytogenetics to study the chromosome number and structure in peripheral blood lymphocytes or other cells has facilitated the prenatal diagnosis and prevention of chromosomal disorders (eg, through amniocentesis).

Single-gene diseases exhibit familial patterns consistent with autosomal recessive, autosomal dominant, or X-linked inheritance. The characteristic of dominant inheritance is that only a single trait-causing allele located on an autosome or an X chromosome is required to express the phenotype. The hallmark of an autosomal recessive phenotype necessitates that both alleles (ie, paternal and maternal) be present for the trait to be expressed. To develop an X-linked recessive disease, a male needs only a single trait-causing allele, whereas the female needs both alleles. Moreover, for expressing an X-linked dominant trait, only a single allele is required in either male or female. Mendelian diseases segregate in families, and full expression of the disease is caused by a few rare mutations of a single gene. In a given family, the same mutation is responsible for the disease phenotype; in another family, a different mutation of the same gene might occur. Single-gene diseases are uncommon in the population; the most frequent is hereditary hemochromatosis, which affects 1 of every 300 individuals. The genetic basis of Mendelian diseases is considered simple because of the direct correspondence of a specific genotype to a phenotype (Figure 1). More than 1000 genes causing Mendelian diseases have been identified. This catalog of human genes linked to genetic diseases is available on-line at *Mendelian Inheritance in Man* (OMIM) (www.ncbi.nlm.nih.gov/omim).

Complex diseases, such as irritable bowel syndrome, nonalcoholic steatohepatitis, and inflammatory bowel disease, are considered multifactorial in etiology. It is believed that these diseases are caused by interaction of several genetic variants with environmental factors; thus, the term "complex" diseases (Figure 2). As a result, the direct correspondence of one genotype to one phenotype that characterizes a single-gene disease does not exist in a complex disorder. This fundamental concept might explain the proposed heterogeneity of complex disease etiology and/or the variation of phenotypes (Figure 2), that is, disease manifestations, progression, and response to treatment. Thus, complex diseases have a genetic component, not strictly Mendelian, but demonstrate familial aggregation, in which the risk of disease among relatives of the proband is greater than the estimated risk in the general population.[9] The term, relative risk ratio of a sibling (λ_s), was coined to define the risk of a sibling developing a specific disease if a biologic brother or sister is already affected. The λ_s is calculated by dividing the prevalence of disease among siblings with the prevalence of disease in the general population.[9] Therefore, the higher the value of λ_s, the greater the evidence for a genetic role in a disease.

Mendelian and complex diseases operate at different ends of a spectrum. Although their prevalence in the population is low, Mendelian diseases are the result of a single-gene with high penetrance of the phenotype. On the other hand, complex diseases are the products of modest effects of multiple genetic variants (genes and non-gene genomic regions) and have high prevalence in the population. It is also important to stress that environmental factors contribute more to the disease phenotype in complex diseases, compared with Mendelian disorders. Identifying the environmental