1

A Conceptual and Historical Framework for Molecular Epidemiology

Paul A. Schulte

We are in the era of molecular research. Between 1970 and 1990,[1] the number of medical journals with the word "molecular" in the title grew from 31 to 90, signaling that the understanding of biologic phenomena has proceeded to the molecular level. This evolution resulted from advances in molecular biology, genetics, analytical chemistry, and other basic sciences. It is now possible to detect smaller amounts of analytes and contaminants and smaller biological changes, as well as to identify mechanisms at the cellular and molecular levels. Progress in the molecular approach to biology and medicine has stimulated and excited both the public and researchers, who now believe these advances can be applied to the study, prevention, and control of health risks faced by human populations. The term "molecular epidemiology" may be used to describe such an approach: the incorporation of molecular, cellular, and other biologic measurements into epidemiologic research.

The use of molecular markers represents a quantum leap in the evolution of epidemiologic ideas. Epidemiology has evolved through development and inclusion of many advances such as the systematic collection and analysis of vital statistics; delineation of the triad of agent, host, and vector (applied in infectious and chronic diseases); refined exposure assessments such as dietary questionnaires and job exposure matrices; clearly delineated study designs (longitudinal and case-control); and heightened computational and statistical capabilities (maximum likelihood estimators, logistic and Poisson regression). To this list now must be added technologically powerful measures of biologic variables, that is, biologic markers indicating events at the physio-

[1] Comparison of National Library of Medicine journal listings for 1970 and 1990.

Molecular Epidemiology: Principles and Practices

logic, cellular, subcellular, and molecular levels. Molecular epidemiology is the use of these biologic markers in epidemiologic research. Although use of biologic markers is not new to epidemiology, the current generation of markers enhances past approaches. The use of validated biologic markers can contribute the following opportunities and capabilities to epidemiologic research:

1. delineation of a continuum of events between an exposure and a resultant disease;
2. identification of exposures to smaller amounts of xenobiotics and enhanced dose reconstruction;
3. identification of events earlier in the natural history of clinical diseases and on a smaller scale;
4. reduction of misclassification of dependent and independent variables;
5. indication of mechanisms by which an exposure and a disease are related;
6. better accounting for variability and effect modification; and
7. enhanced individual and group risk assessments.

Collectively, these capabilities provide additional tools for the epidemiologist studying questions on the etiology, prevention, and control of disease (Fig. 1.1).

Molecular epidemiology is a natural confluence of powerful developments in basic biomedical sciences and the field-tested methods of epidemiology. Although "molecular epidemiology" can be viewed as an evolutionary step in epidemiology, a supplemental set of tools, or even a separate discipline, it generally does not represent a shift in the basic paradigm of epidemiology. This new approach allows for more accurate comparisons among groups, further clarification of mechanisms, and more specialized assessment of individual risk functions, all of which have been established in historical epidemiology. Molecular epidemiology does not have all of the characteristics of a distinct discipline or even of a branch of epidemiology. Rather, it is better seen as a diverse range of approaches and techniques that can supplement the field of epidemiology and boost the field to a new level of opportunity and capability.

Capabilities of Molecular Epidemiology

Delineation of a Continuum of Events between Exposure and Disease

Figure 1.2 shows the evolution of the concept of a continuum useful in molecular epidemiologic research [Perera and Weinstein, 1982; Gann *et al.*, 1985; National Research Council (NRC) 1987; Hulka and Wilcosky, 1988;

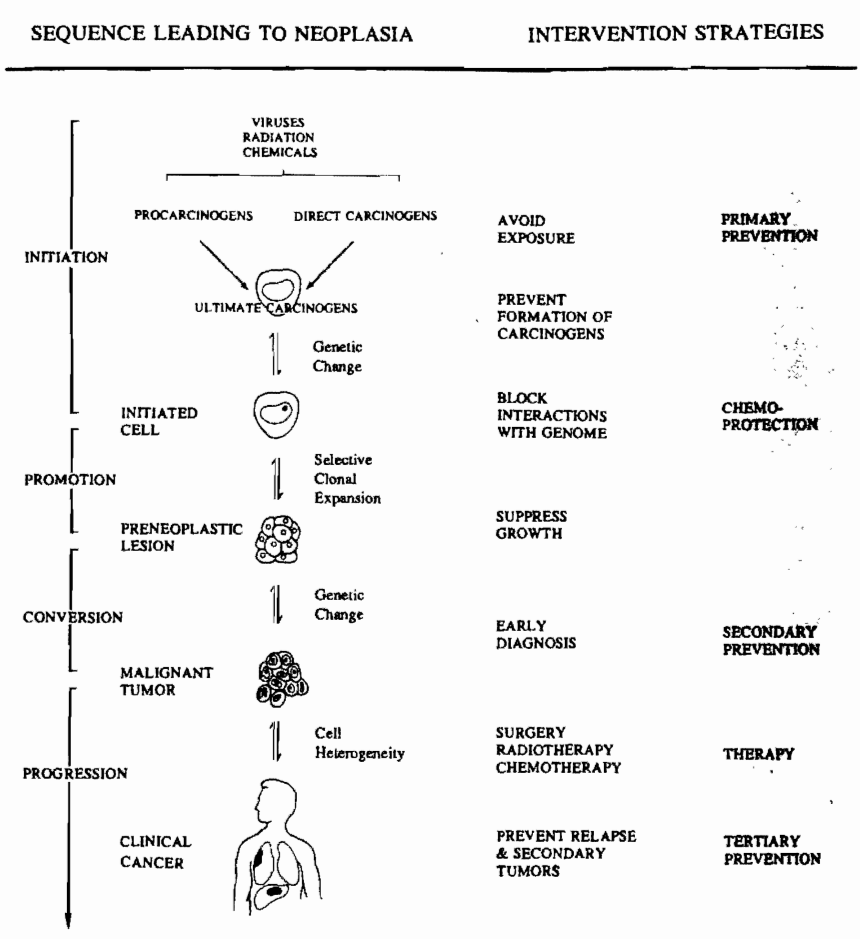

FIGURE 1.1 Intervention strategies that can utilize molecular epidemiologic methods: The example of cancer. Molecular epidemiologic approaches can be used to identify etiologic factors and to assess effectiveness of chemoprevention, secondary and tertiary prevention, and therapy. (Reprinted with permission of Thomas W. Kensler, 1992.)

Schulte, 1989]. A rich level of detail can supplement previous categorical models that linked exposure and disease. The concept of a continuum of events between exposure and disease provides the opportunity to insure that epidemiologic research has a biologic basis for hypotheses and provides the analyses to test these ideas. New opportunities for classical and hybrid epidemiologic study designs can be applied to this continuum (see Chapter 6).

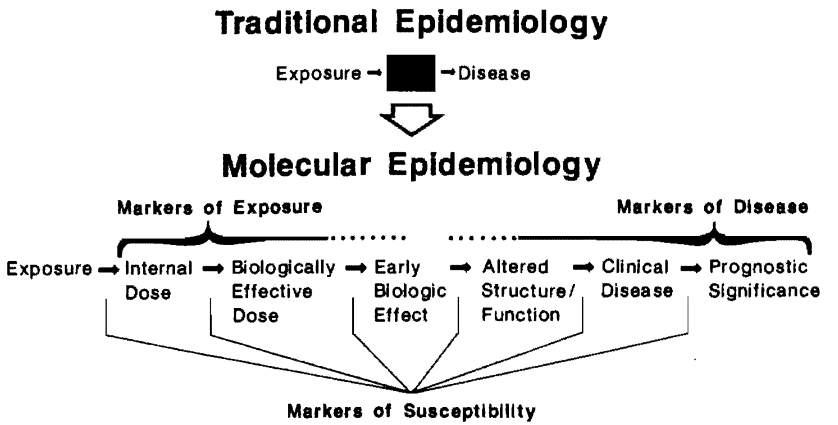

FIGURE 1.2 Evolution of the detailed continuum for molecular epidemiologic research.

Identification of Exposures to Smaller Amounts of Xenobiotics and Enhanced Dose Reconstruction

The powerful tools of molecular biology, analytical chemistry, and related disciplines now allow exposure determinations on the order of 1 part in 10^{18} or 10^{21} (Abdel-Baky and Giese, 1991; Wild and Montesano, 1991). This ability to identify small amounts of a xenobiotic makes more active consideration of "background" levels of target xenobiotics in nominally nonexposed subjects important when designing studies and assessing covariates and confounding factors.

A key capability of molecular epidemiology is assessing past exposures and reconstructing doses received from past exposures by using biologic measurements on samples taken from small groups of subjects (Ehrenberg *et al.*, 1974; Calleman *et al.*, 1978; Thilly, 1982; Wogan and Gorelick, 1989; Groopman *et al.*, 1991; Mendelsohn, 1991; Taylor *et al.*, 1992). This procedure is termed "biologic dosimetry." Biologic dosimetry can complement traditional methods of dose reconstruction by using personal dosimeters to measure ambient exposure, by estimating body burdens through sampling fat, urine, or other materials, or by detecting adducts, gene mutations, chromosome aberrations, or other relevant markers (Mendelsohn, 1991).

Identification of Events Earlier in the Natural History

As shown in Figure 1.2, when a continuum or part of a continuum between an exposure and a disease is identified and understood, it is possible to focus on preclinical rather than clinical events. Thus, asymptomatic individuals who are at increased risk of manifesting clinical disease can be identified (Hatch and Friedman-Jimenez, 1991). Some examples of indicators include decrease in CD4 lymphocytes in HIV-infected persons (Detels *et al.*, 1987),

expression of p300 in bladder cells in people at risk of bladder cancer (Rao *et al.*, 1991), elevated levels of lipoprotein Lp(a) in persons at risk for cardiovascular disease (Murai *et al.*, 1986), and various sperm parameters in individuals at risk of reduced fertility (Bostofte *et al.*, 1988).

The ability to identify prodromal events expands the pool of potential "cases" for epidemiologic studies (Hattis, 1988). It permits studies of interventions that can have impact on the group being studied as well as on the individuals to whom the results can be generalized (Cullen, 1989; Greenwald *et al.*, 1990).

Reduction of Misclassification of Variables

Misclassification of exposure and disease variables is a major weakness of epidemiologic studies (Rothman, 1986; Hogue and Brewster, 1988). Better classification of exposure than that achieved using historical characteristics and measurements may be accomplished by assessing markers of internal and biologically effective doses (Hogue and Brewster, 1988; Hulka and Wilcosky, 1988; Landrigan, 1988; Schulte, 1989). More homogeneous disease groupings can be defined using markers of effect such as specific mutations indicative of exposure (mutational spectra) (Shields and Harris, 1991). The validity and precision of point estimates may be increased as misclassifications are reduced.

Indication of Mechanisms

Delineating a continuum of events between exposure and disease provides opportunities for insight into the mechanism of action (Vogelstein *et al.*, 1988). Much epidemiologic research has been based on theorization about mechanisms, or at least some prior speculation that exposure and outcome are related. Molecular epidemiologic approaches facilitate testing the association between mechanistic events in a defined continuum (Ehrenberg, 1974; Harris *et al.*, 1987; Hatch and Stein, 1987; Perera, 1987b; Schulte, 1989; Harris C. C., 1991; Hatch and Friedman-Jimenez, 1991). Knowledge of the mechanism can guide future research and intervention applications.

Accounting for Variability and Effect Modification

Perhaps one of the greatest contributions of molecular epidemiology is the ability to discern the role of host factors, particularly genetic factors, in accounting for variation in response (Omenn, 1982; Cartwright *et al.*, 1982; Rajput-Williams *et al.*, 1988; Kuller, 1991; Shields and Harris, 1991; Wetmur *et al.*, 1991; van Noord, 1992). Why similarly exposed people do not get the same diseases is a target question for molecular epidemiology. In most disease systems, susceptibility markers are being identified and evaluated. These markers can be incorporated into epidemiologic models as effect modifiers (Hulka *et al.*, 1990).

Enhanced Individual and Group Risk Assessments

The use of epidemiologic data to provide individual and group risk assessments is well established (Paul, 1930; Truett *et al.*, 1967). For example, individual risk functions have played a strong role in cardiovascular disease research and control (Truett *et al.*, 1967), in pulmonary and occupational medicine (Ingram and McFadden, 1981), in infectious disease control (Paul, 1930), and in genetic epidemiology and counseling (Ahearn and Hochberg, 1988). Molecular epidemiology can enhance individual and group risk assessments by providing more person-specific information, allowing extrapolation of risk from one group to another, from animal species to humans, and from groups to individuals (Perera, 1987a; Harris *et al.*, 1987; Alavanja *et al.*, 1987; Harris, 1991; Shields and Harris, 1991). The "parallelogram approach" used in genetic toxicology is a model for animal-to-human extrapolation (Sobels, 1982). A marker appropriate to both species (animal and human) that can be related to exposure–disease relationships in the animal can serve as the basis for predicting effects in exposed humans. Similarly, extrapolation from group to group, group to individual, or individual to group follows the same general model (see examples in Figures 1.3 and 1.4). Identification of a detailed continuum of events between an exposure and disease, coupled with covariates of the event variables in multivariate models,

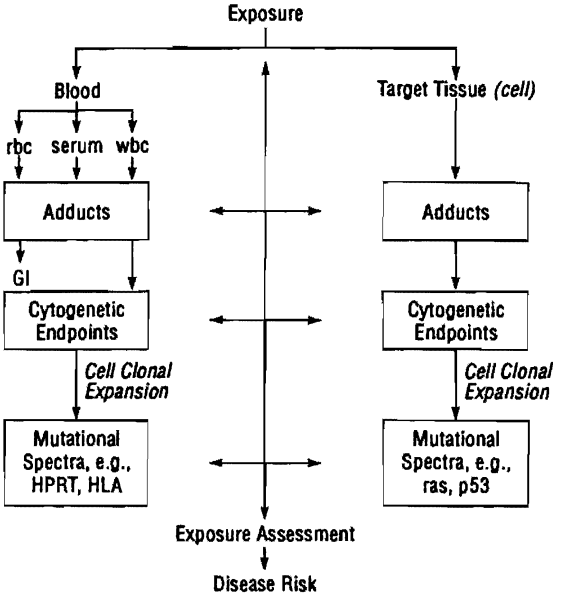

FIGURE 1.3 Exposure assessment paradigm for linking carcinogen–macromolecular adducts, cytogenetic aberrations and mutational spectra. GI, gastrointestinal tract. (Reprinted with permission from C. C. Harris, 1991, copyright CRC Press, Inc.)

permits the calculation of individual risk functions. (See Truett *et al.*, 1967, for an example using serum lipid biomarkers and cardiovascular disease risk functions.) Molecular markers can heighten the specificity of these functions and allow reduced confidence intervals around estimates. Not only is it now possible to say that a middle-aged man with heart disease and a cholesterol level above 240 mg/dl will have a one-in-five chance of dying from a heart attack within 10 years; it may soon be possible to indicate which man that will be (Begley *et al.*, 1991).

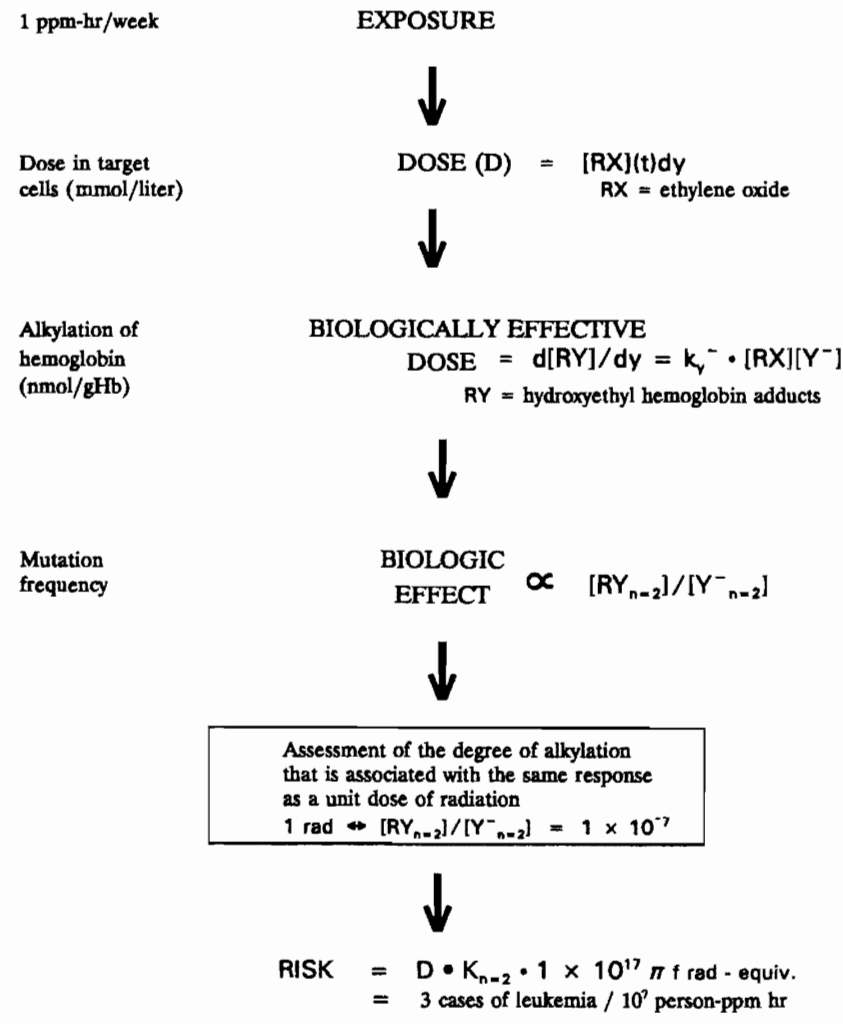

FIGURE 1.4 Example of risk assessment of workers exposed to ethylene oxide at a Swedish plant. (Adapted from Osterman-Golkar and Bergmark, 1988, and Ehrenberg *et al.*, 1980.)

Some readers may view the term "molecular epidemiology" as an oxymoron. Epidemiology is the study of health effects in groups of people. "Molecular" and "cellular" indicate assessment of the individual at the component level. Epidemiology relies on observation and inference of associations between variables. Molecular and cellular sciences use experimental proof of cause and effect. Despite these different scopes, scales, and approaches, molecular sciences and epidemiology are compatible, if not inevitably linked (Kuller, 1991). There is a historical basis for such a link, as well as a conceptual and epistemological framework for epidemiologic research that incorporates biologic measurements of human processes (Paul, 1958; Paul and White, 1973; Kilbourne, 1979; Feinstein, 1991).

In a sense, molecular epidemiology is a signpost that flags the need to incorporate an understanding of biologic phenomena at the physiologic, cellular, and molecular levels into epidemiologic research (Fig. 1.5). Epidemiologists long have used biologic markers (e.g., antibody titers, serum lipids, blood lead). However, in the past when high "exposures" and single outcomes were more prevalent and frequent, epidemiologists argued that knowledge of associations was more useful than understanding the mechanisms, since prevention through control of exposures was often feasible even in the absence of understanding cellular processes. (Snow, 1855; Maclure and MacMahon, 1980; Hatch and Stein, 1987). Previous success in public health led to the identification of the major single or primary cause of diseases. Today, exposures are often smaller and mixed; understanding mechanisms could be more important in determining appropriate intervention strategies. The health conditions of interest today are multicausal; to investigate them requires a wide array of disciplines.

If molecular epidemiology has characteristics of a field or speciality, they are its hybrid interdisciplinary qualities. This new specialty requires attention to new organizational and educational structures and adherence to principles and practices derived from both its molecular biology and epidemiology roots. Molecular epidemiology is not a fundamental departure from the past, but an evolutionary step. In this chapter, I trace the conceptual and historical development of the "field" of molecular epidemiology.

The goal of molecular epidemiology should be to supplement and integrate, not to replace, existing methods. Molecular epidemiology is also a heuristic term used to describe an enhanced capability of epidemiology to understand disease in terms of the interaction of environment and heredity. Although this practice too has been part of the epidemiologic tradition, it generally has been confined to the specialty area of "genetic epidemiology."

Researchers easily can become arrogant in the face of the potential that molecular epidemiology offers. Believing that merely making molecular measurements leads to enhanced understanding of biologic phenomena is tempting. However, this is not always true and can be misleading, because such a reductionist approach fails to pay attention to the social and cultural char-

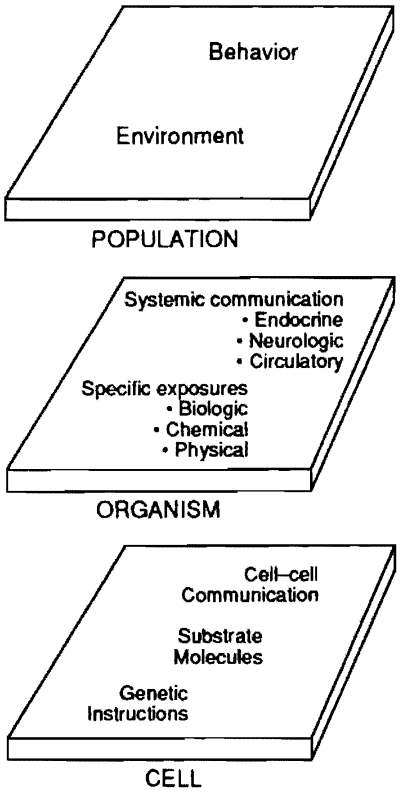

FIGURE 1.5 Component levels of molecular epidemiologic research: The example of cancer. Biologic measurements of individuals and their components, as independent and dependent variables, have a long history of use in epidemiology. The current generation of measurements can be made at the molecular level. (Figure courtesy of John D. Potter.)

acteristics of human populations and to the impact that this type of research has at the population level. Ultimately, research at both the "micro" and the "macro" level is necessary.

Strengths and Limitations of Observational Epidemiology

Epidemiology is the study of the distribution and determinants of health-related states and events in populations and the application of the results of this study to control health problems (Last, 1988). The focus of epidemiology is the group rather than the individual; understanding is gained through inferences drawn from observations within and among groups. Causation is inferred rather than proved. The strengths and weaknesses of epi-

demiology derive from the field's primary goal: identification and control of the causes of human diseases. The strength of epidemiology is that it does not require extrapolation from other species or laboratory experiments. The weakness of the field stems from ethical constraints imposed on human experimentation and, thus, the requirement to study human disease in its natural state. The epidemiologic study of "free living" humans is more problematic than the study of controlled animals. Factors such as history, geography, social characteristics, and status are exceedingly powerful predictors of the health of human populations. Whereas laboratory animals are usually genetically homogeneous and live in controlled conditions, humans are genetically heterogeneous and live in diverse conditions. In short, because of the subject, all major methods of epidemiology (except for clinical trials) are essentially observational and nonexperimental. Drawing inferences about causation from observational studies is considerably more difficult than drawing them from experiments that use random samples and controls [National Academy of Sciences (NAS), 1991b].

Much of epidemiology has been essentially ecologic (Hogue and Brewster, 1988; Hulka and Wilcosky, 1988). Often an individual is assigned the characteristics of a group; categorical descriptors are used to assess risk factors and health status. These descriptors range from dichotomous characterizations (exposed or not; diseased or not) to more quantitative representations of categories (high, medium, or low exposure). Analytic epidemiology uses measurements made on each subject or applies some specific criterion to classify subjects. In many instances, these measurements are surrogates, for example, measures of ambient air exposure for dose or measures of symptoms in lieu of underlying disease. In general, in epidemiology the group is the unit of comparison.

The epidemiologic approach has been highly successful and is the cornerstone of public health. From John Snow's (1855) assessment and control of the cholera epidemic in London to the understanding of the role of environmental factors in cancer and cardiovascular disease, epidemiology has contributed to finding causes and remedies. Despite these contributions, epidemiology has been limited by its fundamental conceptual and technical characteristics. Ecologic characterization of variables has hindered the epidemiologic approach in addressing the interdependence of multiple agents and variations in susceptibility that lead to disease. Epidemiology is also limited in its ability to classify individuals into exposure categories (misclassification), to identify mechanisms of action, and to detect disease at a time in its natural history when intervention in the study population or in other preclinical populations would be most effective (Higginson 1977b; Hogue and Brewster, 1988; Hulka and Wilcosky, 1988).

Emerging abilities to assess exposure and disease at the cellular and molecular levels are promising supplements to traditional epidemiology. Instead of characterizing groups solely by geographic location, job title, or question-

naire-derived history, it is possible to measure dose by collecting biologic specimens and assessing xenobiotic interactions with biologic sites and molecules (Schulte 1987, 1989; Rogan, 1988; Hoffman *et al.*, 1991). At the other end of the spectrum, instead of making comparisons of cases of frank disease with controls, it will be possible to make assessments based on preclinical events such as abnormal DNA content (Hemstreet *et al.*, 1988) or oncogene alteration (Shields and Harris, 1991), once these end points are established as predictive of clinical disease (Rogan, 1988). Additionally, molecular methods make it possible to distinguish subtypes of clinical disease that have potentially different etiologies (Hollstein *et al.*, 1991; Taylor *et al.*, 1992).

A long-observed weakness of epidemiology is its limited ability to address host factors that contribute to variable responses (Rose, 1990). Why similarly exposed people do not all acquire the same disease is a difficult question. Certainly, gross categories of host factors, such as age, race, and sex, are controlled routinely in epidemiologic studies. However, genetic factors generally have been considered to a lesser extent than environmental factors. Evaluation of gene–environment interactions has been minimal. "In the past epidemiologists tended to favor a 'black box' approach to their work: they have measured inputs from external agents and from susceptibility factors and disease outcomes; but they have not been concerned with explaining how the two come to be related" (Rose, 1990).

Molecular epidemiology is not just a term that describes adding new techniques to epidemiology. Rather, it represents an opportunity to use new resolving powers to develop theories of disease causation that acknowledge complex interactions involved in the health–disease process. Considering disease at the molecular level without confronting the other events, such as genetic differences and competing biochemical processes, that occur at the molecular level is not sufficient. The question addressed in this volume is how to integrate these molecular biologic capabilities—measurements made in individuals—into a science that uses comparisons of groups to find causes of disease and opportunities for health protection.

Molecular Epidemiology—The Use of Biologic Markers in Epidemiologic Research

Definitions

Events in the Continuum between Exposure and Disease

A functional definition of molecular epidemiology is the use of biologic markers or biologic measurements in epidemiologic research. Biologic markers (or biomarkers) generally include biochemical, molecular, genetic, immunologic, or physiologic signals of events in biologic systems (NRC, 1987). The events represented can be depicted as parts of a continuum between a

causal initiating event (sometimes an exposure to a xenobiotic substance) and resultant disease. By definition, a continuum is a whole, no part of which can be distinguished from neighboring parts except by arbitrary divisions (The American Heritage Dictionary, 1976). Thus, it is important to remember that the various heuristic components of the continuum shown in Fig. 1.2 are arbitrary.

The proposed continuum between exposure and disease has been described in a number of reports (Perera and Weinstein, 1982; Hatch and Stein, 1987; NRC, 1987; Perera 1987a,b; Schulte, 1989) and is shown in Figure 1.2. Between *exposure* (E) in the environment and the development of *clinical disease* (CD), four generic component classes of biologic markers have been identified: the *internal dose* (ID), the *biologically effective dose* (BED), *early biologic effects* (EBE), and *altered structure and function* (ASF). Clinical disease can be represented not only by biologic markers for the current disease but also by markers for *prognostic significance* (PS). Each marker represents an event in the continuum. The relationships among the markers are influenced by various factors (such as genetic or other host characteristics) that reflect susceptibility to any of the events in the continuum. These indicators for susceptibility also can be represented by markers.

Definition of all the marker events has been elaborated elsewhere (NRC, 1987; Hulka and Wilcosky, 1988) but is summarized briefly here. The continuum between cigarette smoking and lung cancer serves as illustration. ID is the amount of a xenobiotic substance or its metabolites found in a biologic medium (e.g., serum cotinine as an indicator of nicotine). The BED is the amount of that xenobiotic material that interacts with critical subcellular, cellular, and tissue targets, or with an established surrogate tissue (e.g., DNA adducts in peripheral lymphocytes). The BED represents the integration of exposure and effect modification by the host. A marker EBE (e.g., sister chromatid exchange) represents an event correlated with, and possibly predictive of, health impairment. Altered structure or function (e.g., abnormal sputum cytology) and DNA hyperploidy are precursor biologic changes that are more closely related to the development of disease. Markers of CD (e.g., tumor-associated antigen) and of PS (e.g., tumor markers such as CA-125) show the presence or future development of disease, respectively. Markers of susceptibility are indicators of increased (or decreased) risk for any component in the continuum [e.g., extensive debrisoquine metabolizers are at 4- to 6-fold increased risk for lung cancer (Ayesh *et al.*, 1984)].

Relationship between a Marker and the Event It Marks

When considering how biologic markers can be used in epidemiologic research, it is useful to reflect on the nature of the relationship between the marker and the event it marks (Lucier and Thompson, 1987). The semantics of describing a marker in this regard are confusing. Does the marker represent an event, is it an event itself, is it a correlate of the event, or is it a predic-

tor of the event? The answers to these questions may affect who is sampled, how and when they are sampled, and what confounders or effect modifiers are considered. For example, the HPRT gene mutation may be used as a surrogate for other target genes (such as mutated p53 gene) that may be involved in the development of cancer; a hemoglobin adduct, although highly correlated with recent exposures to aklylating agents such as ethylene oxide, may not represent exposures that occur years prior to the collection of a blood specimen, since the dosimetric capacity of the hemoglobin is related to the four-month life of the erythrocyte. Thus, a biologic marker often refers to the use made of a piece of biologic information rather than to a specific type of information (Henderson *et al.*, 1989).

Is a marker different from a test or an assay? Often, biologic markers and tests or assays for a marker are considered the same because, without the assay, the marker cannot be demonstrated. Strictly speaking, they are different and care should be taken not to gloss over the differences. Analytically, a test or assay is said to be valid if it performs "truthfully" in the presence or absence of a marker. The attributes of a test are not necessarily those of the marker. The marker's attributes may pertain to its nature and natural history. A marker may exist although no assay is sensitive enough to detect it. The measurements of a marker involves differentiation of a signal above background noise. A test that has a large signal-to-noise ratio is considered a good test. The signal-to-noise ratio is not generally a set point, but a curve showing the ratio under different conditions. Ultimately, characteristics of the test conditions, the marker, or both may determine the signal-to-noise ratio. These issues are discussed further in Chapters 3 and 4.

Organizational Aspects

Molecular epidemiologic studies require interdisciplinary collaboration between population and field scientists (such as epidemiologists, statisticians, industrial hygienists, exposure assessors, and clinicians) and laboratory scientists from disciplines such as molecular biology, genetics, immunology, biochemistry, pathology, and clinical and analytical chemistry. Interdisciplinary collaboration is not new to epidemiology, but the level and extent of diversity of disciplines that molecular epidemiology will require is unprecedented.

Collaboration requires attention to the underlying assumptions, paradigms, and languages of various disciplines as well as to issues of the institutional context of research (Stein and Jessop, 1988). Every discipline uses assumptions and paradigms to approach a research question. Often these conventions are so fundamental and integrated that investigators may not be conscious of them and, hence, rarely recognize them or make them explicit when interacting with members of other disciplines (Stein and Jessop, 1988). Problems in collaborative research can occur when these fundamentals are

not shared. Epidemiologists generally speak in terms of groups and risks to groups. Laboratory scientists tend to focus on individuals or components of an individual. Different disciplines may use common terms in very different ways. Words such as "sensitivity," "small," "valid," "normal," "bias," and "epidemiology" may have different meanings in various disciplines. Too often, the differences may be subtle, for example, understanding of the term "sensitivity." To the laboratory scientist, it refers to the extent to which an assay is capable of detecting a particular marker at some concentration. To an epidemiologist, the same word means the extent to which people with a marker respond positively on a test.

In addition to conceptual and language barriers, interdisciplinary collaboration often may have administrative barriers (Feinstein, 1991). If a study involves disciplines in two or more academic or organizational departments, several questions might arise that must be answered clearly. Which department is in charge? On whose books is the project accounted? To which journals are papers submitted? Who can access which data? Who is the first author? What is the priority for additional assays on specimens? Many administrative difficulties can be avoided if good communication is established from the outset. Effective collaborative work requires mutual commitment to the value of the work, respect for what each participant can offer, intellectual flexibility, tact, patience, and persistence (Stein and Jessop, 1988).

Protocols for molecular epidemiologic studies often differ from traditional epidemiology, not only in the interdisciplinary nature of the research, but also in the need or desire to maintain flexibility to amend the protocol or study to include discovery of new markers. This variability is best handled by anticipating such possibilities, by including references to such changes in informed consent documents, and by processing and storing samples appropriately.

Crucial to the use of biologic markers in human populations is the requirement that biologic markers be validated adequately before their deployment in hypothesis testing studies (Rogan, 1988). The term "validated" not only has different meanings for the laboratory and population scientists, but also leads to different approaches to validation (see Chapters 3 and 4). Therefore, consistent guidelines for validating markers at the laboratory and the field levels are essential. Laboratory validation requires assessing the adequacy and accuracy of assays. Field validation requires determining the population variability and predictive value of the assay in terms of the exposure or disease of interest.

An organizational model for conducting molecular epidemiologic studies is shown in Figure 1.6. Critical in these types of studies are the need for specialists in study design and for subject selection, exposure assessment, and marker assay development to be integrated in all aspects of the work. The amount of data generated is generally much greater than in a classic epidemiologic study because, in addition to questionnaires and recorded data common to both, data are accumulated from biologic specimen analy-

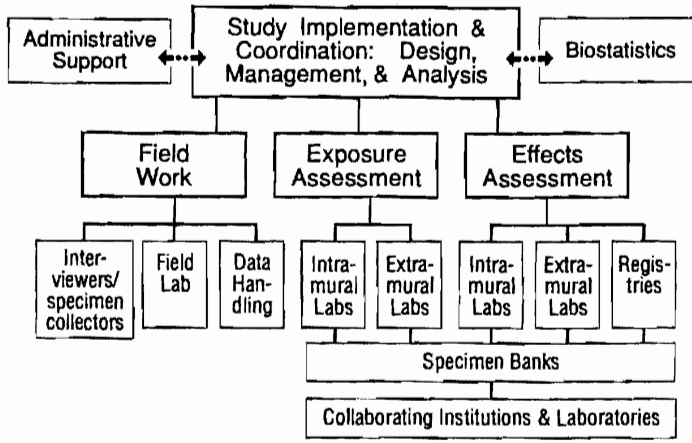

FIGURE 1.6 Project teams required for conducting interdisciplinary molecular epidemiologic studies. (Adapted from Everson, 1989.)

sis. For example, a study of 30 mortuary science students that assessed cytogenetic effects and DNA repair caused by formaldehyde exposure required 1800 slides and generated more than 3600 data points (Suruda *et al.*, 1992). Some of these data came from questionnaires and some from biologic measurements.

The organizational model in Figure 1.6 pertains to studies for validating or using biologic markers. Validation studies require a similar effort as field studies but contain certain differences that may involve animal or *in vitro* research. Validation often requires large, prospective population-based studies that are costly, time consuming, and resource intensive. Prior to such studies, small-scale pilot studies should be made. Pilot studies can provide estimates of assay variability and interindividual variability that are necessary for power calculations. The field of molecular epidemiology has been characterized by pilot studies in highly exposed or high-risk groups or in patient series. Based on the findings of these studies, larger studies in general groups are considered. (See Chapter 6 for a typology of molecular epidemiologic studies.)

Many problems are attendant to developing a marker assay for use in epidemiologic studies. One problem is the difficulty of transforming research tools and research laboratories into productive facilities capable of handling large numbers of specimens in a cost-effective and timely manner. Other challenges include testing for dose response, marker persistence, correlation with other markers, inter- and intraperson variation, and correlation with clinical responses.

As new and more reliable assays are developed, scientists will be pressured to use them to replace previous assays. When a marker assay is new, measurements may differ from those in previous assays for the same marker.

Although it may be reasonable to wait until laboratory techniques and estimates of variability display consistency, it is not reasonable or feasible to wait until a technique is so standard that no refinement of it is likely to occur. Pressure to use the new technique also promotes assessment of comparability of old and new results and, ultimately, making adjustments or corrections. A good balance must be achieved between old and new approaches. In some cases, employing both techniques in a study may be useful.

Molecular epidemiology also requires paying attention to the collection, handling, and storage of biologic specimens (see Chapter 8). Specimens must be stored to be usable for analyses not yet developed in addition to currently known analyses. The actual collection of biologic specimens also may require attention to timing, so specimens are collected when the influence of factors such as therapeutic treatments will be minimal. For example, a study of the role of oncogenes and colon cancer will be more informative if specimens are collected before DNA-altering chemotherapy or radiation is administered. Other "new" issues also arise, such as how to address the difficulty of obtaining biologic specimens from controls. For example, in a case-control study of colon cancer, how can usable colon specimens be obtained from controls?

Use of markers of susceptibility also may have a great impact on the organization and conduct of molecular epidemiologic research. Biologic monitoring data derived from molecular epidemiologic studies are likely to be sought by a diversity of societal groups ranging from insurers and employers to potential mates. The possibility that individual markers can predict future characteristics and outcomes makes them a target of interest. The use of susceptibility markers will be required increasingly in molecular epidemiologic studies that assess exposure or effects at the DNA level, since effect modification or confounding caused by inherited or other nonstudy factors can influence associations under study. The analysis of biologic marker data may create new obligations for researchers that were not present in traditional epidemiology (Schulte, 1991). What is the responsibility for follow-up of "abnormal" results, how will individuals with these results be treated, and what will be done if other pathologic conditions are observed (see Chapter 9)?

The use of biologic markers of effect that permit focus on and detection of preclinical and extremely early disease also raises questions about the true preventability and treatability of these conditions (Yach, 1990). At present, insufficient information is available to determine whether particular levels of a certain biologic marker reflect normal ranges of that marker in a person over time or whether they reflect early stages of a preventable disease (Yach, 1990). Researchers currently face this challenge.

Finally, from an organizational viewpoint, epidemiologists may potentially be polarized into two worlds: one of molecular epidemiologists who emphasize the molecular and genetic causes of disease and the other of social epidemiologists who stress the role of social, psychological, and economic

factors in health. Neither approach alone will satisfactorily address the health issues of the current era. History has shown that complete reliance on reductionist approaches is antithetical to public health, yet failure to use the powerful tools available also will not safeguard public health. A synthesis of the two approaches is needed that can address the entire scope of health issues (Yach, 1990).

Historic Contributions to Molecular Epidemiology

If molecular epidemiology can be described as the use of biologic markers and measurements in epidemiologic research, then the historic contributions to this approach come from those disciplines that have made advances in relating biologic measurements to health and disease. The principles and practices discussed in this volume and other publications on the topic (Hulka *et al.*, 1990; Gledhill and Mauro, 1990; Garner *et al.*, 1991; Groopman and Skipper, 1991) build on a rich and diverse history of biologic measurements. Many of the current troubling issues facing molecular epidemiology were encountered when the tools leading to molecular epidemiology were developed. Although not entirely exhaustive, the list of some disciplines that have contributed to molecular epidemiology include bacteriology, immunology, and infectious disease epidemiology; pathology and clinical chemistry; carcinogenesis and oncology; occupational medicine and toxicology; cardiovascular epidemiology; genetics, molecular biology, and genetic epidemiology; and traditional epidemiology and biostatistics. In these disciplines, the techniques and building blocks of molecular epidemiology have their history.

Bacteriology, Immunology, and Infectious Disease Epidemiology

The first consideration of cellular biomarkers in medical research appeared in the study of infectious disease. Even before the development of the microscope, Italian physician Girolamo Fracastoro wrote in 1546 that the "seeds" or germs of contagious diseases were carried from person to person (cited in Clendening, 1942). This speculation later was confirmed, after the development of the microscope by Van Leeuwenhoek (1632–1723) and the conclusion by Schwann in 1839 that the cell was the fundamental unit of living matter (cited in Venzmer, 1968). Subsequently, Pasteur, Koch, Gram, Von Pettenkofer, and others in the late 1800s began to isolate specific organisms responsible for disease. The detection of bacteria in biologic specimens indicated exposure or disease (depending on the study). Immunology was built on the use of biologic markers indicative of exposure, effect, or susceptibility. In contrast to bacteriology, in which bacteria were "markers for themselves," immunologists used indirect markers indicative of infection, for example, white cell counts and antibody titers. In the 1920s, epidemiologists found a

way to use these biomarkers to address the difficult problem of measuring degrees of susceptibility and resistance to disease in the human host (Paul and White, 1973). Before this time, resistance and susceptibility were identified by asking the patient or family whether he had experienced various common contagious diseases. However, with advances such as the Schick and tuberculin skin tests, a method was found that made use of a biologic marker to assess exposure and susceptibility in individuals or groups.

Frost's (1928) work in New York City and Baltimore was among the earliest to use biomarkers to identify age-specific immunity patterns. The use of serologic tests, such as the Wassermann test for syphilis in segments of urban populations in Baltimore in the 1920s, signaled the beginning of one of the earliest precursors to molecular epidemiology, "seroepidemiology" (Williams, 1920). A classic example was the work of Aycock and Kramer (1930), which showed a rural–urban difference in both diphtheria and poliomyelitis using skin and serum biomarkers, leading to the conclusion that both infections were spread by close human contact. The term "serological epidemiology" is attributed to John Paul in 1935, who was instrumental in the development and use of serum surveys in epidemiology (White, 1973). Paul identified issues and applications of seriologic epidemiology that foreshadowed what accompanies the current problems and issues in molecular epidemiology. These issues include attention to selecting an appropriate sample group in a population; determination of sample size; storage and handling of specimens; and development of longitudinal studies.

Other contributions to molecular epidemiology are assays that take advantage of immunologic binding and specificity. The introduction of the radioimmunoassay (Yalow and Berson, 1960) is considered one of the most important advances in biologic measurement (Chard, 1990). The immunoassay is a major tool of molecular epidemiology. Much of the work detecting the interaction of xenobiotics and macromolecules such as DNA and RNA involves immunoassay methods. This type of assessment has become known as "molecular dosimetry" (Perera, 1987a) or "biological dosimetry" (Mendelsohn, 1991). The approach seeks to focus on the biologically effective dose by measuring the bound xenobiotics.

The term "molecular epidemiology" may have been used first in modern infectious disease research. In the early 1970s, the term appeared in a number of papers, Kilbourne (1973) published *The Molecular Epidemiology of Influenza.* Subsequently, occasional papers were published in the infectious disease literature; among the earliest were Pereira *et al.* (1976), Kilbourne (1979), Pappenheimer and Murphy (1983), and Follett (1984).

Kilbourne described the approach as involving molecular determinants of epidemiologic events and used influenza to illustrate the approach. (See Figure 1.7 for the level of detail and specificity the molecular approach provided.) The rationale for such an approach also has been described by Balayeva in relation to rickettsiae: "Conventional serological and biological tech-

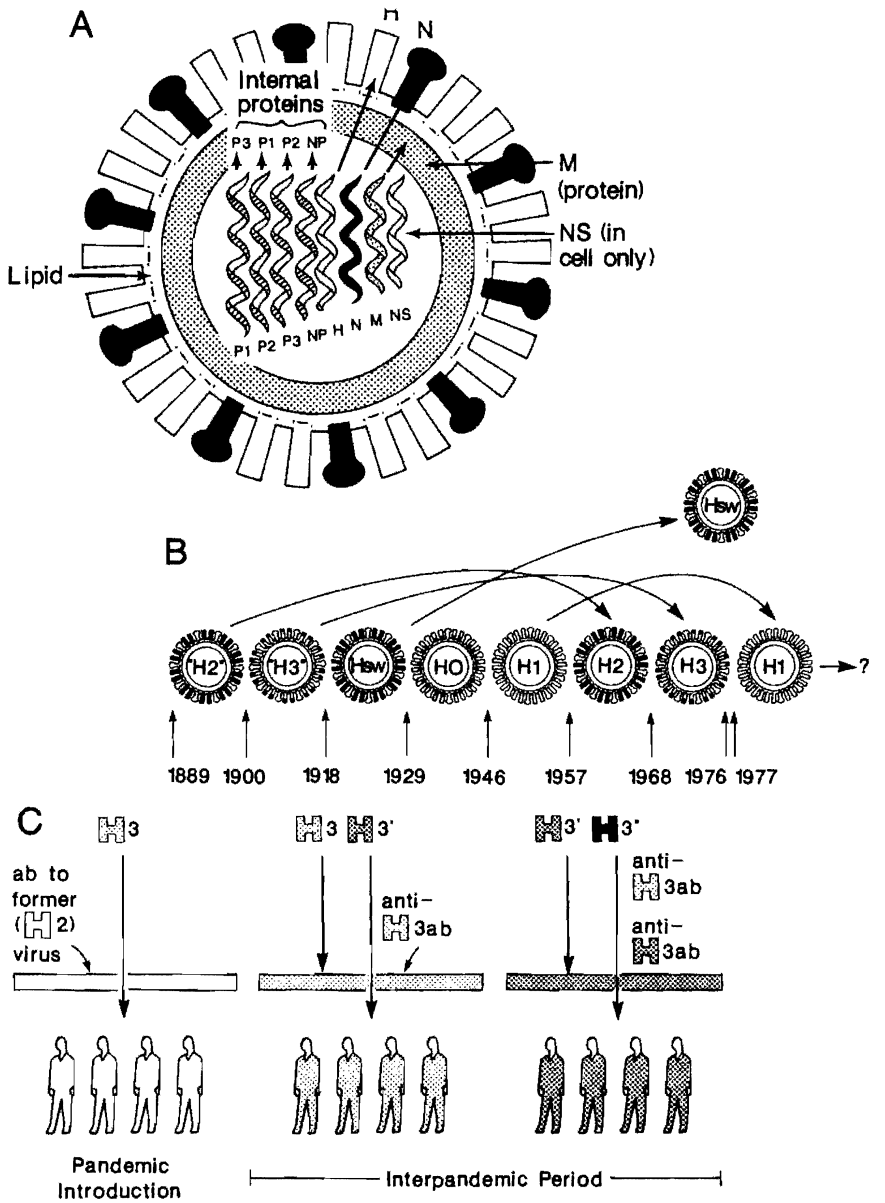

FIGURE 1.7 Molecular determinants of epidemiologic events. (With permission from Kilbourne, 1979.) (A) Schematic of influenza virion, indicating the 8 viral genes and their polypeptide products. Relative sizes of the RNA genes are only approximate; structural details of their relationship to one another and to the internal proteins P1–P3 and NP are not known. H, hemagglutinin glycoprotein; N, neuraminidase glycoprotein. (B) Influenza pandemics of the past century. Schematic of periods of prevalence of major hemagglutinin subtypes of influenza A viruses. Year of introduction of "new" subtypes indicated by arrows. H0, H1, etc., refer to hemagglutinin subtype of the viruses without reference to their neuraminidase antigens. Subtypes Hsw, H0, and H1 are more closely related antigenically than they are to H2 or H3. Identification of viral subtypes prior to 1933 is inferential on the basis of studies of human antibodies. (C) Selection of antigenic mutants are as a function of population antibody. New pandemic viral subtype H3 transcends barrier of antibody to unrelated previously prevalent virus H2 and readily infects the population. When H3 infects a critical percentage of the population its survival is impeded, and antigenically changed mutant H3′, and later H3″, have survival advantage (minor and antigenic variation or antigenic "drift").

niques for characterizing rickettsiae do not seem to be capable of providing answers to all the ecological and epidemiological questions surrounding these organisms. New techniques for identification and epidemiological assessment are required" (Balayeva, 1989).

Pathology and Clinical Chemistry

A century ago, Virchow (1821–1902) foreshadowed what contemporary molecular epidemiology could be by relating clinical disease to cellular pathology (Virchow, 1858). Subsequently, understanding of the cellular basis for disease grew and cellular changes were characterized according to whether they indicated homeostasis, normal variations, or pathogenesis.

Many of the basic principles for the practice of molecular epidemiology are derived from the discipline of clinical chemistry, which comprises the development and performance of specific chemical analyses for diagnostic purposes. Clinical chemistry cannot be sharply separated from the related sciences of pathology, hematology, immunology, and bacteriology (Richterich, 1969). Modern clinical chemistry can be traced back to 1850, when measurements of serum electrolytes were used to distinguish cholera patients from healthy individuals during an epidemic (Schmidt, 1850). However, at that time these measurements were not considered to be of any pathogenic, diagnostic, or therapeutic value. More than 50 years elapsed before electrolyte measurements were considered significant in the field of pediatrics (Richterich, 1969). As clinical chemistry evolved, the fundamental principles of "range of normal," individual variation, sensitivity, and specificity of a test were developed. These principles are part of the foundation of molecular epidemiology. To this end, the classic work of Vecchio (1966), *Predictive value of a single diagnostic test in unselected populations*, and that of Galen and Gambino (1975), *Beyond Normality: The Predictive Value and Efficiency of Medical Diagnoses*, are significant because they provide a quantitative approach to assessing validity. Interestingly, the development of the work of Galen and Gambino, two clinical pathologists, was influenced by the epidemiologist Mervyn Susser.[2] Susser presented the concepts of sensitivity, specificity, and predictive value; Vecchio, and Galen and Gambino applied these concepts to clinical laboratory data.

Quantitative methods in pathology have a long history that contributes directly to molecular epidemiology. In the late 1800s and early 1900s, the many hypotheses of cancer causation relied on identifying quantitative changes in chromosomes in cancer cells (Baak and Tosi, 1991). The beginning of modern cytometry often is dated to the early 1930s when Caspersson and colleagues showed that proteins and nucleic acids in epithelial tumor cells differed from those in normal cells (see Caspersson and Zech, 1972).

[2] In 1972, Galen attended an epidemiology seminar led by Professor Mervyn Susser, who described the terms "sensitivity," "specific," and "predictive value."

Automation and quantitative analysis of DNA, cell surface receptors, and other markers is now a major tool in many branches of medicine (Anderson *et al.*, 1984; Koss, 1987).

Carcinogenesis and Oncology

Environmental Carcinogenesis

The study of environmental carcinogenesis, whether caused by chemical, biologic, or physical agents, has contributed some of the broad concepts that form the foundation of molecular epidemiology. The understanding of cancer as a multistage process (Berenblum and Shubik, 1949) and the efforts to find tumor markers provided the background for the concept of a continuum of events between a xenobiotic exposure and a resultant disease. The two-stage and multistage models of chemical carcinogenesis were based on the work of Rous and Kidd (1941) and expanded by Berenblum and colleagues (1949). These models provided a framework within which to explain the temporal continuum between a chemical exposure and development of a cancer. Such models, coupled with an increased understanding of the natural history of cancer, presage current thinking about cancer biomarkers.

In 1958, Foulds described the natural history of cancer by suggesting that the biologic properties and behaviors of neoplastic cells during progression are determined by numerous "unit characteristics" such as growth rate, histologic type, responsiveness to hormones, and invasiveness. Beginning in the late 1960s, Ehrenberg (1974), Calleman *et al.* (1978), Ostermann-Golkar (1983), and colleagues have contributed to an understanding of the components of the continuum between exposure to electrophilic carcinogens and cancer. These investigators have demonstrated how the dose of genotoxic agents can be assessed by measuring the products of their covalent bonds to hemoglobin and DNA, a process known as molecular or biologic dosimetry. Subsequently, the researchers compared the alkylation effect of a dose of a genotoxic chemical to the known alkylation effect of radiation and using these "rad equivalents," performed risk assessments for leukemia in workers exposed to those chemicals.

Other insights into the continuum of events between exposure and disease came from the field of radiation carcinogenesis. In this area, indicators of dose and markers of early biologic effects were used to predict health outcomes (Beebe *et al.*, 1978). The concepts of "internal dose" and "biologically effective dose") were developed in the area of radiation carcinogenesis (Gledhill and Mauro, 1990). The assessment of biologic effects, specifically mutations, in animal and human populations exposed to radiation used biologic markers of metabolic phenotypes, karyotypic differences, and gene rearrangements. The term "mutational epidemiology" was applied to these uses of biomarkers, first with radiation and later with mutagenic chemicals (Miller, 1983).

More recent history of molecular epidemiology can be traced to efforts in the 1970s and 1980s to address environmental carcinogensis. Wynder and Reddy (1974) are credited with using the term "metabolic epidemiology" to describe the incorporation of measurements of bile acids and cholesterol into epidemiologic studies to assess the role of diet in colon cancer. Later, the term "biochemical epidemiology" was used to describe testing of the molecular hypotheses at the clinical and population levels (Harris *et al.*, 1985). For example, whereas it was once common to identify individuals at high risk of cancer due to lifestyle or occupation, the susceptible individual now may be identified at the molecular level for some cancers. Outside the realm of infectious disease, the use of the term "molecular epidemiology" has been ascribed to Higginson (1977) in his paper, "The role of the pathologist in environmental research and public health." However, seminal thinking in this area was found in the work of Lower (Lower *et al.*, 1979; and Lower, 1982) and Perera and Weinstein (1982). Lower's work addressed the interaction between exposure to a carcinogen and metabolic phenotypes. He described an interdisciplinary approach to molecular epidemiologic research. The work of Perera and Weinstein (1982) demonstrated in detail the potential for the use of biologic markers of dose to provide a more accurate appraisal of exposure than ecologic descriptors can. Perera and Weinstein (1982) first used the term "molecular cancer epidemiology" to describe this approach. They discussed carcinogen–DNA adducts measured by highly sensitive antibodies as examples of promising markers in studies of environmental carcinogenesis. They concluded that adduct formation ultimately might be useful in identifying individuals at high risk of cancer.

Tumor Markers

One of the clearest historical precursors of the use of biologic markers was the quest to discover cancer at an early and intervenable stage. In the early 1930s, Zondek (1942) considered using human chorionic gonadotropin in body fluids to diagnose tumors of gestational and germ cell origin. The term "markers," as in "tumor markers," has been used widely in oncology and cancer research. The search for a tumor marker or test that would detect all or most forms of cancer was based on the supposition that neoplasia in all its many forms results in a unique change, either in a component of body fluids or in a host response phenomenon (Bagshawe, 1983). No such marker has been found. However, numerous markers are used routinely as part of the process to diagnose or confirm a diagnosis for specific cancer types. These include, for example, carcinoembryonic antigen (CEA) for gastrointestinal tumors, serum acid phosphatase for carcinoma of the prostate, 5-hydroxy indoleacetic acid in the urine for carcinoid tumors, α-fetoprotein for liver cancer, and thyrocalcitonin for modular carcinoma of the thyroid (Ghosh and Rob, 1987; Klee and Go, 1987). The introduction in 1975 of monoclonal antibody technology revolutionized serologic and biochemical analy-

sis by allowing the development of sensitive and specific probes of human cancer (Wright and Cox, 1987). Tumor markers have been considered for use in screening, diagnosing, staging, and monitoring treatment and recurrence of disease and in determining the efficacy of specific forms of therapy. From an epidemiologic viewpoint, most of the studies of tumor markers have focused on efforts to validate whether a marker indicates a diagnosis of cancer and whether the predictive value of a positive test was sufficient to render a cost-effective method for screening the general population.

A great lesson for molecular epidemiologists can be learned from the history of the quest for tumor markers. As one reviewer noted, ". . . it is not uncommon for scientists to tumble on some phenomenon that appears to distinguish patients with cancer from other people and before long the scientist is liable to be carried along on a wave of blind conviction and messianic fervour" (Bagshawe, 1983). The lesson is that some of the papers that report on a "cancer test" provide inadequate technical information; early studies may have included researchers not blinded to the ultimate disease status of the subject or a host of other factors that can fool an investigator. Critical evaluation of the initial milieu in which a marker was identified and developed is important.

The history of tumor marker research also provides excellent examples of past attempts to validate a marker and use it in a screening program. Some of the best examples of early tumor markers have been developed in the fields of cancer cytology and cytogenetics. Since the conceptualization of the cell theory by Schwann (1839), researchers have been on a quest for the basic cellular markers of disease. Cellular changes have been the primary focus of study for many diseases, particularly cancer. The very definition of cancer is based on characteristics indicative of loss of cellular growth control. As refinements in technology progressed, subcellular markers such as chromosomal changes that could be linked to human cancers were found. In 1972, the 9,22 translocation was identified in chronic myeloid leukemia (Rowley, 1973). To date, at least 70 recurring translocations have been detected in human malignant cells (Rowley, 1990). Using chromosome aberrations as markers, genes relevant to the process of malignant transformations have been located. Ultimately, it has been feasible to measure DNA bases, which are key to the carcinogenic process. Measurement of the DNA characteristics in cells build on the work of Feulgen and Rossenbeck (1924), who described the stain that subsequently proved specific for double-stranded DNA (Koss, 1990). In the 1930s, Caspersson and colleagues began measuring the fluorescent intensity of DNA and RNA in cells (Koss, 1990; Baak and Tosi, 1991). This work was part of the foundation of quantitative methods in pathology that foreshadowed the measurement of biologic changes as variables in epidemiology.

The use of Papanicolaou cytology as a marker of preclinical cervical cancer demonstrates how a good marker can lead to effective intervention when

a disease is in a treatable state. The work of Papanicolaou and Traut (1943) in staining and classifying cells led to this important marker of cancer diagnosis. The history of the Pap test provides a lesson in how a new biomarker may not be used in a timely fashion. Approximately 27 years lapsed between the development and the adoption of the Pap test. Greenwald (Greenwald *et al.*, 1990) identified three main reasons for this delay: (1) failure to recognize and pursue the potential of cervical cytology, (2) lack of clinical trials to verify the test's effectiveness, and (3) failure to use screening in those population groups at highest risk.

Mutational Spectra

The location and type of mutations in a specific sequence of nucleotides define a mutational spectrum (Hollstein *et al.*, 1991). The ability to determine mutational spectra has been demonstrated in genetic toxicology and cancer research and may prove to be a major component to molecular epidemiology. The discovery of chemical-induced chromosome mutations first occurred in studies of plants (Oehlkers, 1943; Basler, 1987). More detailed analysis of mutations occurred when Benzer (1961) first demonstrated the sequence specificity of spontaneous and induced mutations in the *rII* locus of the bacteriophage T4. Two key observations were made: (1) For a given agent, several sites in the gene displayed higher mutation frequencies than others. (2) The distribution of mutations as a function of base-pair position was different for spontaneous mutations and for mutations induced by a wide variety of chemicals (Thilly *et al.*, 1982).

From this and other research we have learned that spontaneous and induced mutations are not distributed randomly with respect to type and position in the genome. Various researchers, including Albertini (1982) in the 1970s and early 1980s, developed assays on T lymphocytes as practical approaches to monitoring human mutagenicity. Thilly *et al.* (1982) demonstrated an approach to distinguishing between spontaneous and induced mutations using specific forward mutation assays, each of which detects only missense mutations at a small number of base pairs in the human genome. In the late 1970s and early 1980s, the term "mutational epidemiology" began to be used to describe surveillance and monitoring of individuals exposed to known and suspected mutagens (Hook, 1982). The phrase also encompassed etiological epidemiologic studies (Miller, 1983).

In 1991, Hollstein *et al.* focused on the pattern of base substitution mutations in the *p53* gene observed in human cancers. Different malignancies showed different patterns of mutations. By comparing the mutation spectra at the same locus, for a variety of tumors of different proposed etiology, it may be possible to distinguish the etiological contributions of exogenous and endogenous factors to human carcinogenesis. (See Figure 1.8 for an illustration of specific mutations for various cancers.) For example, for liver tumors

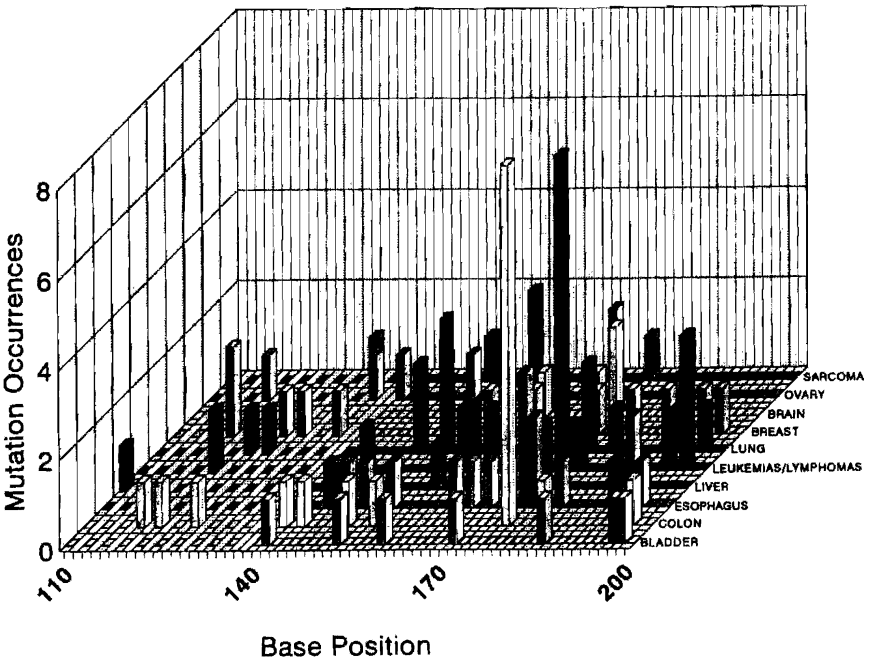

FIGURE 1.8 Localization of p53 base substitutions in human cancers. Only a section (codons 110–200) of p53 is shown to illustrate this mutational spectra. Adapted from Hollstein *et al.*, 1991. In many instances, analysis of the p53 gene in tumors was limited to exons 5 through 8, corresponding to residues 126 through 306. Five of the 280 mutations were found outside the 200-codon stretch presented here, four of which resulted in chain-terminating codons.

in persons from geographic areas in which aflatoxin B is a cancer risk factor, most mutations of the *p53* gene are at one nucleotide pair (G to T) of codon 249 (Bressac *et al.*, 1991; Hsu, 1991), suggesting a major role for aflatoxin B. The findings in mitogenesis experiments, in which aflatoxin B was administered to rodents and found to induce G to T transversions, suggest an ability to distinguish cancer caused by aflatoxin exposure and that caused by other factors (Muench *et al.*, 1983 McMahon, *et al.*, 1990). Chemical or physical agents are believed to produce a spectrum of mutational events that may be specific for each exposure, creating a mutational fingerprint (See Figure 1.9).

Occupational Medicine and Toxicology

The study of occupational disease and toxicology is another historical root of molecular epidemiology. Since the early part of the twentieth century, occupational exposures have been confirmed by biologic monitoring (Hamil-

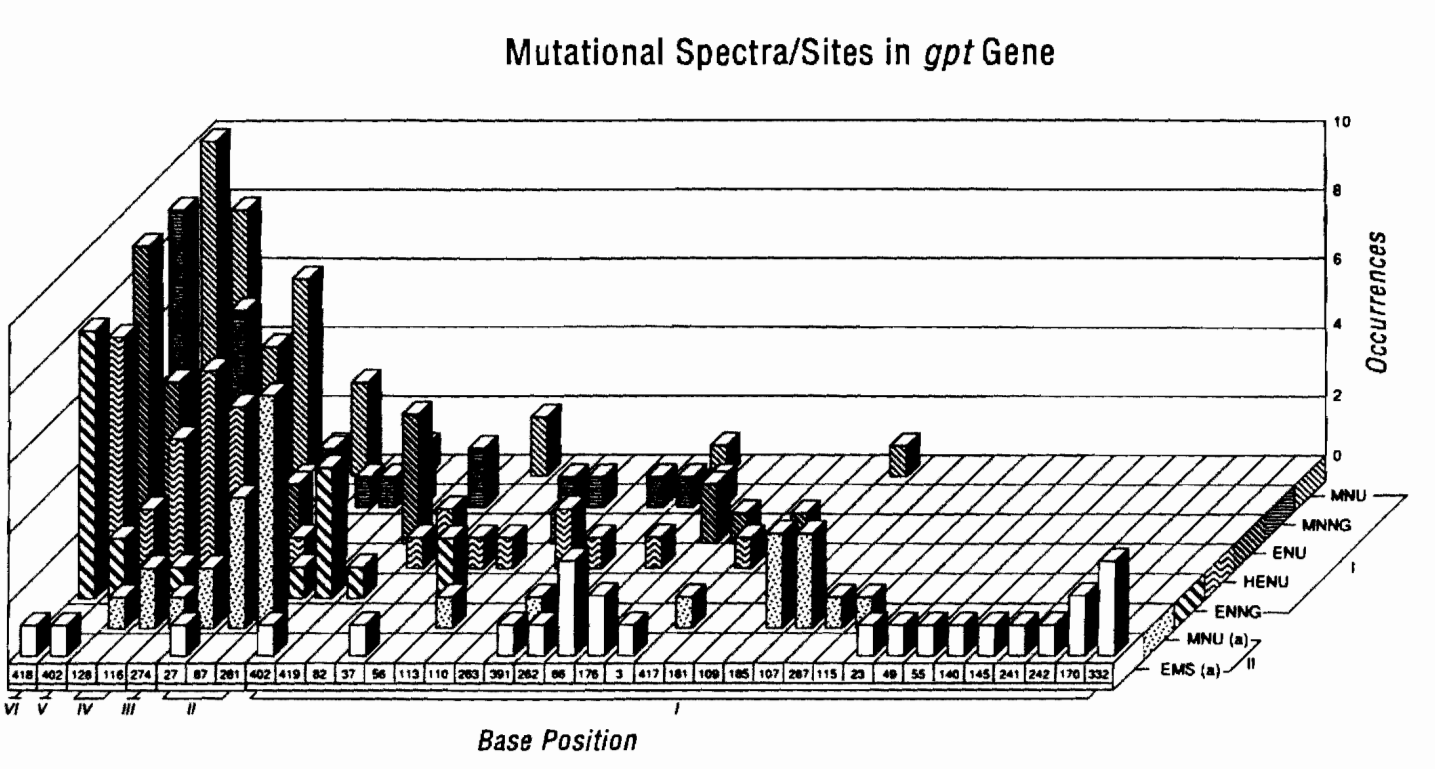

FIGURE 1.9 Frequency of mutated sites in the *gpt* gene of *E. coli* exposed to selected alkylating agents. EMS, ethyl methanesulfonate; MNU, *N*-methyl-*N*-nitrosourea; ENNG, *N*-ethyl-N′-nitro-*N*-nitrosoguanidine; HENU, 1-(2-hydroxyethyl)-1-nitrosourea; ENU, *N*-ethyl-*N*-nitrosourea, MNNG, *N*-methyl-N′-nitro-*N*-nitrosoguanidine. Roman numerals represent either base positions or chemical types that clustered in a multivariate analysis. This figure represents data collected by Benigni *et al.*, (1992). (Reprinted with permission.)

ton, 1943). Biologic monitoring of markers of exposure was advocated in the 1950s as an essential element of industrial hygiene programs (Elkins, 1954). The forerunner of the continuum of biologic markers shown in Figure 1.2 came from the work of Zielhuis (1978), Lauwerys (1983), Hernberg and Aitio (1987), and others. Four conditions for meaningful biologic monitoring have been identified from their work:

1. The substance and/or its metabolites are present in some tissue or body fluid suitable for sampling.
2. Valid practical and analytical methods are available.
3. The measurement is correct.
4. The result can be interpreted in terms of health risk or exposure (Tola and Hernberg, 1981).

That external exposure data is a poor approximation of individual dose is of particular concern in occupational epidemiology. If an individual dose is the integration of exposures from various routes and sources, both known and unknown, there may not be a high correlation of dose to a measure of external exposure. This problem was addressed by a World Health Organization study group (1975), which found that "an important conceptual advance in the assessment of toxic hazards is the increasing understanding that a good biological exposure test need not correlate with measurements of workroom air." Such correlations often are hampered by several factors, including exposures via nonrespiratory routes, additional nonoccupational exposures that accumulate in the person, use of respirators, and interindividual variations with respect to personal hygiene, smoking, metabolic processes, and physical work demands. This very discrepancy makes biologic monitoring a useful supplement to measurements of workroom air as an indicator of individual risk (Droz *et al.*, 1991).

A major historic parallel for molecular epidemiology is the testing of pulmonary function in worker populations exposed to airborne pollutants (Ingram and McFadden, 1981). This study exemplifies an early use of physiologic markers to establish dose–response relationships and predict risk. The challenge to assess the variability, sensitivity, accuracy, and precision of such tests serves as a lesson for those attempting to use molecular markers as surrogates for outcome and to assess individual risk.

Cardiovascular Epidemiology

Cardiovascular epidemiology is another historic precursor of molecular epidemiology. Biomarkers have played a significant role in the epidemiologic research of cardiovascular disease. The physiologic marker blood pressure has been assessed widely and, when elevated, has been found in epidemiologic studies to be a risk factor for cardiovascular disease. At the biochemical level, cholesterol and lipoproteins are included among the numerous risk fac-

tors for atherosclerosis. Studies of familial hyperlipidemias represent early examples of the use of biologic markers in assessment of gene–environment interactions. Research using these serum or plasma markers [e.g., the Framingham Heart Study (Kannel *et al.*, 1976), the Tecumseh Study (Epstein *et al.*, 1970), and other studies (Lipid Research Clinics Program (LRCP), 1984)] provided early lessons on the need for standardizing laboratory techniques for handling biomarkers. This need was most evident in the Lipid Research Clinics Program to determine the prevalence of various levels of lipids and lipoproteins in 11 North American populations. Triglycerides and cholesterol levels were determined for more than 75,000 individuals (LRCP, 1984). Serum cholesterol and other lipids have been used in epidemiologic studies as markers of exposure (to dietary fat), as markers of effect (risk factor for atherosclerosis), and as markers of susceptibility (LDL receptor defect) in genetic epidemiologic studies.

Cardiovascular disease research also provides a model for individual risk characterization, which has been described as a major advance of molecular epidemiology (Shields and Harris, 1991). Truett and colleagues (1967) described the calculations of individual risk functions based on multivariable analysis of covariates (including biologic markers and demographic and behavioral factors).

Genetics, Molecular Biology, and Genetic Epidemiology

The science of genetics was built on the concept that offspring phenotype is a biomarker of events at the gene level. This field provides good examples of how science uses observation and inference at the macro level. Visualization of cells—and eventually chromosomes—facilitated definition of mitosis and meiosis, so by the 1890s enough evidence was available to establish a chromosomal theory of inheritance (Mayr, 1982). In the twentieth century, progress in technique has allowed precise determination of chromosomal number; development of methods to detect how chromosomes function; recognition of Q-, G-, R-, T-, and C-bands of stained chromosomes (Caspersson and Zech, 1972); and understanding of chromatin structure (Jeppesen and Bower, 1989).

Genetics and molecular biology overlap as research extends to the level of the gene. Even before it was "seen," the gene was hypothesized, albeit among numerous competing hypotheses, as the unit of inheritance (see Mayr, 1982, for review), based on observations at the macro rather than the micro level. However, a more complete understanding evolved from research on the biochemistry of living processes.

Genetics and molecular biology rely in part on an understanding of the information macromolecules: proteins and nucleic acids. The central dogma of molecular biology is that information flows from DNA to RNA to proteins (Figure 1.10). This dogma also describes the flow of genetic information be-

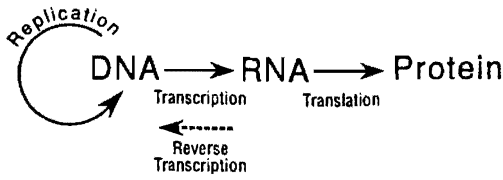

FIGURE 1.10 Information flow in molecular biology. Reprinted with permission from Maxon and Daugherty, 1985.

tween generations of somatic cells (mitosis) and germinal cells (meiosis). However, even this view is subject to change, as evidenced by the finding that RNA, once thought of as a passive carrier of genetic information, can sometimes function as an enzyme as well (Woldrop, 1992).

The sampling frame for molecular epidemiology is likely to be composed largely of the nucleotide sequences that constitute DNA (Figure 1.11). In 1883, Roux recognized that the basic process of transmission of genetic information was the division of the cell nucleus into two identical halves. Subsequently, however, the crucial event was recognized as the actual doubling of the genetic material, followed by its segregation into two daughter cells. In 1958, replication of DNA was demonstrated by Meselson and Stahl to be semi-conservative, that is, each daughter double helix receives one complete parental stand of DNA (Meselson and Stahl, 1958). This single strand serves as a template on which a new strand is synthesized from free nucleotides. The significance of this discovery is that the genetic information in a cell was shown to be reproducible indefinitely and—barring mutations—precisely.

Genetic analysis at the macro and micro levels has provided an opportunity to quantify variability in populations. "Variation" has been the topic that sparked creative thinking in biology as well as a controversial and perplexing one for molecular epidemiology (and promises to continue to be so). However, Mayr (1982) observed that Western thinking for 2000 years after Plato was dominated by "essentialism," or the belief that there were a limited number of immutable essences. Much of epidemiologic (and, for that matter, biomedical scientific) thinking to date has been essentialist. For example, the early statistics used in public health by Graunt (1620–1674) and Quetelet (1796–1874) attempted to calculate true values to overcome the confusing effects of variation. Quetelet, (see Mailley, 1875), who attempted to use a mathematical function of height and weight as a biomarker, hoped to calculate characteristics of the average person. To him and like thinkers, variations were nothing but "errors" around mean values. In the nineteenth century, new ways to view nature began to spread; a concept called "population thinking" developed that stresses the uniqueness of everything in the organic world. To population thinkers, there is no "typical" individual and mean values are abstractions (Mayr, 1982).

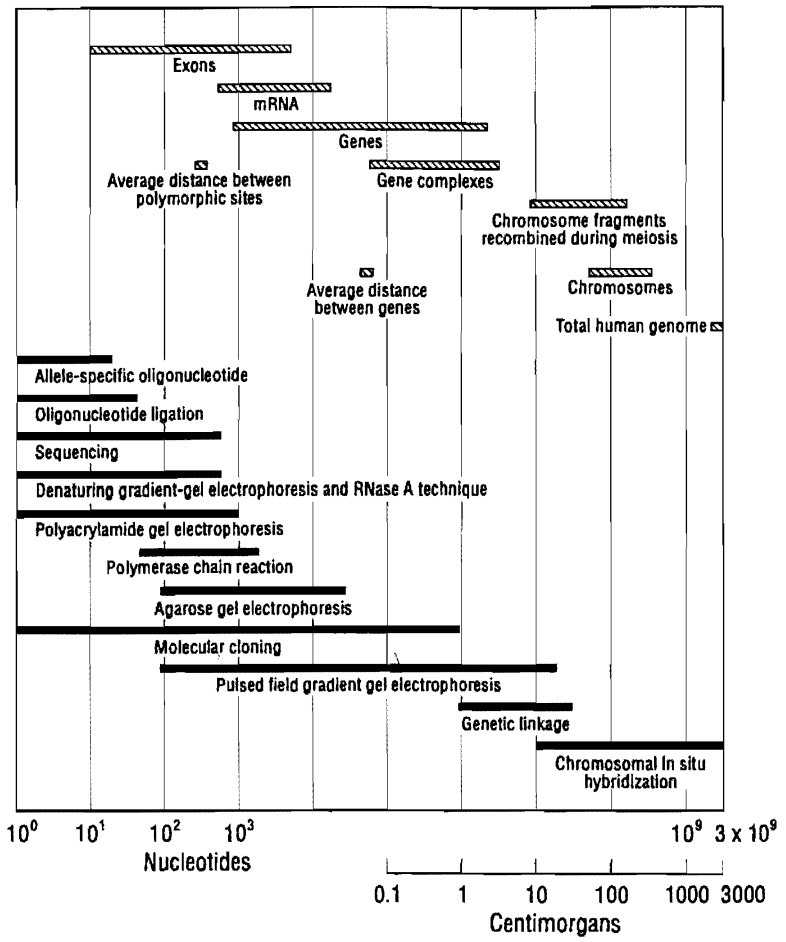

FIGURE 1.11 The size of the human genome in relationship to the various genetic and physical methods that are being applied to obtaining a map. Hatched lines, Regions of informational units. Solid lines, Techniques used in mapping. (Modified from Landegren *et al.*, 1988 in Weatherall, 1991.)

With the development of electrophoretic and other analytical techniques, researchers have discovered far more variability in proteins and their underlying genetic information than population geneticists and evolutionary biologists expected (Janetos, 1988). Similarly, with the introduction of recombinant DNA technology, the genomes of prokaryotes and eukaryotes have been found to be far more diverse and plastic than originally thought (Janetos, 1988).

Another lesson from genetics has been the perspective on the relative roles of heredity and environment in causing or exacerbating disease. Despite

a long-standing "nature–nurture" debate, most scientists support the concept that phenotypic variation generally results from both environmental and hereditary influences. "Genetic epidemiology" has been particularly useful in disentangling the effects of heredity and environment. Biologic markers have been used in genetic epidemiology at the macro level (an individual trait, such as eye color) and at the micro level (e.g., HLA antigens) in twin studies, adoption studies, path analysis, analysis of cultural transmission of disease risk factors, and studies of association between specific genotypes and diseases. Since the early 1950s, the terms "pharmacogenetics" (Vogel, 1959), "ecogenetics" (Brewer, 1971), and "occupational ecogenetics" (Mulvihill, 1986) have been used to describe fields studying the interaction of genetic factors and exogenous agents. By 1975, a report by the National Academy of Sciences (1975) listed 92 human genetic disorders believed to predispose the affected individuals to the toxic effects of pollutants. DNA polymorphisms have been used to assess genetic predisposition to various diseases, such as cardiovascular disease, cancer, and diabetes, that may put persons with exposure to xenobiotics at increased risk (Cartwright *et al.*, 1982; Rajput-Williams *et al.*, 1988; Trucco and Dorman, 1989; Caporaso *et al.*, 1990; Weatherall, 1991).

Since the time of Mendel, genetics and genetic epidemiology have used the concept of "genetic markers." A genetic marker is a variant allele that is used to label a biologic process or structure throughout the course of an experiment or epidemiologic study (Suzuki *et al.*, 1986). That Mendel used the genes for pea shape, pea color, and so on was largely irrelevant; he used them as markers to trace the hereditary processes of segregation and assortment. DNA technology has provided a method for obtaining a large number of new genetic markers, which facilitates studying heritable differences in DNA sequences. Unlike classical expressed markers, DNA polymorphisms can be detected whether or not a given sequence encodes a protein (Conneally *et al.*, 1984).

Ultimately, for epidemiologic purposes, the technical developments for amplifying small samples of DNA (the polymerase chain reaction; Saiki 1985) and for determining the sequences of genes will be considered landmark events. These approaches allow the use of small amounts of sample collected from contemporary subjects or from stored or archived specimens. When applied to a properly designed study, the polymerase chain reaction (PCR) techniques can allow powerful comparisons among various groups. Moreover, a high resolution map of the entire human genome could be developed using restriction fragment length polymorphisms (described in Botsein *et al.*, 1980).

Molecular biology, which has undergone a meteoric rise since the discovery of the structure of DNA (Watson and Crick, 1953), has been the source of powerful tools in genetics that will be employed in molecular epidemiology. The discoveries of molecular biology have simplified and unified biol-

ogy, even as they have shown the immense diversity. These discoveries are a driving force behind molecular epidemiology, since they promise to permit more accurate classification of subjects, more detailed appraisal of mechanisms, and earlier opportunities for prevention and control.

Epidemiologic and Statistical Context

A number of authors have discussed how "molecular epidemiology" fits into the historic framework of epidemiology (Perera and Weinstein, 1982; Hatch and Stein, 1987; Schulte, 1987; 1989; Vandenbroucke, 1988; Hulka *et al.*, 1990; Loomis and Wing, 1990; Yach, 1990). Vandenbroucke (1988) believes the use of molecular biology to identify individual steps in disease causation is reminiscent of the nineteenth century conflict between the miasma and contagion theories of disease. For the miasmists, the competing idea was the germ theory; for the epidemiologists of today, it is molecular biology.

Loomis and Wing (1990) see the contest between the miasma and contagion theories as growing from a tension as old as Western medicine between the often unglamorous work of the hygienist and the heroic capabilities of the healer. The miasma and contagion theories share a derivation from a global scientific paradigm that has been called "Cartesian reductionism." As Loomis and Wing caution:

> In epidemiology this paradigm equates causal inquiry with discovering the inherent dose–response relationships between agents and diseases through rigorous hypothesis testing. The weakness of causal theories arising from the Cartesian paradigm is their failure to recognize that the action of interdependent parts (agents, aspects of the environment, and individuals) is not an immutable, historical characteristic of those parts but is, instead, dependent on the properties of the whole system in which they operate. (Loomis and Wing, 1990)

Data on biologic markers often have been used as topical examples in the evolution of statistical epidemiologic ideas. Generally, use of the markers was not the driving force behind an idea, but a case in point. For example, Woolf's paper (1955), "On estimating the relation between blood group and disease," presented the first opportunity to assess the common odds ratio estimation for a set of two-by-two tables and a test for heterogeneity of the odds ratio across tables (Greenland, 1987).

Measurements of lipids, lipoproteins, and blood pressure were among the variables used in heart disease research that forced researchers to introduce concepts of multivariate analysis to overcome the "thinness of data" encountered in ordinary stratified analysis (Greenland, 1987). Similarly, when Gordon (1974) addressed the hazards of using multivariate analysis, such as the assumption of a linear combination of variables, biomarkers were used as examples. They also were used to illustrate the issue of combining categorical and continuous variables.

Many of the issues about validity and reproducibility of biological mark-

ers came from work on nutritional epidemiology (Keys *et al.*, 1965; Willet *et al.*, 1983; Romieu *et al.*, 1990; Bates, *et al.*, 1991). The primary efforts have involved validation of biochemical markers as predictors of dietary intake and the use of biologic markers to validate food frequency questionnaires.

A significant contribution to the statistical handling of biologic marker data was the seminal work of Finney (1952), *Statistical Method in Biological Assay*. Finney traced the history of the biologic assay dating from the late nineteenth century. His focus encompassed Erlich's investigations into the standardization of diphtheria antitoxin; developments in pharmacology, endocrinology, and plant pathology; and the works by Coward (1947) and Gaddum (1948) that directed attention to the statistical considerations. Key in Finney's work is the elucidation of concepts of the statistical validity of a bioassay and his recommendation pertinent to molecular epidemiology:

> Unless the assayist himself knows enough about the statistical logic and structure of programs, he may feel compelled to use methods suited to his experiments or may overlook indications that anomalies or unsuspected errors make a particular set of data misleading. The great expansion in some types of assay (notably radioimmunoassays) that are essential to patient care surely renders neglect of quality in statistical analysis as blameworthy as neglect of proper maintenance of equipment or negligence in identifying and measuring samples. (Finney, 1952)

Much of what is portrayed currently as molecular epidemiology appears not to have heeded this lesson. Such studies often fail to involve discussion of study design, statistical analysis, sources of bias, misclassification, or errors. Considering changes found in a gene to be indicative of the "truth" about a phase in the natural history of disease may be acceptable as part of a pilot process in molecular epidemiology, but will not support a firm foundation for inference, comparison, or generalization without attention to statistical considerations.

Many of the issues that will face practitioners of molecular epidemiology will be similar to those identified by John Paul (1958) in his publication, *Clinical Epidemiology* and include issues pertaining to host susceptibility, understanding the natural history of disease, interacting with the subject, and collecting biologic specimens. Clinical epidemiology is a forerunner of molecular epidemiology in these ways.

The recent history of molecular epidemiology was built on work done in the 1980s and early 1990s in molecular dosimetry (reviewed in Gledhill and Mauro, 1990; Garner *et al.*, 1991; Groopman and Skipper, 1991), and in pilot efforts to incorporate biologic markers into epidemiologic research. These efforts were mostly in the area of environmental health research. Papers and monographs (Perera and Weinstein, 1982; Tannenbaum and Skipper, 1984; Harris *et al.*, 1985; Wogan and Gorelick, 1985; Schulte, 1987; Hulka and Wilcosky, 1988) in the 1980s formed the systematic body of literature on methods and techniques of this nascent field. The publication by Hulka *et al.* (1990), *Biological Markers in Epidemiology,* was the first

epidemiology-oriented text in the field. This work, although mostly focused on cancer-related biomarkers, described many of the basic issues that, in retrospect, will be deemed characteristic of this approach, including when to use a biologic marker in epidemiologic research, properties of the marker, sample size, and control of confounding, analysis, and interpretation of marker. Other related publications (Bertazzi and Duca, 1987; Hatch and Stein, 1987; Perera *et al.*, 1987, 1990; Schulte, 1987, 1989; Vineis *et al.*, 1990; NRC, 1991a, b) also provide guidance on methodologic issues.

Conclusion

Currently, molecular epidemiology is an evolving set of techniques rather than a well-conceived and practiced discipline. The "field" may become more rigorous by application of lessons learned from other disciplines that have a history of obtaining and using biologic markers. Biomedical science can be characterized by a process of increasingly refined biologic measurements and their use in etiologic, preventive, and therapeutic research. I have described this history briefly, with particular reference to predecessor disciplines that contribute to a body of knowledge that could be useful to molecular epidemiology. This is a rich history that reveals how many of the concerns about molecular epidemiology have been confronted previously. Researchers can learn from this history and use it as a foundation to establish good principles in molecular epidemiology.

References

Abdel-Baky, S., and Giese, R. W. (1991). Gas chromatography electron capture negative-ion mass spectrometry at the zeptomole level. *Anal. Chem.* **63**, 2986–2989.

Ahearn, J. M., and Hochberg, M. C. (1988). Epidemiology and genetics of ankylosing spondylitis. *J. Rheumatol.* **15** (**Suppl. 6**), 22–28.

Alavanja, M., Aron, J., Brown, C., and Chandler, J. (1987). Cancer risk-assessment models: Anticipated contributions from biochemical epidemiology. *J. Natl. Cancer Inst.* **4**, 633–643.

Albertini, R. J., Castle, K. L., Borcherding, W. R. (1982). T-cell cloning to detect the mutant 6-thioquanine-resistant lymphocytes present in peripheral blood. *Proc. Nat. Acad. Sci. U.S.A.* **79**, 6617–6621.

Anderson, N. L., Hoffman, J. P., Gemmell, A., and Taylor, J. (1984). Global approaches to quantitative analysis of gene-expression patterns observed by use of two-dimensional gel electrophoresis. *Clin. Chem.* **30**, 2031–2036.

Aycock, W. L., and Kramer, S. D. (1930). Immunity to poliomyelitis in normal individuals in urban and rural communities as indicated by the neutralization test. *J. Prev. Med.* **4**, 189–200.

Ayesh, R., Idle, J. R., Richie, J. C., Crothers, M. J., and Hetzel, M. R. (1984). Metabolic oxidation phenotypes as markers for susceptibility to lung cancer. *Nature* **312**, 169–170.

Baak, J. P. A., and Tosi, P. (1991). Historical background of quantitative pathology. *In* "Manual

of Quantitative Pathology in Cancer Diagnosis and Prognosis" (J. P. A. Baak, ed.), pp. 3–5. Springer-Verlag, Berlin.

Bagshawe, K. D. (1983). Tumour markers—Where do we go from here? Br. J. Cancer 48, 167–175.

Balayeva, N. (1989). Approaches to the molecular epidemiology of rickettsioses. *Eur. J. Epidemiol.* 5, 414–419.

Basler, A. (1987). Scientific justification of testing chromosome mutations and regulatory requirements for the assessment of mutagenicity. *In* "Cytogenetics" (G. Obe and A. Basler, eds.), pp. 379–393. Springer-Verlag, Berlin.

Bates, C. J., Thurnham, D. I., Bingham, S. A., Margetts, B. M., and Nelson, M. (1991). Biochemical markers of nutrient intake. *In* "Design Concepts in Nutritional Epidemiology" (B. M. Margetts and M. Nelson eds.), pp. 192–265. Oxford University Press, New York.

Beebe, G. W., Kato, H., and Land, C. E. (1978). Studies of the mortality of A-bomb survivors. 6. Mortality and radiation dose 1950–74. *Radiat. Res.* 75, 138–201.

Begley, S., Glick, D., and Wilson, L. (1991). Not in the stars but in our genes. *Newsweek* October 21, 56–57.

Benigni, R., Palombo, F., and Dogliotti, E. (1992). Multivariate statistical analysis of mutational spectra of alkylating agents. *Mutation Research* 267, 77–88.

Benzer, S. (1961). On the topography of the genetic fine structure. *Proc. Natl. Acad. Sci. U.S.A.* 47, 403–415.

Berenblum, I., and Shubik, P. (1949). An experimental study of the initiating state of carcinogenesis and re-evaluation of the somatic cell mutation theory of cancer. *Brit. J. Cancer* 3, 109–118.

Bertazzi, P. A., and Duca, P. S. (1987). The use of biological indicators in the surveillance of groups of individuals. *In* "Occupational and Environmental Chemical Hazards: Cellular and Biochemical Indices for Monitoring Toxicity" (V. Foa, E. A. Emmett, M. Maroni, and A. Colombi, eds.), pp. 61–68. Ellis Horwood, Chichester, England.

Bostofte, E., Bagger, P., Michael, A., and Stakemann, G. (1988). Fertility prognosis for infertile men: Results of follow-up study of semen analysis from two different populations evaluated by the Cox regression model. *Urology–Andrology* 54, 1100–1106.

Botstein, D., White, R. L., Skolnick, M., and Davis, R. W. (1980). Construction of genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32, 314–331.

Bressac, B., New, M., Wunds, J., and Ozturk, M. (1991). Selective G to T mutations of p53 gene in hepatocellular carcinoma from Southern Africa. *Nature* (London) 350, 429–431.

Brewer, J. C. (1971). Annotations: Human ecology, and expanding role for the human geneticist. *Am. J. Hum. Genet.* 23, 92–94.

Bull, R. J. (1989). "Decision Model for the Development of Biomarkers of Exposure." U.S. Environmental Protection Agency. EPA 600/x-89/163. U.S. Government Printing Office, Washington, D.C.

Calleman, C. J., Ehrenberg, L., Jansson, B. Ostermann-Golkar, S., Segerback, D., Svansson, K., and Wachtmeister, C. A. (1978). Monitoring and risk assessment by means of alkyl groups in hemoglobin in persons occupationally exposed to ethylene oxide. *J. Environ. Pathol. Toxicol.* 2, 427–442.

Caporaso, N. E., Tucker, M. A., Hoover, R. N., Hayes, R. B., Pickle, N. W., Issaq, H. J., Muschik, G. M., Green-Gallo, L., Burvys, P., Arsnor, S., Resau, J. H., Trump, B. F., Tollerud, D., Weston, A., and Harris, C. C. (1990). Lung cancer and the debrisoquine metabolic phenotype. *J. Natl. Cancer Inst.* 82, 1264–1272.

Cartwright, R. A., Glashen, R. W., Rodgers, H. J., Ahmad, R. A., Barham-Hall, T., and Higgins, E. (1982). The role of *N*-acetyltransferase phenotypes in bladder carcinogensis. *Lancet* ii, 842–896.

Caspersson, T., and Zech L. (eds.) (1972). "Chromosome Identification." Nobel Foundation, Stockholm.

Chard, T. (1990). "An Introduction to Radioimmunoassay and Related Techniques." Elsevier Science, Amsterdam.

Clendening, L. (1942). "Source Book of Medical History," pp. 106–121. Dovery Publications, New York.

Conneally P. M., Wallace, M. R., Gusella, J. F., and Wexler, N. S. (1984). Huntington Disease: estimation of heterozygote status using linked genetic markers. *Genet. Epid.* 1, 81–88.

Coward, K. H. (1947). "The Biological Standardisation of the Vitamins," (2d Ed.) Baillieve, Tindal, & Cox, London.

Cullen, M. R. (1989). The role of clinical investigations in biological markers research. *Environ. Res.* 50, 1–10.

Detels, R., Visscher, B. R., Fahey, J. L., Sever, J. L., Gravell, M., Madden, D. L., Schwartz, K., Dudley, J. P., English, D. A., Powers, H., Clark, V. A., and Gottlieb, M. S. (1987). Predictors of clinical AIDS in young homosexual men in a high-risk population. *Int. J. Epidemiol.* 16, 271–276.

Droz, P. O., Berode, M., and Wu, M. M. (1991). Evaluation of concomitant biological and air monitoring results. *Appl. Occ. Env. Hyg.* 6, 465–474.

Ehrenberg, L. (1974). Genetic toxicity of environmental chemicals. *Acta Biol. Ingosl. Genetika* 6, 367–398.

Ehrenberg, L., Hallstrom, T., Ostermann-Golkar, S., and Wennberg, I. (1974). Evaluation of genetic risks of alkylating agents: Tissue doses in the mouse from air contaminated with ethylene oxide. *Mutat. Res.* 24, 83–103.

Elkins, H. B. (1954). Analyses of biological materials as indices of exposure to organic solvents. *Arch. Ind. Hyg. Occ. Med.* 9, 212–220.

Epstein, F. H., Napier, J. A., Block, W. D., Hayner, N. S., Higgins, M. P., Johnson, B. C., Keller, J. B., Mitzner, H. L., Montoyo, H. J., Ostrandor, L. D., and Ullman, B. M. (1970). The Tecumseh study: Design, progress, and prospectives. *Arch. Environ. Health* 21, 402–407.

Everson, R. B. (1989). Conceptual models for biochemical epidemiology. *Proc. Am. Assoc. Cancer Res. Ann. Mg.* 30, 678–679.

Feinstein, A. R. (1991). Scientific paradigms and ethical problems in epidemiologic research. *J. Clin. Epidemiol.* 44 (**Suppl.** 1), 1195–1235.

Feulgen, R., Rossenbeck, H. (1924). Mikrosopisch—chemischer Nachweis einer Nukleinsäure vom Typus der Thymonukleinsäure und die daraut beruhende elektive Färbung von Zellkernen in mikroscopische Präparaten. *Z. Physiol. Chem.* 135, 203–224.

Finney, D. J. (1952). "Statistical Method in Biological Assay." Charles Griffin, London.

Follett, E. A. C., Sanders, R. C., Beards, G. M., Hundley, F., and Desselberger, U. (1984). Molecular epidemiology of human rotaviruses. *J. Hyg. Camb.* 92, 209–222.

Foulds, L. (1958). The natural history of cancer. *J. Chron. Dis.* 8, 2–37.

Frost, W. U. (1928). Infection, immunity, and disease in the epidemiology of diphtheria. *J. Prev. Med.* 2, 305.

Gaddum, J. H. (1948). "Pharmacology." Oxford University Press, London.

Galen, R. S., and Gambino, S. R. (1975). "Beyond Normality: The Predictive Value and Efficiency of Medical Diagnoses." John Wiley and Sons, New York.

Gann, P. H., Davis, D. L., and Perera, F. (1985). "Biologic markers in environmental epidemiology." Paper presented at the 5th workshop of the Scientific Group on Methodologies for the Safety Evaluation of Chemicals (SGOMSEC), sponsored by the World Health Organization, 12–16 August, Mexico City.

Garner, R. C., Farmer, P. B., Steel, G. T., and Wright, A. S. (eds.) (1991). "Human Carcinogen Exposure. Biomonitoring and Risk Assessment." IRL Press, Oxford.

Ghosh, B. C., and Rob, C. G. (1987). Tumor markers. *In* "Tumor Markers and Associated Antigens" (B. C. Ghosh and L. Ghosh, eds.), pp. 1–10. McGraw-Hill, New York.

Gledhill, B. L., and Mauro, F. (eds.) (1990). "New Horizons in Biological Dosimetry." Wiley-Liss, New York.

Gordon, T. (1974). Hazards in the use of logistic function with reference to data from prospective cardiovascular studies. *J. Chron. Dis.* 27, 97–102.

Graunt, J. (1662). "Natural and Political Observations Made upon the Bills of Mortality." Republished by The Johns Hopkins University Press, Baltimore, 1939.

Greenland, S. (ed.) (1987). "Evolution of Epidemiologic Ideas." Epidemiology Resources, Chestnut Hill.

Greenwald, P., Cullen, J. W., and Weed, D. (1990). Introduction: Cancer prevention and controls. *Sem. Oncol.* 17, 383–390.

Groopman, J. D., and Skipper, P. L. (1991). "Molecular Dosimetry and Human Cancer: Analytical, Epidemiological, and Social Considerations." CRC Press, Boca Raton, Florida.

Groopman, J. D., Sabbioni, G., and Wild, C. P. (1991). Molecular dosimetry of human aflatoxin exposures. *In* "Molecular Dosimetry and Human Cancer: Analytical, Epidemiological, and Social Considerations" (J. D. Groopman and P. L. Skipper, eds.), pp. 303–324. CRC Press, Boca Raton, Florida.

Hamilton, A. (1943). "Exploring the Dangerous Trades." Little Brown, Boston.

Harris, A. L. (1991) Telling changes of base. *Nature (London)* **350**, 377–378.

Harris, C. C. (1991). Molecular epidemiology: Overview of biochemical and molecular basis. *In* "Molecular Dosimetry and Human Cancer: Analytical, Epidemiological and Social Considerations." (J. D. Groopman and P. L. Skipper, eds.), pp. 15–26. CRC Press, Boca Raton, Florida.

Harris, C. C., Vahakangus, K., Autrup, H., Trivers, G. E., Shamsuddin, A. K. M., Trump, B. F., Bowman, B. M., and Mann, D. L. (1985). Biochemical and molecular epidemiology of human cancer risk. *In* "The Pathologist and the Environment." (D. Scarpelli and J. Craighead, eds.) International Academy of Pathology Monograph No. 26, pp. 140–167.

Harris, C. C., Weston, A., Willey, J. C., Trivers, G. E., and Mann, D. L. (1987). Biochemical and molecular epidemiology of human cancer: Indicators of carcinogen exposure, DNA damage, and genetic predisposition. *Environ. Health Perspect.* 75, 109–119.

Hatch, M. C., and Friedman-Jimenez, G. (1991). Using reproductive effect marker to observe subclinical events, reduce misclassification, and explore mechanism. *Env. Health Perspect.* **90**, 255–259.

Hatch, M. C., and Stein, Z. A. (1987). The role of epidemiology in assessing chemical induced disease. *In* "Mechanisms of Cell Injury: Implications for Human Health" (B. A. Fowler, ed.), pp. 303–314. John Wiley and Sons, New York.

Hattis, D. (1988). The use of biological markers in risk assessment. *Stat. Sci.* 3, 358–366.

Hemstreet, G. P., Schulte, P. A., Ringen, K., Stringer, W., and Altekruse, E. B. (1988). DNA hyperploidy as a marker for biological response to bladder carcinogen exposure. *Int. J. Cancer* **42**, 817–820.

Henderson, R. F., Bechtold, W. E., Bond, J. A., and Sun, J. D. (1989). The use of biological markers in toxicology. *Crit. Rev. Tox.* 20, 65–82.

Hernberg, S. (1987). Report of Session 1. *In* "Occupational and Environmental Chemical Hazards: Cellular and Biochemical Indices for Monitoring Toxicity" (V. Foa, E. A. Emmett, M. Maroni, and A. Colombi, eds.), pp. 121–124. Ellis Horwood, Chichester, England.

Hernberg, S. and Aitio, A. (1987). Validation of biological monitoring tests. *In* "Occupational and Environmental Chemical Hazards: Cellular and Biochemical Indices for Monitoring Toxicity" (V. Foa, E. A. Emmett, M. Maroni, and A. Colombi, eds.), pp. 41–49. Ellis Horwood, Chichester, England.

Higginson, J. (1977a). The role of the pathologist in environmental medicine and public health. *Am. J. Pathol.* **86**, 460–484.

Higginson, J. (1977b). Changing concepts in cancer prevention: Limitations and implications for future research in environmental carcinogenesis. *Cancer Res.* **48**, 1381–1389.

Hoffman, D., Hayley, N. J., Djordjevic, K. D., Brunnemann, K. D., and Hecht, S. S. (1991). Biomarkers to exposure to tobacco products. In "Human Carcinogen Exposure" (R. C.

Garner, P. B. Farmer, G. T. Steel, and A. S. Wright, eds.), pp. 275–284. Oxford: New York.

Hogue, J. R., and Brewster, M. A. (1988). Developmental risks: Epidemiologic advances in health assessment. *In* "Epidemiology and Health Risk Assessment" (L. Gordis, ed.), pp. 61–81. Oxford University Press, New York.

Hollstein, M., Sidransky, D., Vogelstein, B., and Harris, C. C. (1991). p53 mutations in human cancers. *Science* 253, 49–53.

Hook, E. (1982). Perspectives in mutation epidemiology: Epidemiologic and design aspects of studies of somatic chromosome breakage and sister chromatid exchange. *Mutat. Res.* 99, 373–383.

Hsu, I. C., Metcalf, R. A., Sun, T., Welsh, J. A., Wang, N. J., and Harris, C. C. (1991). Mutational hotspot in the p53 hepatocellular carcinomas. *Nature* (London) 350, 427–428.

Hulka, B. S., and Wilcosky, T. (1988). Biological markers in epidemiologic research. *Arch. Environ. Health* 43, 83–89.

Hulka, B. S., Wilcosky, T. C., and Griffith, J. D. (1990). "Biological Markers in Epidemiology." Oxford University Press, New York.

Ingram, R. H., Jr., and McFadden, E. R., Jr. (1981). Physiologic measurements providing enhanced sensitivity in detecting early effects of inhalants. *In* "Occupational Lung Diseases: Research Approaches and Methods" (H. Weill and M. Turner-Warwick, eds.), pp. 87–98. Marcel Dekker, New York.

Janetos, A. C. (1988). Biological variability. *In* "Variation in Susceptibility to Inhaled Pollutants" (J. D. Brain, B. D. Beck, A. J. Warren, and R. A. Shaikh, eds.), pp. 9–29. The Johns Hopkins University Press, Baltimore.

Jeppesen, P., and Bower, D. J. (1989). Towards understanding the structure of *eukaryotic* chromosomes. *In* "Cytogenetics" (G. Obe and A. Basler, eds.), pp. 1–29. Springer-Verlag, Berlin.

Kannel, W. B., McGee, D., and Gordon, T. (1976). A general cardiovascular risk profile: The Framingham Study. *Am. J. Cardiol.* 38, 46–51.

Keys, A., Anderson, J. T., and Grande, F. (1965). Serum cholesterol response to changes in the diet. IV. Particular saturated fatty acids in the diet. *Metabolism* 14, 376–387.

Kilbourne, E. D. (1973). The molecular epidemiology of influenza. *J. Infect. Dis.* 127, 478–487.

Kilbourne, E. (1979). "Molecular Epidemiology: Influenza as Archetype." *The Harvey Lectures* 73, 225–228.

Klee, G. G., and Go, V. L. W. (1987). Carcinoembryonic antigen and its role in clinical practice. *In* "Tumor Markers and Tumor Associated Antigens" (B. C. Ghosh and L. Ghosh, eds.), pp. 22–43. McGraw-Hill, New York.

Koss, L. G. (1987). Automated cytology and histology. A historical perspective. *Anal. Quant. Cytol. Histol.* 9, 369–374.

Koss, L. G. (1990). The future of cytology. *Acta Cytologica* 34, 1–9.

Kuller, L. (1991). Epidemiology is the study of "epidemics" and their prevention. *Am. J. Epidemiol.* 134, 1051–1056.

Landrigan, P. J. (1988). Relation of body burden measures to ambient measures. *In* "Epidemiology and Health Risk Assessment" (L. Gordis, ed.), pp. 139–147. Oxford University Press, New York.

Last, J. M. (ed.) (1988). "A Dictionary of Epidemiology." 2nd Edition, Oxford University Press, New York.

Lauwerys, R. R. (1983). "Industrial Chemical Exposure: Guidelines for Biological Monitoring." Biomedical Publications, Davis, California.

Lipid Research Clinics Program (1974). "Manual of Laboratory Operations. Vol. 1. Lipid and Lipoprotein Analysis," 2d Ed. U.S. Government Printing Office, Washington, D.C.

Lipid Research Clinics Program (1984). The Lipid Research Clinics Coronary Primary Prevention Trial Results. I. Reduction in incidence of coronary heart disease. *J. Am. Med. Assoc.* 251, 351–364.

Loomis, D., and Wing, S. (1990). Is molecular epidemiology a germ theory for the end of the twentieth century? *Int. J. Epidemiol.* 19, 1–3.

Lower, G. M. (1982). Concepts in causality: chemically induced human urinary bladder cancer. *Cancer* 49, 1056–1066.

Lower, G. M., Nilsson, T., Nelson, C. E., Wolf, H., Gamsley, T. E., and Bryan, G. T. (1979). N-acetyltransferase and risk in urinary bladder cancer; approaches in molecular epidemiology. Preliminary results in Sweden. *Environ. Health Perspect.* 29, 71–79.

Lucier, G. W., and Thompson, C. L. (1987). Issues in biochemical applications to risk assessment: When can lymphocytes be used as surogate markers. *Environ. Health Perspect.* 76, 187–191.

Maclure, K. M., and MacMahon, B. (1980). An epidemiologic perspective of environmental carcinogenesis. *Epidemiol. Rev.* 2, 19–48.

Mailley, E. (1875). Essai sur la vie et les ouverage de L.A.J. Quetelet. F. Hayes, Brussells.

McMahon, G., Davis, E., Huber, L. J., Kim, Y., and Nognn, G. (1990). Characterization of C-Ki-*ras* and N-*ras* oncogenes in aflatoxin B1-induced rat liver tumors. *Proc. Natl. Acad. Sci. U.S.A.* 87, 1104–1108.

Maxon, L. R., and Daugherty C. H. (1985). "Genetics." William C. Brown, Dubuque, Iowa.

Mayr, E. (1982). "The Growth of Biological Thought." Harvard University Press, Cambridge.

Mendelsohn, M. L. (1991). An introduction to biological dosimetry. *In* "New Horizons in Biological Dosimetry" (B. L. Gledhill and F. Mauro, eds.), pp. 1–9. Wiley-Liss, New York.

Meselson, M., and Stahl, F. (1958). The replication of DNA in *Escherichia coli. Proc. Nat. Acad. Sci., U.S.A.* 44, 671–682.

Miller, J. R. (1983). Perspectives in mutation epidemiology. 4. General principles and considerations. *Mutat. Res.* 114, 425–447.

Muench, K. F., Misra, R. P., and Humayun, M. Z. (1983). Sequence specificity in aflatoxin B1-DNA interactions. *Proc. Natl. Acad. Sci. USA* 80, 6–10.

Mulvihill, J. J. (1986). Occupational ecogenetics: Gene–environmental interactions in the workplace. *J. Occup. Med.* 28, 1093–1095.

Murai, A., Miyahara, T., Fujimoto, N., Matsuda, M., and Kameyama, M. (1986). Lp[a] lipoprotein as a risk factor for coronary heart disease and cerebral infarction. *Atherosclerosis* 59, 199–204.

National Academy of Sciences. (1975). Special risks due to inborn errors of metabolism. In "Principles for Evaluating Chemicals in the Environment." p. 331. National Academy Press, Washington, D.C.

National Academy of Sciences (1991a). "Human Exposure Assessment for Airborne Pollutants. Advances and Opportunities." National Academy Press, Washington, D.C.

National Academy of Sciences (1991b). "Environmental Epidemiology." National Academy Press, Washington, D.C.

National Research Council (1987). Biological markers in environmental health research. *Environ. Health Perspect.* 74, 1–191.

Oehlkers, F. (1943). Die Auslösung von Chromosomenmutationen in der Meiose durch eine Wirkung von Chemikalien. *Z. Ind. Abst. Verbundsl.* 81, 313–341.

Omenn, G. S. (1982). Predictive identification of hypersusceptible individuals. *J. Occup. Med.* 24, 369–374.

Ostermann-Golkar, S. (1983). Tissue doses in man: Implications in risk assessment. *In* "Developments in the Science and Practice of Toxicology" (A. W. Hayes, R. C. Schnell, and T. S. Miya, eds.), pp. 289–298. Elsevier Science, Amsterdam.

Papanicolaou, G. N., and Traut, H. F. (1943). Diagnosis of uterine cancer by the vaginal smear. Commonwealth Fund, New York.

Pappenheimer, A. M., Jr., and Murphy, J. R. (1983). Studies on the molecular epidemiology of diphtheria. *Lancet* October 22, 923–926.

Paul, J. R. (1958). "Clinical Epidemiology." University of Chicago Press, Chicago.

Paul, J. R., and White, C. (1973). "Serological Epidemiology." Academic Press, New York.

Pereira, L., Cassai, E., Honess, R. W., Roizman, B., Terni, M., and Nahmias, A. (1976). Variability in the structural polypeptides of herpes simplex virus 1 strains: Potential application in molecular epidemiology. *Infect. Immun.* 13, 211–220.

Perera, F. P. (1987a). The potential usefulness of biological markers in risk assessment. *Environ. Health Perspect.* 76, 141–145.

Perera, F. P. (1987b) Molecular cancer epidemiology: A new tool in cancer prevention. *J. Natl. Cancer Inst.* 78, 887–898.

Perera, F. P., and Weinstein, I. B. (1982). Molecular epidemiology and carcinogen–DNA adduct detection: New approaches to studies of human cancer causation. *J. Chron. Dis.* 35, 581–600.

Perera, F. P., Schulte, P. A., Santella, R. M., and Brenner, D. (1990). DNA adducts and related markers in assessing the risk of complex mixtures. *In* "Genetic Toxicology of Complex Mixtures" (M. D. Waters, F. B. Daniel, J. Lewtas, M. M. Moore, and S. Nesnow, eds.), pp. 271–290. Plenum Press, New York.

Rajput-Williams, T. J., Knott, Wallis, S. C., Sweetman, P., Yarnell, J., Cox, N., Bell, G. I., Miller, N. E., and Scott, J. (1988). Variation of apoliprotein-B gene is associated with obesity, high blood cholesterol levels, and increased risk of coronary heart disease. *Lancet* ii, 1442–1446.

Rao, J. Y., Hemstreet, G. P., Hurst, R. E., Bonner, R. A., Jones, P. L., Kyung, W. M., and Fradet, Y. (1991). "Biochemical Mapping of Bladder Cancer Tumorigenesis" (submitted).

Richterich, R. (1969). "Clinical Chemistry." pp. 3–5. S. Karger, Basel.

Rogan, W. J. (1988). Relation of surrogate measures to measures of exposure. *In* "Epidemiology and Health Risk Assessment" (L. Gordis, ed.), pp. 148–158. Oxford University Press, New York.

Romieu, I., Stampfer, M. J., Stryker, W. S., Hernandez, M., Kaplan, L., Sober, A., Bosner, B., and Willett, W. C. (1990). Food predictors of plasma beta-carotene and alpha-tocopherol: Validation of a food frequency questionnaire. *Am. J. Epidemiol.* 13, 864–876.

Rose, G. (1990). Preventive cardiology: What lies ahead. *Prev. Med.* 19, 97–104.

Rothman, K. (1986). "Modern Epidemiology." Little Brown, Boston.

Rous, P., and Kidd, J. G. (1941). Conditional neoplasms and subthreshold neoplastic states. *J. Exp. Med.* 73, 365–389.

Roux, W. (1883). "Über die Bedeutung der Kerntheilungs Figuren." Engelmann, Leipzig.

Rowley, J. D. (1973). A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature (London)* 243, 290–293.

Rowley, J. D. (1990). Molecular cytogenetics: Rosetta Stone for understanding cancer. Twenty-ninth GHA Clowes Memorial Award Lecture. *Cancer Res.* 50, 3816–3825.

Saiki, R. K., Arnheim, N., and Erlich, H. A. (1985). A novel method for detection of polymorphic restriction sites by clearance of oligonucleotide probes: Application to sickle cell anemia. *BioTechnology* 3, 1008–1012.

Suzuki, D. T., Griffiths, A. J. F., Miller, J. H., and Lewontin, R. C. (1986). "An Introduction to Genetic Analysis." Freeman, New York.

Schmidt, C. (1850). "Characteristics of Epidemic Cholera." Rehyer, Leipzig.

Schulte, P. A. (1987). Methodologic issues in the use of biologic markers in epidemiologic research. *Am. J. Epidemiol.* 126, 1006–1016.

Schulte, P. A. (1989). A conceptual framework for the validation and use of biologic markers. *Environ. Res.* 48, 129–144.

Schulte, P. A. (1991). Contributions of biological markers to occupational health. *Am. J. Ind. Med.* 20, 435–446.

Schwann, T. (1839). "Mikroskopische Untersuchungen über die Übereinstimung in der Struktur und dem Wachstum der Tiere and Pflanzen." Springer-Verlag, Berlin.

Shields, P. G., and Harris C. C. (1991). Molecular epidemiology and the genetics of environmental cancer. *J. Am. Med. Assoc.* 246, 681–687.

Snow, J. (1855). "On the Mode of Communication of Cholera. (2nd ed.) Churchill, London. Reproduced in "Snow and Cholera." Common-Wealth Fund, New York, 1936. Reprinted by Hafner, New York, 1965.

Sobels, F. H. (1982). The parallelogram: an indirect approach for the assessment of genetic risks from chemical mutagens. *Prog. Mut. Res.* **3**, 323–327.

Soussi, T, Caron de Fromentel C., May, P. (1990). Structural aspects of the p53 protein in relation to gene evolution. *Oncogene* **5**, 945–52.

Stein, R. E. K., and Jessop, D. J. (1988). Thoughts on interdisciplinary research. *J. Clin. Epidemiol.* **41**, 813–815.

Suruda, A., Schulte, P., Boeniger, M., Hayes, R. B., Livingston, G. K., Steenland, K., Stewart, P., Herrick, R., Douthitt, D., and Fingerhut, M. A. (1992). Cytogenetic effects of formaldehyde exposure in students of mortuary science. (*in press*)

Tannenbaum, S. R., and Skipper, P. L. (1984). Biological aspects to the evaluation of risk: Dosimetry of carcinogens in man. *Fund. Appl. Toxicol.* **4**, 5367–5370.

Taylor, J. A., Sandler, D. P., Bloomfield, C. D., Shore, D. L., Ball, E. D., Neubauer, A. R., McIntyre, O. R., and Liu, E. (1992). *Ras* oncogene activation and occupational exposures in acute myeloid leukemia. *J. Natl. Cancer Inst.* **84**, 1626–1632.

The American Heritage Dictionary (1976). (W. Morris, ed.), p. 289. Houghton Mifflin, Boston.

Thilly, W. G., Phaik-Mooi, L., and Skopek, T. R. (1982). Potential of mutational spectra for diagnosing the cause of genetic change in human cell populations. *Banbury Report* **13**, 453–465.

Thompson, S. G. (1983). A method of analysis of laboratory data in an epidemiological study where time trends are present. *Stat. Med.* **2**, 147–153.

Tola, S., and Hernberg, S. (1981). Strategies of biological monitoring. *In* "Recent Advances in Occupational Health" (J. C. McDonald, ed.), pp. 199–209. Churchill Livingstone, Edinburgh.

Trucco, M., and Dorman, J. S. (1989). Immunogenetics of insulin-dependent diabetes mellitus in humans. *CRC Crit. Rev. Immunol.* **9**, 201–245.

Truett, J., Cornfield, J., and Kannel, W. (1967). A multivariate analysis of coronary heart disease in Framingham. *J. Chronic Dis.* **20**, 511–524.

Vandenbroucke, J. P. (1988). Is 'the causes of cancer' a miasma theory for the end of the twentieth century? *Int. J. Epidemiol.* **17**, 708–709.

van Noord, P. A. H. (1992). "Selenium and Human Cancer Risk: Nail Keratin as a Tool in Metabolic Epidemiology." Thesis Publishers, Amsterdam.

Vecchio, T. J. (1966). Predictive value of a single diagnostic test in unselected populations. *N. Engl. J. Med.* **274**, 1171–1173.

Venzmer, G. R. (1968). "Five Thousand Years of Medicine." Taplinger, New York.

Vineis, P., and Caporaso, N. (1988). Applications of biochemical epidemiology in the study of human carcinogenesis. *Tumori* **74**, 19–26.

Vineis, P., Faggiano, F., and Terracini, B. (1990). Biochemical epidemiology: Uses in the study of human carcinogenesis. *Terat. Carcin. Meta.* **10**, 231–237.

Virchow, R. (1858). [English translation 1971]. "Cellular Pathology." Dover, New York.

Vogel, F. (1959). Modern problems in human genetics. *Ergeb. Inn. Med. Kinderheilkd.* **12**, 52–125.

Vogelstein, B., Fearon, E. R., Hamilton, S. R., Kern, S. E., Preisinger, A. C., Leppert, M., Nakamura Y., White, R., Smits, A. M. M., and Bos, J. L. (1988). Genetic alterations during colorectal tumor development. *N. Eng. J. Med.* **319**, 525–532.

Waldrop, M. M. (1992). Finding RNA makes proteins give 'RNA world' a big boost. *Science* **256**, 1396–1397.

Watson, J. D., and Crick, F. H. (1953). Molecular structure of nucleic acids in a structure for deoxyribose nucleic acid. *Nature (London)* **171**, 737–740.

Weatherall, D. J. (1991). "The New Genetics and Clinical Practice" 3d Ed. Oxford University Press, Oxford.

Wetmur, J. G., Lehnert, G., and Desnick, R. J. (1991). The delta-aminolevulinate dehydratase polymorphism: Higher blood lead levels in lead workers and environmental exposed children with the 1-2 and 2-2 isozymes. *Environ. Res.* **56**, 109–119.

White, C. (1973). Preface. *In* "Serological Epidemiology" (J. R. Paul and C. White, eds.) Academic Press, New York.

Wild, C. P., and Montesano, R. (1991). Detection of alkylated DNA adducts in human tissues. *In* "Molecular Dosimetry and Human Cancer" (J. D. Groopman and P. L. Skipper, eds.), pp. 263–280. CRC Press, Boca Raton, Florida.

Willett, W. C., Stampfer, M. C., Underwood, B. A., Speizer, F. E., Rosner, B., and Hennekins, C. H. (1983). Validation of a dietary questionnaire with plasma carotenoid and alpha-tocopherol. *Am. J. Clin. Nutr.* **38**, 631–639.

Williams, J. W. (1920). The value of the Wasserman reaction in obstetrics based upon the study of 4547 consecutive cases. *Johns Hopkins Hosp. Bull.* **31**, 335–342.

Wogan, G. N., and Gorelick, N. J. (1985). Chemical and biochemical dosimetry of exposure to genotoxic chemicals. *Environ. Health Perspect.* **62**, 5–18.

Woolf, B. (1955). On estimating the relation between blood group and disease. *Ann. Hum. Genet.* **19**, 251–253.

World Health Organization (1975). "Early Detection of Health Impairment in Occupational Exposure to Health Hazards. (WHO Technical Report No. 571). World Health Organization, Geneva.

Wright, G. L., Jr., and Cox, A. D. (1987). Monoclonal antibodies to human tumor antigens. *In* "Morphological Tumor Markers" (G. Seifert, ed.), pp. 1–18. Springer-Verlag, New York.

Wynder, E. C., and Reddy, B. S. (1974). Metabolic epidemiology of colorectal cancer. *Cancer* **34**, 801–806.

Yach, D. (1990). Biological markers: Broadening or narrowing the scope of epidemiology. *J. Clin. Epidemiol.* **43**, 309–310.

Yallow, R. S., and Berson, S. A. (1960). Immunoassay of endogenous plasma insulin in man. *J. Clin. Invest.* **39**, 1157–1175.

Zielhuis, R. L. (1978). Biological monitoring. *Scand. J. Work Environ. Health* **4**, 1–18.

Zondek, B. (1942). The importance of increased production and excretion of gonadotrophic hormone for diagnosis of hydatidiform mole. *J. Obstet. Gynecol. Br. Emp.* **49**, 397.

Molecular Epidemiology

Principles and Practices

Edited by

Paul A. Schulte

Centers for Disease Control and Prevention
National Institute for Occupational Safety and Health
Robert A. Taft Laboratories
Cincinnati, Ohio

Frederica P. Perera

Department of Epidemiology
Columbia University
New York, New York

Academic Press, Inc.
Harcourt Brace & Company

San Diego New York Boston London Sydney Tokyo Toronto

This book is printed on acid-free paper. ♾

Academic Press, Inc.
1250 Sixth Avenue, San Diego, California 92101-4311

United Kingdom Edition published by
Academic Press Limited
24–28 Oval Road, London NW1 7DX

Library of Congress Cataloging-in-Publication Data

Molecular epidemiology : principles and practices / edited by Paul A. Schulte, Frederica Perera.
p. cm.
ISBN 0-12-632345-3
1. Molecular epidemiology. 2. Biochemical markers. I. Schulte, Paul A. II. Perera, Frederica P.
[DNLM: 1. Biological Markers. 2. Epidemiologic Factors. 3. Molecular Biology. QH 506 M71925]
RA652.5.M65 1993
614.4–dc20
DNLM/DLC
for Library of Congress 92-49193
CIP

PRINTED IN THE UNITED STATES OF AMERICA
93 94 95 96 97 MP 9 8 7 6 5 4 3 2 1