CHAPTER 4

GENETIC EFFECTS IN HUMAN DIETARY STUDIES

ELIZABETH R. DE OLIVEIRA E SILVA, MD; RUTH MCPHERSON, MD, PHD; AND DARRELL L. ELLSWORTH, PHD

INTERACTIONS BETWEEN GENES AND NUTRITION

Metabolic responses to dietary change are complex and vary considerably among individuals. Genetic factors are known to account for a portion of the normal population variation in parameters ranging from fasting serum cholesterol concentration to basal metabolic rate. Genetic variables thus are likely to contribute to individual responses to dietary change and must be considered during in-depth analyses of the biochemical effects of specific nutritional interventions. Studies on the genetic basis of variable dietary response in humans are complicated, however, by the inability to adequately control for a number of confounding environmental factors. For example, the background (prestudy) diet is often difficult to quantify but may have important effects on the response to a specific dietary intervention. Also, although diet can be rigorously controlled if resources permit, other environmental variables that affect the dietary response are more difficult to quantify and regulate.

The biochemical nature of the interactions between dietary response and genetic factors is not well understood, although nutrients have been shown to exhibit specific effects on various levels of gene expression including transcription and messenger RNA (mRNA) processing and stability (1). More research is needed on the effects of individual nutrients on the expression of specific genes and on modulation of gene expression by other genes in order to understand the metabolic basis of individual differences in nutritional responses.

Major gene disorders such as familial hypercholesterolemia have been well characterized despite the fact that they account for only a small percentage of the total population variability in dietary response. Polygenic diversity is also of primary importance in nutrition studies, because multiple genes likely account for a significant portion of the total genetic variance contributing to response *phenotypes.* For example, individual susceptibility to the hyperlipidemic effects of dietary cholesterol and saturated fatty acids, excess alcohol intake, and obesity is believed to have a considerable genetic component. Despite their potential importance, progress in the identification and characterization of the many genes contributing to variability in dietary response has been limited.

GENES INFLUENCING OBESITY AND BODY FAT DISTRIBUTION

Factors influencing obesity and body fat distribution may be important considerations in dietary studies because of the increasing prevalence of obesity in developed countries (2) and associations between weight gain and increased susceptibility to cardiovascular disease (3). The distribution of body fat appears to be an important determinant of the morbidity that is associated with obesity, and abdominal visceral fat in particular is believed to play a critical role in the met-

abolic complications that are commonly observed in obese patients (4–6).

Genetic (7) and environmental (8) factors are both important determinants of interindividual variability in abdominal visceral fat levels and total body fat content. Intervention studies using identical twins first suggested the influence of genetic factors on body fat storage and mobilization (9), and subsequent research has identified numerous candidate genes that may influence the level and distribution of body fat. A summary of all genes and anonymous genetic markers potentially related to obesity syndromes and obesity-related phenotypes is presented within the human obesity gene map (10). To date approximately 133 candidate loci have been identified from various lines of evidence, including association and linkage studies as well as cross-breeding experiments and animal models of obesity. Identifying susceptibility genes associated with candidate loci and evaluating the relative contributions of known candidate genes to obesity in different human populations will likely be an important focus of future research (11).

Study Designs in Genetic Research

Ascertainment of Study Subjects

The purpose and goals of a genetic study examining various phenotypes associated with diet will determine the selection strategy for subjects as well as the type of data that will be collected. Unrelated individuals in matched case-control designs can provide useful genetic information and can be used in association studies to detect relationships between specific measures of a particular phenotype and polymorphisms in candidate genes. However, it is often important to observe the transmission of the phenotype of interest from parents to their children and to identify regions of the genome that are shared by siblings or family members who exhibit similar response phenotypes.

Family studies are often crucial to defining the contributions of genetic factors to complex phenotypes (such as dietary response). Such studies may evaluate as few as two affected siblings or as many as several hundred individuals spanning multiple generations. The goals and specific aims of the research, however, are what dictate the type and number of families that would be optimal. Different family structures have advantages and limitations (see Table 4–1) that must be considered before investigators commit to the expensive and time-consuming tasks of recruitment and screening (12). The method of ascertainment to be used in an investigation must be specified a priori and will be dependent on the research questions being investigated (Table 4–2). Commonly used family groups include monozygotic (MZ) and dizygotic (DZ) twins, sibling pairs, nuclear families (parents and their children), multiplex (multiple-generation) families, and extended pedigrees containing more distant relatives. (Also see Chapter 9, ''Children as Participants in Feeding Studies.'')

Initial Genetic Analyses

Before researchers conduct DNA-based studies with the ultimate goal of localizing and characterizing genes that influence dietary response, it is prudent to determine as much as possible about the genetic basis of the trait of interest. An obvious first step is to define the relative contributions of genetic and environmental factors to the phenotype(s) under study. *Heritability analyses* estimate the proportion of the total phenotypic variance that is caused by genetic variation, as opposed to environmental variation and measurement error. Twin studies, which compare MZ and DZ twins, are used for estimating the proportion of a complex phenotype that is attributable to genetic factors (13, 14). The basic premise is that MZ twins, which result from the division of a single fertilized ovum, share 100% of their genes and are thus genetically identical, whereas DZ twins are expected to share only about 50% of their genes and hence are no more genetically similar than siblings of the same parents born after separate pregnancies. Greater concordance (ie, similarity with respect to the phenotype) between MZ twin pairs relative to DZ twin pairs is believed to reflect a genetic contribution to the trait (15, 16). This is a particularly important issue for complex phenotypes, which may be significantly influenced by the environment. It may be an oversimplification to assume that the greater similarity between MZ twins relative to DZ twins is solely attributable to genetic factors.

Environmental influences may differ between MZ and DZ twins even though twin pairs are generally assumed to share a common environment. Even the concordance of the prenatal environment may not be the same for MZ and DZ twins. MZ twins are formed from a division of a single fertilized ovum. If the division occurs at an early stage—between the zygote and the formation of the embryonic disc—the two embryos will have separate sets of fetal membranes, resulting in dichorionic-diamniotic twins. More commonly, the division occurs after the blastocyst has been formed, resulting in monochorionic-diamniotic twins. Rarely, if the division occurs much later, it results in monochorionic-monoamniotic twins. All dizygotic twins, however, are dichorionic-diamniotic, because they result from the separate fertilization of two ova by two spermatozoa, and each embryo develops within its own set of fetal membranes. Because the placenta develops from the chorionic structures, those MZ twins who share a single placenta will experience a greater similarity of maternal/paternal effects than DZ and MZ twins who develop with separate placentas.

Further studies on the genetics of nutrition and dietary response may be able to delineate models of inheritance to determine the number of genes (and their relative magni-

TABLE 4-1

Population Sampling Strategies for Genetic Research in Dietary Studies

Study Group	Attributes and Applications
Twins	Monozygotic (MZ) and dizygotic (DZ) twins are useful for estimating the proportion of a complex phenotype that is attributable to genetic factors (in heritability and concordance analyses), because they share different proportions of their genes (MZ 100%, DZ 50%) but are assumed to experience similar environmental exposures. Discordant twin pairs provide information regarding environmental influences on complex traits. Environmental influences, however, may differ between MZ and DZ twins and may complicate comparisons.
Nuclear families	Children and their parents are perhaps the easiest family structure to obtain and tend to be most representative of a particular phenotype in the general population. Nuclear families have traditionally been used to localize genes responsible for single-gene (Mendelian) diseases but may not be appropriate for complex phenotypes. This is because it may be labor intensive to identify complete nuclear families in which to observe hereditary transmission for traits with late age of onset, and it may be difficult to detect major genes if multiple genes contribute significantly to the phenotype.
Extended pedigrees	Such pedigrees are often ascertained through a single proband and then extended to more distant relatives. A single large pedigree may provide evidence for linkage, thereby reducing the potential problem of genetic heterogeneity (genetic differences) among families. Transmission of the phenotype of interest and inheritance of genetic markers may be traced across several generations, providing important information for segregation, linkage, and mapping studies. Extended pedigrees may not be representative of most families because the gene(s) contributing to a particular phenotype in a single pedigree may be relatively rare in the general population. Genetic heterogeneity may complicate phenotypic expression if additional genes influencing the phenotype are introduced by spouses into the extended family.
Siblings	Affected sibling-pair methods have recently become popular for localizing disease susceptibility genes because these methods are amenable to nonparametric (model-free) methodologies; they use allele-sharing methods that are conceptually and computationally simple. Discordant siblings provide significant power to detect linkage, which reduces the number of individuals needed in the analyses. Despite their popularity, sib-pair methods have several disadvantages, including reliance on identity-by-state (IBS) estimates of allele sharing rather than identity-by-descent (IBD), and potentially large sample sizes because many sibling pairs may be required to detect genes having relatively modest effects on complex phenotypes.
Inbred or isolated populations	Inbred or genetically isolated populations may facilitate the identification of genes influencing a complex trait due to reduced genetic heterogeneity—only a small number of genes influencing the trait may be present in that population. Conversely, genes affecting the trait may not be identifiable because many of the susceptibility genes that affect the general population may not be present in the isolate.
Unrelated individuals	Unrelated individuals may be used in association studies that compare the frequency of alleles between affected individuals (cases) and healthy individuals of similar age and gender (controls). Subjects are typically recruited based on the phenotype of interest which may be present or absent (in the case of disease). However, recruitment may also sample individuals from the extremes of the population distribution for a given trait. Unrelated individuals do not provide information on the transmission of genes or phenotypes from parents to children and thus are not suitable for linkage analyses.

TABLE 4-2

Study Designs for Genetic Research in Dietary Studies[1]

Objective	Potential Study Groups
Estimate the proportion of a complex phenotype that is attributable to genetic factors.	Monozygotic (MZ)/dizygotic (DZ) twins Siblings raised separately
Explain inheritance patterns and the number of genes contributing to a particular trait.	Nuclear families ascertained through (1) an affected parent (2) two or more affected siblings Extended pedigrees
Localize genes influencing a complex phenotype.	Affected sibling pairs Nuclear and extended families Inbred or genetically isolated populations
Examine relationships between DNA variants and phenotypes of interest.	Unrelated individuals (cases vs controls) Trios (a subject and his or her parents)

[1]Adapted from DS Postma and DA Meyers, in SB Liggett and DA Meyers, eds, *The Genetics of Asthma* (New York: Marcel Dekker, Inc; 1996), p 447, by permission of Marcel Dekker, Inc.

tude) influencing a particular trait. *Segregation analyses* attempt to explain the inheritance patterns of a given trait in families (pedigrees) by ''fitting'' various models to the observed patterns of inheritance and expression (17). Models range from simplistic, in which a single gene is sufficient to account for the observed familial correlations, to highly complex, in which many genes interact in controlling inheritance and phenotypic expression.

Localizing and Identifying Genes

Two basic approaches may be implemented to identify genes influencing individual metabolic responses to diet. The study of *candidate genes* is restricted to examining known genes of presumed metabolic relevance (18). Genes become candidates for involvement in complex traits when detailed biological and physiological information about the phenotype suggests that the action or product of the gene may be involved in early developmental pathways (presymptomatic), in the biochemical and cellular processes of progression, or in clinical manifestations (19). Complex phenotypes that may be pertinent to dietary studies include metabolic processes (such as thrombogenesis, the regulation of plasma low-density lipoprotein [LDL] kinetics, or calcium absorption), and may extend to common chronic diseases such as atherosclerosis and heart disease. The candidate gene approach is often limited because many important genes involved in regulating dietary response are unknown or have not been well characterized, and because investigators often have differing opinions regarding which genes should be considered viable candidates for influencing a particular phenotype. (See Figure 4–1 for an example of gene structure.)

An alternate approach for localizing genes associated with complex phenotypes such as metabolic response to diet involves a *genome-wide scan* in families or affected sibling pairs using genetic markers at regularly spaced intervals throughout the genome (eg, 20, 21). Linkage statistics can help to define initial candidate regions that must be further characterized by fine-mapping and physical mapping techniques to locate the gene(s) of interest.

Linkage analyses are commonly used to identify genes contributing to a particular phenotype or disease; the strength of the evidence for linkage is compared to the strength of the evidence for no linkage (22). Linkage is caused by the close proximity of two genes (or markers) along a chromosome and is inferred to exist when cosegregation is detected more often than would be expected by chance between a genetic marker and a gene affecting the phenotype of interest.

Traditional linkage approaches for localizing genes contributing to single-gene disorders may not be appropriate for the study of complex phenotypes such as dietary response, which is likely influenced by many genes. This is because *alleles* (alternate forms of a gene that differ in their DNA sequence) associated with individual differences in response may be common in the general population and any one gene may contribute only modestly to the response phenotype. Affected sib-pair methods have become popular for localizing genes that confer increased susceptibility to complex diseases. Such methods have increased in popularity because they are amenable to nonparametric (model-free) statistical methodologies and use techniques that are relatively simple from both conceptual and computational standpoints (23–25). For these reasons, nonparametric methods may generally be appropriate for nutrition studies, which often examine complex phenotypes.

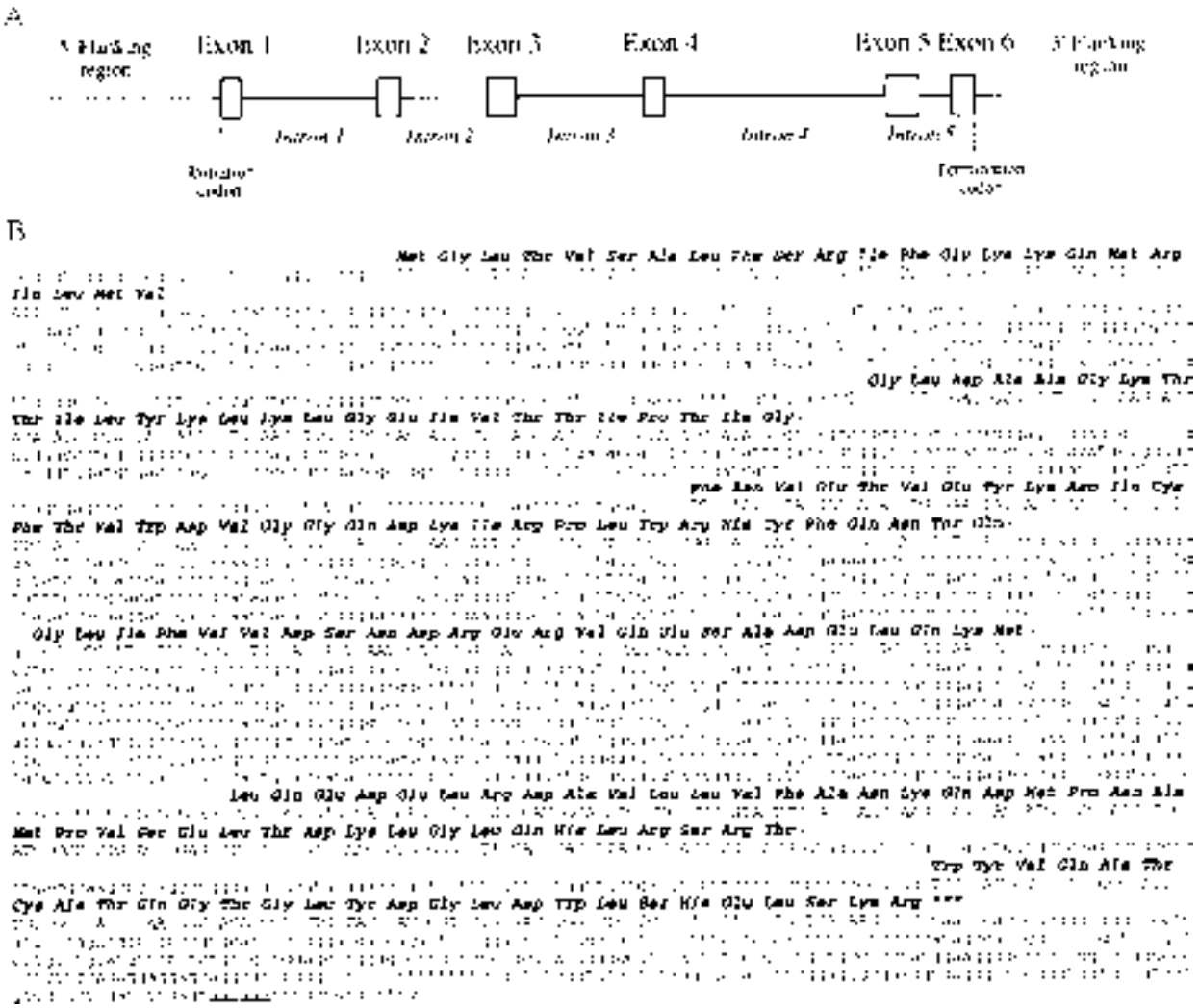

FIGURE 4-1. (A) Schematic drawing representing the overall structure of a typical gene. The gene depicted is the human ADP-ribosylation factor 5 (ARF5) gene with exon and intron sizes drawn approximately to scale. Nucleotide sequence has been determined for the main portion of the gene (solid line) but is unavailable for part of the flanking regions (dashed line). The 5′ flanking region contains the promoter—the region of the DNA to which an RNA polymerase binds to initiate transcription (production of messenger RNA). Exons contain the information that is translated into the amino acid sequence of the gene product (protein), whereas introns (intervening sequences) are removed from the messenger RNA and thus do not encode amino acid sequence. The 3′ flanking region contains the polyadenylation signal that functions in processing the messenger RNA transcript once it has been synthesized. (B) Nucleotide sequence of the human ARF5 gene with exons depicted in uppercase letters and introns/flanking regions in lowercase letters. The predicted amino acid sequence (abbreviated) is listed above the corresponding DNA sequence, the termination codon is indicated by asterisks, and the polyadenylation signal (aataaa) is underlined. Reprinted with permission of Academic Press from McGuire et al (94).

Identification and Analysis of DNA Sequence Variation

Once potential candidate genes have been identified, through either a compilation of genes of known function or a genome scan, genetic polymorphisms (variations in the DNA sequence) within these genes may be examined. In the event that the candidate genes have not been well characterized, an exhaustive search must be conducted for DNA variation within the candidate gene region. A variety of scanning techniques is available for the initial detection of unknown mutations (26).

Characterized genetic polymorphisms may be genotyped in a particular study population in one of several ways. *Restriction fragment length polymorphisms (RFLPs)* are genetic variants that occur within a restriction site, a short DNA sequence recognized by restriction enzymes that cut the DNA only at that specific sequence (Figure 4–2). Many copies of the region can be generated by *polymerase chain reaction (PCR)* amplification, followed by restriction enzyme digestion and resolution of the resulting fragments on acrylamide or agarose gels (eg, 27, 28). Polymorphisms that do not occur within restriction sites can be assayed with allele-specific oligonucleotides and the *oligonucleotide*

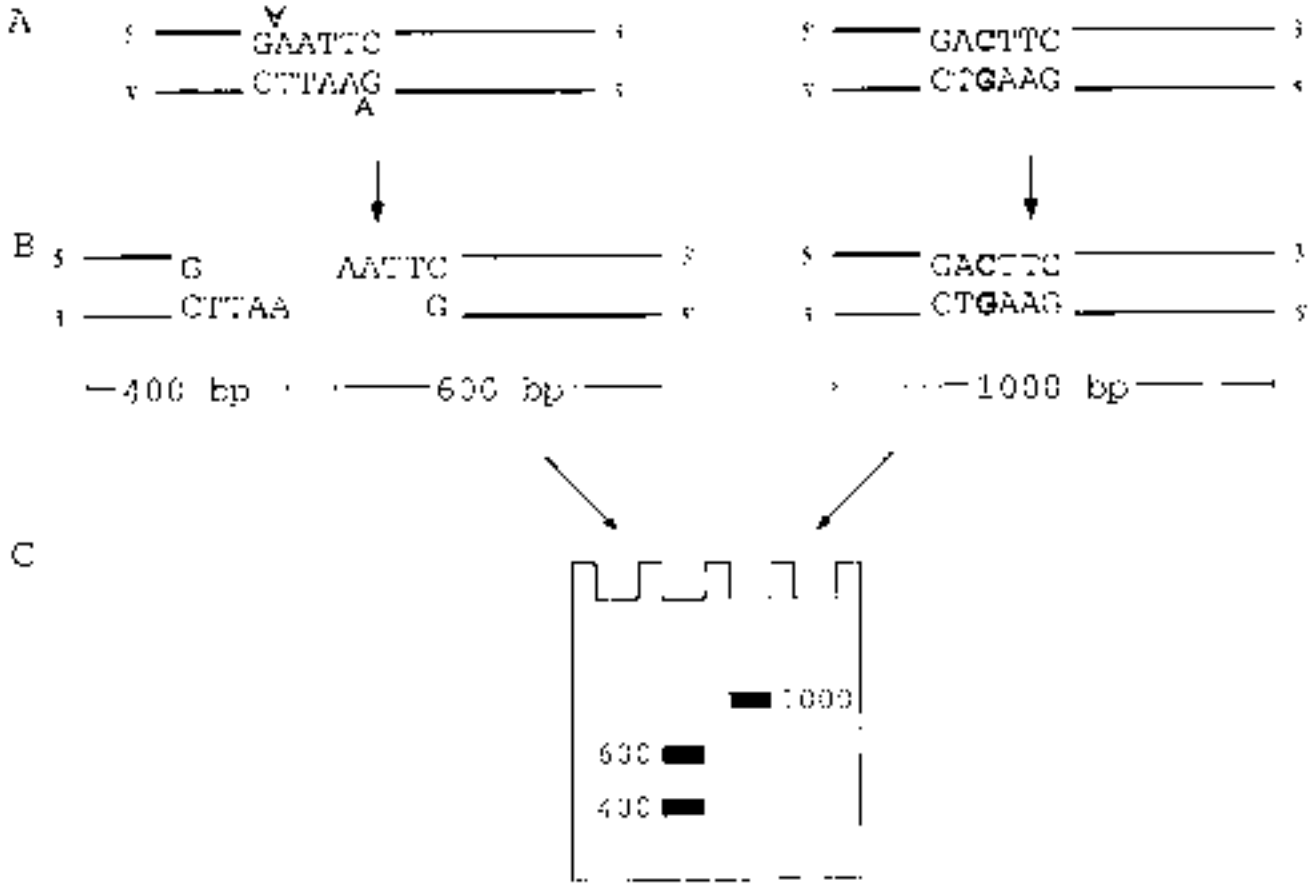

FIGURE 4-2. Representation of a restriction fragment length polymorphism (RFLP). To characterize an RFLP genetic marker, many copies of the DNA region containing the RFLP can be generated using the polymerase chain reaction. (A) A short DNA sequence (5′-GAATTC-3′) that is recognized by the restriction enzyme *Eco* RI (isolated from the bacterium *Escherichia coli*) is located within the amplified fragment of DNA on the left. The restriction enzyme will cut both strands of the DNA within this sequence between the bases indicated by open arrowheads. A single base substitution (in bold) has altered the restriction site (5′-GACTTC-3′) such that *Eco* RI will not cut the DNA fragment on the right. (B) The fragment on the left has been cut into two smaller fragments. One fragment is 400 base pairs in length whereas the other fragment is 600 base pairs long. The fragment on the right has not been cut and remains 1,000 base pairs long. (C) The fragments resulting from the restriction digestion can be separated and visualized on an agarose or polyacrylamide gel. The DNA is applied to the gel in the wells at the top and migrates toward the bottom of the gel when an electric current is applied. Fragments are separated by size and when visualized produce a characteristic banding pattern.

ligation assay (OLA) (29). Emerging techniques such as genetic bit analysis (GBA) (30) and DNA chip technologies (31, 32), which use miniaturized arrays of custom DNA fragments (DNA chips), are automated, high-throughput methodologies that have the capacity to rapidly generate *genotype* and DNA sequence information.

Genetic polymorphisms can then be analyzed by statistical methods in populations or families with respect to particular outcomes (such as lipoprotein response to diet). For example, association studies compare the frequency of alleles between unrelated individuals who exhibit the phenotype of interest (cases) relative to those who do not (controls). An allele is considered to be associated with a given phenotype if that allele occurs at a significantly higher frequency in the case group relative to the controls (17).

Many genetic polymorphisms identified in population surveys do not alter gene function or expression either because they do not result in an amino acid substitution in the corresponding protein (silent or synonymous substitutions) or because they occur in noncoding regions of the gene such as introns. However, such markers may be statistically associated with a particular phenotype because they are in *linkage disequilibrium* with true functional mutations. Alleles exhibiting linkage disequilibrium often occur at loci that are physically close to one another (tightly linked) and tend to be inherited together because they are located on the same chromosome.

Transmission disequilibrium tests (TDT) may help avoid some of the potential problems that may be encountered in simple association studies as a result of linkage disequilibrium. Most TDT statistics examine the transmission of alleles from heterozygous parents to those children exhibiting the phenotype of interest. The transmission of an allele associated with the phenotype will be greater than that ex-

pected under random (Mendelian) segregation if the marker is located close to the phenotype-associated gene (33, 34).

Dietary and Genetic Influences on Plasma Lipids and Lipoproteins

Metabolic and Genetic Basis of Dietary Cholesterol Responsiveness

Variability among individuals in response to diet-induced hypercholesterolemia has been observed in numerous animal species and has been attributed to a variety of genes that regulate lipid and lipoprotein metabolism. Insights from animal models first suggested the influence of genetic factors in cholesterol responsiveness to diet (35) and have subsequently identified biochemical mechanisms and novel candidate genes that may alter dietary sensitivity (36, 37).

Cholesterol absorption efficiency and alterations in bile acid synthetic capacity appear to be significant correlates of the hypercholesterolemic effect of dietary fat and cholesterol in animals. For example, an inbred substrain of New Zealand white rabbits (CRT/mlo) can markedly upregulate or increase bile acid synthesis in response to a cholesterol challenge. By increasing expression of cholesterol 7 α-hydroxylase, the putative rate-limiting enzyme in bile acid synthesis, CRT/mlo rabbits are able to dispose of excess dietary cholesterol and thus resist hypercholesterolemia and atherosclerosis (38).

Similarly, elevated expression of sterol 27-hydroxylase, another important enzyme regulating bile acid synthesis, appears to mitigate diet-induced hypercholesterolemia in baboons (*Papio* spp.) subjected to a high-cholesterol, high-fat diet (39). African green monkeys (*Cercopithecus aethiops*) may downregulate intestinal cholesterol absorption by reducing bile acid production in response to increased dietary cholesterol. Regulation of cholesterol 7 α-hydroxylase activity and abundance may attenuate bile acid secretion; this may be an important adaptive strategy for reducing cholesterol absorption in nonhuman primate species (40).

Early metabolic studies in humans recognized variability in dietary response among individuals and attempted to quantify intrinsic cholesterol responsiveness to diet (41, 42). Marked heterogeneity of response to dietary cholesterol and fatty acid intake has been consistently observed in virtually all dietary studies despite rigorous dietary control (43–47). For example, increases in LDL-cholesterol (LDL-C) varied from 0% to 62% in young men consuming 80 or 320 mg cholesterol per 1,000 kcal in an otherwise constant diet (47). Similarly, Mistry et al (43) demonstrated that consumption of 1,500 mg cholesterol per day for 2 weeks increased plasma total cholesterol concentrations by 4 to 75 mg/dL in 32 subjects, but in 5 subjects the diet produced no change or a slight decrease in cholesterol. Most studies indicate that approximately 10% of individuals are hyperresponders (ie, show significant changes), whereas a similar proportion exhibit little or no change in plasma cholesterol in response to specific alterations in dietary fat and cholesterol.

Although dietary responsiveness appears to be inversely correlated with habitual cholesterol consumption (a greater response is typically associated with lower baseline cholesterol levels) (48), investigations of individual response to dietary cholesterol suggest that multiple metabolic variables may play a role. These variables may include: failure to downregulate endogenous cholesterol synthesis in response to dietary cholesterol (43, 45, 49); linear increases in cholesterol absorption with increasing dietary cholesterol intake (50); and, in individuals most sensitive to the hypercholesterolemic effects of dietary fat (51), low fractional catabolic (degradation) rates of LDL apoB, suggesting a marked downregulation of hepatic LDL receptors.

The numerous genes involved in the complex pathways of lipoprotein metabolism are likely to contribute to variation in dietary responsiveness (52). Individual responses to changes in dietary fatty acid and cholesterol intake may be driven by genetic factors regulating cholesterol absorption and excretion, chylomicron remnant clearance, and downregulation of the LDL receptor and HMG-CoA reductase. In addition, metabolic responses to dietary change, including alterations in apolipoprotein A-I (apoA-I) and apoB synthesis and the activity of hepatic lipase and lipoprotein lipase, are subject to genetic variation. Increasing knowledge of the human genome and advances in molecular biology are anticipated to uncover novel genes and genetic mechanisms that influence complex phenotypes (53) and provide the framework for testing the involvement of putative candidate genes in dietary response.

Candidate Genes Regulating Plasma Lipid Response to Diet

Numerous studies have reported associations between common genetic variation at the apolipoprotein E (apoE) gene and variation in plasma lipid levels, atherosclerotic involvement, and cardiovascular disease (54–56). ApoE is a key component of plasma lipid metabolism and plays a central role in the catabolism of triglyceride-rich lipoprotein remnants (57). The variant forms of apoE have been shown to effect differences in total serum cholesterol levels because of differential binding affinities for lipoprotein receptors in the liver and differences in the metabolism of lipoproteins.

The average effect of the ε4 allele is to raise total plasma cholesterol levels, possibly through efficient uptake and conversion of lipoprotein particle remnants to LDL (58, 59). Substantial data indicate an association between the ε4 allele and elevated levels of LDL-C and total cholesterol that are strongly related to atherosclerotic potential (60, 61). Conversely, the ε2 allele is often associated with lower plasma cholesterol levels and may thus have a protective effect against the development of coronary disease (62, 63).

Clinical and population studies on the effects of the apoE gene on sensitivity to dietary fat and cholesterol report

that the ε4 allele is associated with marked increases in LDL-C in response to dietary cholesterol (64–66). Although Kesaniemi et al (67) reported an increase in cholesterol absorption in subjects with an apoE ε4 allele, inconsistent results have been reported from a number of additional studies, such as Lefevre et al (68), suggesting that multiple genes are involved in determining dietary sensitivity and that the relative physiological effects of these genes may differ among populations (69, 70).

Additional genes that may affect dietary fat and cholesterol absorption and/or influence lipoprotein metabolism include apolipoprotein A-IV (apoA-IV), apolipoprotein B (apoB), and lipoprotein lipase. A glutamine → histidine substitution (Gln360His) near the C-terminus of the apoA-IV protein distinguishes apoA-IV-1 and apoA-IV-2 alleles that are present in white populations at frequencies of approximately 90% and 8%, respectively (71). The glutamine (in the apoA-IV-1 form) → histidine (in apoA-IV-2) amino acid substitution adds a positive charge to the apoA-IV protein but does not appear to alter its primary or secondary structure. The average effect of the A-IV-2 allele is to raise high-density cholesterol (HDL) and to lower triglyceride levels (72). A recent dietary-modification study (73) suggested that the A-IV-2 allele may attenuate short-term hypercholesterolemic responses to high levels of dietary cholesterol.

Genetic variation within the apoB gene is associated with alterations in plasma lipid and triglyceride levels and is also linked to obesity, high cholesterol levels, and an increased risk of coronary artery disease (74, 75). Specific mutations in apoB may alter the binding affinity of low-density lipoproteins for the low-density lipoprotein receptor and thus significantly affect rates of LDL catabolism and clearance (76, 77). Individuals homozygous for the absence of an *Xba* I restriction site in the apoB gene appear to exhibit the greatest sensitivity to dietary saturated fat and cholesterol (69). However, the precise mechanism for the relationship between the *Xba* I polymorphism and dietary sensitivity remains unclear because the site does not appear to alter the receptor binding region of apoB.

Susceptibility to hypertriglyceridemia in response to obesity and alcohol intake also appears to be influenced by genetic factors. For example, a mutation in the lipoprotein lipase gene that causes asparagine to be replaced by serine at position 291 may predispose individuals to elevated plasma triglyceride levels in conjunction with certain physiological characteristics such as increased body mass index (BMI) (78).

High-Density Lipoprotein Turnover Studies

Environmental and hormonal influences account for a significant fraction of the variance for HDL-cholesterol (HDL-C) in the general population; however, family studies provide evidence that genetic factors are also important. Twin studies have shown a much higher correlation between HDL-C levels in MZ twins (0.68–0.74) relative to DZ twin pairs (0.34–0.46) (79–81).

A series of lipoprotein turnover studies in the laboratory of one of the authors (De Oliveira) has been examining the genetic and environmental influences of HDL-C response to diet by studying HDL turnover (a measure of HDL metabolism) under metabolic ward conditions in MZ and DZ twins. Intrapair correlations can be more accurately estimated in an inpatient setting because environmental variables can be minimized and repeat sampling can be conducted. The HDL turnover studies evaluate healthy individuals with various HDL levels to determine the relative contributions of synthesis and catabolism of HDL proteins in determining plasma HDL levels. During the 4-week trials, subjects consume a carefully controlled (metabolic) diet, receive a protein tracer injection, and donate blood and urine samples on an approximately daily basis during the final 2 weeks of the diet period. Information about the logistical aspects of conducting dietary studies with twins is provided in Tables 4–3 and 4–4.

Differences in the intrapair correlations for a given parameter such as the rate of HDL metabolism between MZ and DZ twins can be used to estimate heritability. Heritability estimates use MZ and DZ twins to determine the relative contribution of genetic factors to a particular phenotype under the assumption that the environments for both identical and fraternal twins are similar. However, this assumption may not be valid because identical twins are often subjected to more similar microenvironments (diet, activity, education) than fraternal twins and may even have a greater similarity of dietary habits. In fact, food and nutrient intake may be influenced by the genetic background of the individual. Dairy products, for example, are eaten less often by persons with lactase deficiency. Genetic mechanisms also may affect food selection apart from the components of total energy intake that are related to body weight and energy requirement; the distribution of dietary energy among macronutrient sources (ie, protein, fat, and carbohydrate) has been reported as more similar in MZ twins than in DZ twins (82). (Also see reference 81.)

Furthermore, twin studies have traditionally been conducted with free-living individuals not subject to dietary control or repeat sampling. Violation of the identical environments assumption and lack of dietary control may affect the intrapair correlations and may bias (inflate) estimates of genetic heritability. Despite the possibility that MZ twins may share a more similar environment than DZ twins, twin studies are highly suggestive of genetic influences on HDL-C levels.

Heritability Estimates, Sample Sizes, and Power Calculations

Heritability estimates have been used extensively to examine the genetic components of a broad range of phenotypes, and a variety of complex traits and disease syndromes have been

TABLE 4-3

Challenges of Conducting Controlled Diet Studies with Twins

Issues and Implications	Relevant Information and Potential Solutions
Twins are relatively rare in populations: It can be difficult to recruit adequate numbers to achieve sufficient statistical power within a satisfactory time frame. Recruiting activities may require extra time, money, and labor.	Twins comprise about 1% of live births. The rate of twinning varies according to ethnic group and geographical region, a factor that can affect recruitment strategies. Demographically diverse study populations should be recruited for twin studies, as for other research projects, unless justified on scientific grounds. Recruitment techniques for twin studies include: advertisements in local and national newspapers, magazines, and radio stations; signs posted at colleges and hospitals; informal ''word-of-mouth'' referrals from enrolled participants; recruitment materials included in mailings from twins organizations; and materials distributed at information booths at national twins meetings. Accrual of twin pairs can be slow, with consequences for budget, personnel, data collection, and interpretation and publication of results. Slow accrual indicates that recruitment efforts, and possibly the scientific goals of the protocol, should be evaluated. Recruitment materials can highlight (in realistic terms) the unique opportunities and benefits of participation, such as: confirming zygosity status, learning about personal health and medical situations, experiencing once more a common living situation with their twin (perhaps after years of being geographically separated). Twins may be surprisingly receptive to participation in research protocols because they often have long experience of other people's curiosity and may even have a sense of being a part of a ''natural experiment.''
Study outcomes usually require complete data for both twins: If one twin leaves the study, the partial data for both twins would not be usable in most cases. Screening procedures thus must ensure that both twins are eligible and likely to complete the study. Management techniques to enhance compliance and retention are critical to prevent loss of data.	Recruitment for twin studies is rendered more difficult than for other types of controlled diet studies by the requirements that both members of the pair must fit the eligibility criteria and be able and willing to participate. The obstacles to enrollment of twins are similar to those faced by other individuals. These include family commitments, financial problems, work schedules, reluctance to undergo study procedures, and unwillingness to eat the study food. Logistical support with arrangements may help potential participants to overcome difficulties in enrollment. This support structure may include providing transportation subsidies, allowing spouses and relatives to visit periodically and to stay nearby, providing long-distance telephone services, and covering other expenses associated with participating. Financial incentives may be helpful if allowed by the research institution. Family social obligations (such as weddings and graduations) or emergencies (such as sickness or death) are likely to affect both members of a twin pair. Allowing a one-day absence with prepacked study meals can help to address such situations.

Continued

shown to have significant heritable components. These estimates are expressed as h^2, which is derived from the fraction of trait variance caused by heritable factors. Values range from 0 (no genetic contribution) to 1 (complete genetic control). Assuming a trait is 100% genetic ($h^2 = 1$), the correlation between DZ twins would be 50% of the correlation between MZ twins. However, if a trait were not influenced by genetic factors ($h^2 = 0$), correlations between MZ and DZ twins would be identical. To estimate statistical power in the HDL-C study mentioned earlier, the heritability

TABLE 4-3

Issues and Implications	Relevant Information and Potential Solutions
	Particular care must be exercised in the initial selection of participants. A thorough medical and social history often will help researchers in identifying which twins can be considered for enrollment, and whether problems are likely to develop. The twin pair must be considered as the unit of recruitment, and both individuals must be good candidates for compliance and retention. It is prudent to require a series of ''run-in'' activities that will evaluate the cooperation of both members; these may include attending clinic visits, participating in interviews, filling out questionnaires, and eating several sample study meals. The potential subjects should be provided with a detailed explanation of all study protocols. The importance of good compliance should be discussed in detail.
Identification errors are likely to occur: Monozygotic twins are expected to be visually identical, but dizygotic twins also can look very much alike. Twins also often have similar-sounding names with identical first and last initials. Misidentification errors thus can occur during the distribution of food and medication and during the collection and labeling of biological samples such as blood or urine.	Conduct rigorous training for staff regarding the rationale of the protocol and the procedures involved. This should include advance consideration of procedures and activities susceptible to error. Hold a joint briefing for subjects and staff before the study begins, including question and answer sessions. This enhances the psychological ''investment'' of the subjects in the outcome of the study. Expand the role of subjects to include an appropriate degree of responsibility for their own protocols. This can include checking labels and procedure times and telling the staff about misidentifications. Build safeguards into the protocol, study materials, and procedures, including: double-checking labels and verifying names before distributing food and medications and conducting procedures; using full first and last names, rather than initials, for labeling food items, sample tubes, etc; establishing unique identifiers for collected samples; color-coding dietary treatments, food and medicine labels, and other items. Labeling or identification errors made with one twin probably will be repeated with the other twin (for example, in providing food items or timing blood draws).
Each twin has individual rights as a research subject: Decisions to participate or withdraw must be made independently by each twin.	Investigators must guard against coercive pressures either from the co-twin or from members of the research team. Respect for the individual integrity of each member of the twin pair must be maintained; staff should avoid referring to participants as ''the twins'' or ''the set.''

The information presented in this table draws on the experience of one of the authors (De Oliveira) at the General Clinical Research Center, Rockefeller University, New York, NY, in conducting controlled feeding studies in twins.

of HDL-C was hypothesized to be 0.5–1.0. Given this assumption, a sensitivity analysis ensured that the study possessed adequate power to detect differences between MZ and DZ twin intrapair correlations at a significance level of $P < 0.05$.

Fisher's Z transformation may be used to evaluate the null hypothesis (H_0: $r_{MZ} = r_{DZ}$) that the intrapair correlation coefficients for MZ (r_{MZ}) and DZ (r_{DZ}) twins are identical (83). The distribution of this quantity is approximately normal with a mean of $\frac{1}{2}\log_e[(1+r)/(1-r)]$ and a variance of $1/(n-3)$. H_0 represents a simple test that compares two normally distributed statistics with known variances $1/(n_{MZ}-3)$ and $1/(n_{DZ}-3)$ using the following formula:

$$Z = \frac{\frac{1}{2}\log_e\left(\frac{1 + r_{MZ}}{1 - r_{MZ}}\right) - \frac{1}{2}\log_e\left(\frac{1 + r_{DZ}}{1 - r_{DZ}}\right)}{\sqrt{\frac{1}{n_{MZ}-3} + \frac{1}{n_{DZ}-3}}}$$

The experimental design used equal numbers of MZ and DZ twins. Assuming $r_{MZ} > r_{DZ}$, the Z for the one sided al-

TABLE 4-4

Registries and National Organizations for the Identification and Recruitment of Twins

Organization Name and Address	Telephone/Fax/E-mail Address/Internet Address
International Twins Association (ITA) 6898 Channel Road Minneapolis, MN 55432	Telephone: (612) 571-3022
The Twins Foundation PO Box 6043 Providence, RI 02940-6043	Telephone: (401) 729-1000 Fax: (401) 751-4642 E-mail: twins@twinsfoundation.com Internet: http://www.twinsfoundation.com
National Organization of Mothers of Twins Clubs (NOMOTC) PO Box 23188 Albuquerque, NM 87192-1188	Telephone: (505) 275-0955 or (800) 243-2276 Internet: http://www.nomotc.org
Parents of Multiple Births Association of Canada (POMBA) 240 Graff Ave, Box 22005 Stratford, Ontario, Canada N5A 7V6	Telephone: (519) 272-2203 Fax: (519) 272-1926 E-mail: office@pomba.org Internet: http://www.pomba.org
Twins Days Festival Committee, Inc PO Box 29 Twinsburg, OH 44087	Telephone: (330) 425-3652 Fax: (330) 426-7280

Note: Many twin registries and other similar organizations require that information provided for research purposes be kept confidential. Written consent for use of the information may be required. Some organizations have committees that review and approve research proposals seeking to recruit from their membership. Once approved, the recruitment requests are posted in newsletters and other publications.

ternative was 1.645 at $\alpha = 0.05$. We estimated r_{MZ} to be 0.90 based on preliminary data from 12 MZ twin pairs where the intrapair correlations (under our experimental conditions) were 0.94 for LDL and 0.92 for HDL (84). Under the assumption that $h^2 = 1.0$, r_{DZ} was determined to be 0.45. Substituting these values into the above equation yielded a value for n of 8.55. Similarly, under the assumption that $h^2 = 0.50$, $r_{DZ} = 0.68$ and $n = 15.72$. Therefore, a sample size of 16 MZ and 16 DZ twins was predicted to be adequate to detect 50% heritability of LDL-C, HDL-C, and any other parameters in the protocol given that the assumption $r_{MZ} = 0.90$ is valid.

Ethical and Social Issues in Genetic Studies of Dietary Response

Participant Recruitment

Target Individuals and Exclusion Criteria

Recruitment strategies for genetic studies of dietary response depend on the specific research questions of the particular study. Experimental designs should include both men and women with adequate representation of minority groups.

Often it may be advisable to conduct separate analyses on men and women if there are fundamental gender-specific differences in the etiology and/or pathophysiology of the parameter(s) of interest. Under special circumstances, single-gender studies may be appropriate to investigate phenotypes influenced by sex hormones, such as the effects of estrogen on the genesis of osteoporosis. Medical history and laboratory screening should be conducted to evaluate the overall health status of potential participants. A variety of general exclusion criteria such as the presence of renal, hepatic, hematologic, and immunologic disorders may be appropriate. Specific exclusion criteria for metabolic studies should include alcohol consumption and use of medications known to alter physiological levels of compounds associated with the phenotype under study.

Verifying Family Relationships

Misspecification of family relationships can be problematic in genetic studies and is of particular concern when twins or sibling pairs are examined. Unreported half-sibling (''half-sib'') relationships are often readily evident in nuclear families as departures from Mendelian segregation (the child in question will possess alleles not present in either of the purported parents). However, it may be difficult to determine the actual biological relationships between two individuals believed to be full siblings (''full-sibs'') if no additional

family members (particularly parents) are available. One method, which can be used with data from genome-wide scans that examine many highly polymorphic markers, is to compare the distributions of alleles for the individuals in question with the distributions of alleles observed for known true full-sibling pairs. Alleles *identical by state* will be found more frequently among true full-sibs than among half-sibs (20).

Zygosity determination is an important issue in twin studies because misclassification of twins will bias the resulting heritability estimates. Zygosity has traditionally been determined by self-description and direct observation of physical similarity (85). Initial screenings of study subjects customarily include queries regarding the degree of physical similarity among siblings (ie, questions about whether they are repeatedly confused by family members and friends, or whether they are considered like ''two peas in a pod''; positive answers often suggest monozygosity) (86, 87). Although these methods have been frequently used and found to be fairly reliable (88), questionnaires and self-reports may be inconclusive. A more accurate approach would include verification of zygosity status by molecular genetic techniques. Previous investigations used blood group typings or other serological markers, but with the advent and rapid development of DNA-based ''fingerprinting'' technologies (reviewed in 89), more robust and reliable approaches are now available.

Informed Consent

Within the scientific community it is generally accepted that adequate informed consent must be obtained when investigators conduct research. This includes informing prospective participants that biological specimens (tissue, blood, saliva, or other bodily fluids that can serve as a source of DNA) will be used for genetic analyses. Informed consent procedures use easily understood language to disclose detailed information regarding: the nature and objectives of the research project, potential risks and anticipated benefits that may result from participation, future contact for additional information, the extent to which confidentiality of all research data including genetic results will be maintained, and procedures that will be implemented to minimize inadvertent release of personal information (90, 91).

Participants are told that they have the right to refuse to enroll in the study, as well as the option to withdraw at any time in the future. They are apprised of their right to make specific inquiries about the investigation should they choose to participate. Information is also provided to prospective subjects regarding the conditions for long-term storage of biological samples and potential future use of their DNA in research that may be unrelated to the objectives of the present investigation (92). Informed consent may not be required for retrospective access to residual samples previously collected in conjunction with clinical care, provided the samples are anonymous (lack identifiers) or are anonymized (identifiers removed). Once potential participants have been adequately informed, agreements to participate in the research should be strictly voluntary and should be obtained under conditions free of coercion and undue influence. (Note: Other aspects of informed consent and research ethics are discussed in Chapter 5, ''Ethical Considerations in Dietary Studies.'')

Ethics and Issues of Confidentiality

There is growing public concern regarding the confidentiality of personal medical and genetic information that may be generated as a consequence of participation in research programs. Research records and the identity of program participants must be strictly controlled by the investigators and must remain confidential on publication of research findings. Although ordinary clinical results such as measurements of blood pressure and cholesterol levels are typically provided to each participant (or to his or her personal physician with written permission), genetic findings should not be incorporated into the medical records of participants. Privacy and strict confidentiality of the genetic data are necessary to prevent personal and social stigmatization as well as possible discrimination in insurance or employment (93).

Participants are not usually given the option to receive genetic information about themselves. However, the ability to obtain personal information may be medically beneficial to the participant as well as his/her family. The benefit occurs because individuals who know that they have a genetic predisposition to disease may choose to make beneficial lifestyle changes (selecting a more healthful diet, increasing physical activity, discontinuing tobacco and alcohol use) or to seek preventive treatments that may decrease their risk of developing symptoms. Conversely, such information may lead to ethical dilemmas and difficult personal or reproductive decisions. Personal genetic information should therefore be presented by a certified genetic health care professional or genetic counselor who can educate the subjects as to the benefits and limitations of molecular diagnostic tests and provide counsel regarding therapeutic options or personal issues.

Conclusion

The effects of genetic variation on individual response to dietary change often are of interest in nutrition studies, even if this is not the primary focus of the investigation. Consideration of appropriate and specific genetic effects is likely to enhance the study overall and may strengthen or corroborate the main conclusions. Interactions may be detected between changes in metabolic parameters and genetic variables, which then may provide insight into the interpretation of discrepant outcomes. Even in large family- or population-based dietary studies that do not examine genetic factors, it is of paramount importance to properly collect and store bio-

logical samples (blood samples or buccal swabs) that may serve as a source of DNA for future retrospective examinations of gene-nutrient interactions. Likewise it is vital to obtain proper informed consent that will allow for possible future genetic research.

Novel candidate genes for a variety of complex phenotypes and human diseases are likely to emerge as the Human Genome Project advances and methods for mapping and characterizing genes become more refined. Because most dietary response phenotypes appear to be influenced by multiple genes (each with a relatively small effect) that may interact with other genes and/or the environment, continuing developments in genome technology and statistical methods of analysis will aid our ability to decipher the complexities of dietary response in humans.

Glossary of Terms

Alleles: alternate forms of a gene or genetic locus that differ in DNA sequence. Such differences may or may not affect the function of the RNA or protein product.

Candidate genes: genes believed to influence the expression of complex phenotypes because of known biological and/or physiological properties of their products.

Genome-wide scan: scans used to localize genes contributing to complex diseases by analyzing many highly polymorphic DNA markers spaced at approximately regular intervals throughout the genome. Statistical (linkage) analyses are then used to identify one or more chromosomal regions in which susceptibility genes are believed to reside.

Genotype: the genetic constitution of an organism. The term is often used to refer to the identity of the two alleles at a specific gene or genetic locus.

Heritability (h^2): for a quantitative trait, the proportion of the total phenotypic variation that is attributable to genetic factors. Heritability in the broad sense represents the degree to which a trait is genetically determined; heritability in the narrow sense is the degree to which a trait is transmitted from parents to offspring.

Identity by state (IBS): alleles are considered to be identical by state if they appear to be identical (for example, alleles at microsatellite markers may appear to be the same size and hence inferred to be identical). Conversely, alleles are identical by descent (IBD) if they are descended from the same allele in an ancestral generation. For linkage analyses, IBD information is more informative than IBS.

Linkage analysis: linkage between a genetic marker and a disease susceptibility gene is detected when specific forms (alleles) of the marker and the gene are inherited together (cosegregate) more often than would be expected by chance and is caused by the close proximity of the marker and the gene on a chromosome.

Linkage disequilibrium: the nonrandom transmission from parents to offspring of alleles from genes that are located on the same chromosome. Alleles at tightly linked (located close together) loci are often inherited together; therefore linkage disequilibrium is useful for identifying regions of the genome that historically have been inherited as a linkage group and may identify the approximate location of genes that contribute to disease.

Oligonucleotide ligation assay (OLA): a gel-free assay normally used to characterize known biallelic polymorphisms in a large population. Oligonucleotides (short synthetic fragments of DNA) that are specific to each form of the DNA sequence (each allele) at a particular genetic locus are used to determine the allelic composition of that locus in a large number of individuals.

Phenotype: the observable physical appearance or other properties of an organism that are primarily determined by the genotype but may be influenced by the environment.

Polymerase chain reaction (PCR): a method that has revolutionized the fields of genetics and molecular biology because it can be used to make many copies of (amplify) a particular region of DNA that can then be used in a number of molecular biology applications.

Restriction fragment length polymorphism (RFLP): a genetic marker based on DNA fragments that differ in length due to the presence or absence of specific sequences (restriction sites) that are recognized by certain enzymes. A restriction enzyme will cut the DNA if its recognition sequence is intact but will be unable to cut the DNA if the restriction site has been altered by a mutation (Figure 4–2).

Segregation analysis: analysis whose objective is to explain the inheritance pattern of a particular trait in families by inferring the number of genes influencing the trait, the relative contribution of these genes to the observed phenotype, frequencies of the ''normal'' and ''disease-associated'' alleles, and the proportion of individuals with a given genotype that will exhibit the expected phenotype under specific environmental conditions (penetrance).

Transmission disequilibrium test (TDT): a method used to test for linkage between a genetic marker and a complex phenotype when an association between the phenotype and a genetic variant at the marker has previously been detected. The TDT examines the transmission of the disease-associated allele from a heterozygous parent to an affected child. A particular allele in linkage disequilibrium with the disease susceptibility allele will be transmitted more frequently to affected children than would be expected if all alleles were transmitted randomly.

References

1. Clarke SD, Abraham S. Gene expression: nutrient control of pre- and posttranscriptional events. *FASEB J.* 1992;6:3146–3152.
2. Kuczmarski RJ, Flegal KM, Campbell SM, Johnson CL. Increasing prevalence of overweight among US adults. The National Health and Nutrition Examination Surveys, 1960 to 1991. *JAMA*. 1994;272:205–211.
3. Manson JE, Colditz GA, Stampfer MJ, Willett WC, Rosner B, Monson RR, Speizer FE, Hennekens CH. A

prospective study of obesity and risk of coronary heart disease in women. *N Engl J Med.* 1990;322:882–889.
4. Kissebah AH, Freedman DS, Peiris AN. Health risks of obesity. *Med Clin North Am.* 1989;73:111–138.
5. Després JP, Moorjani S, Lupien PJ, Tremblay A, Nadeau A, Bouchard C. Regional distribution of body fat, plasma lipoproteins, and cardiovascular disease. *Arteriosclerosis.* 1990;10:497–511.
6. Johnson D, Prud'homme D, Després JP, Nadeau A, Tremblay A, Bouchard C. Relation of abdominal obesity to hyperinsulinemia and high blood pressure in men. *Int J Obes Relat Metab Disord.* 1992;16:881–890.
7. Bouchard C, Rice T, Lemieux S, Després J-P, Pérusse L, Rao DC. Major gene for abdominal visceral fat area in the Québec Family Study. *Int J Obes Relat Metab Disord.* 1996;20:420–427.
8. Rice T, Pérusse L, Bouchard C, Rao DC. Familial clustering of abdominal visceral fat and total fat mass: the Quebec Family Study. *Obes Res.* 1996;4:253–261.
9. Bouchard C, Tremblay A, Després JP, Nadeau A, Lupien PJ, Thériault G, Dussault J, Moorjani S, Pinault S, Fournier G. The response to long-term overfeeding in identical twins. *N Engl J Med.* 1990;322:1477–1482.
10. Chagnon YC, Pérusse L, Bouchard C. The human obesity gene map: the 1997 update. *Obes Res.* 1998;6:76–92.
11. Chagnon YC, Pérusse L, Bouchard C. Familial aggregation of obesity, candidate genes, and quantitative trait loci. *Curr Opin Lipidol.* 1997;8:205–211.
12. Postma DS, Meyers DA. Approaches to family studies of asthma. In: Liggett SB, Meyers DA, eds. *The Genetics of Asthma.* New York, NY: Marcel Dekker, Inc; 1996:443–453.
13. Feinleib M, Garrison RJ, Fabsitz R, Christian JC, Hrubec Z, Borhani NO, Kannel WB, Rosenman R, Schwartz JT, Wagner JO. The NHLBI twin study of cardiovascular disease risk factors: methodology and summary of results. *Am J Epidemiol.* 1977;106:284–295.
14. Austin MA, Newman B, Selby JV, Edwards K, Mayer EJ, Krauss RM. Genetics of LDL subclass phenotypes in women twins: concordance, heritability, and commingling analysis. *Arterioscler Thromb.* 1993;13:687–695.
15. Duffy DL. Twin studies in medical research. *Lancet.* 1993;341:1418–1419.
16. Phillips DI. Twin studies in medical research: can they tell us whether diseases are genetically determined? *Lancet.* 1993;341:1008–1009.
17. Khoury MJ, Beaty TH, Cohen BH. *Fundamentals of Genetic Epidemiology.* Cambridge, UK: Cambridge University Press; 1993.
18. Lusis AJ. Genetic factors affecting blood lipoproteins: the candidate gene approach. *J Lipid Res.* 1988;29:397–429.
19. Hajjar DP, Nicholson AC. Atherosclerosis: an understanding of the cellular and molecular basis of the disease promises new approaches for its treatment in the near future. *Am Scientist.* 1995;83:460–467.
20. Hanis CL, Boerwinkle E, Chakraborty R, Ellsworth DL, Concannon P, Stirling B, Morrison VA, Wapelhorst B, Spielman RS, Gogolin-Ewens KJ, Shephard JM, Williams SR, Risch N, Hinds D, Iwasaki N, Ogata M, Omori Y, Petzold C, Rietzsch H, Schröder H-E, Schulze J, Cox NJ, Menzel S, Boriraj VV, Chen X, Lim LR, Lindner T, Mereu LE, Wang Y-Q, Xiang K, Yamagata K, Yang Y, Bell GI. A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. *Nat Genet.* 1996;13:161–166.
21. Nichols WC, Koller DL, Slovis B, Foroud T, Terry VH, Arnold ND, Siemieniak DR, Wheeler L, Phillips JA 3rd, Newman JH, Conneally PM, Ginsburg D, Loyd JE. Localization of the gene for familial primary pulmonary hypertension to chromosome 2q31–32. *Nat Genet.* 1997;15:277–280.
22. Ott J. *Analysis of Human Genetic Linkage.* Baltimore, Md: Johns Hopkins University Press; 1991.
23. Weeks DE, Lange K. The affected-pedigree-member method of linkage analysis. *Am J Hum Genet.* 1988;42:315–326.
24. Amos CI. Robust variance-components approach for assessing genetic linkage in pedigrees. *Am J Hum Genet.* 1994;54:535–543.
25. Risch N, Zhang H. Extreme discordant sib pairs for mapping quantitative trait loci in humans. *Science.* 1995;268:1584–1589.
26. Landegren U, ed. *Laboratory Protocols for Mutation Detection.* Oxford, UK: Oxford University Press; 1996.
27. Hixson JE, Vernier DT. Restriction isotyping of human apolipoprotein E by gene amplification and cleavage with *Hha* I. *J Lipid Res.* 1990;31:545–548.
28. Frosst P, Blom HJ, Milos R, Goyette P, Sheppard CA, Matthews RG, Boers GJH, den Heijer M, Kluijtmans LAJ, van den Heuvel LP, Rozen R. A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. *Nat Genet.* 1995;10:111–113.
29. Nickerson DA, Kaiser R, Lappin S, Stewart J, Hood L, Landegren U. Automated DNA diagnostics using an ELISA-based oligonucleotide ligation assay. *Proc Natl Acad Sci USA.* 1990;87:8923–8927.
30. Nikiforov TT, Rendle RB, Goelet P, Rogers YH, Kotewicz ML, Anderson S, Trainor GL, Knapp MR. Genetic bit analysis: a solid phase method for typing single nucleotide polymorphisms. *Nucleic Acids Res.* 1994;22:4167–4175.
31. Hacia JG, Brody LC, Chee MS, Fodor SPA, Collins FS. Detection of heterozygous mutations in *BRCA1* using high-density oligonucleotide arrays and two-colour fluorescence analysis. *Nat Genet.* 1996;14:441–447.
32. Woolley AT, Sensabaugh GF, Mathies RA. High-speed DNA genotyping using microfabricated capillary array electrophoresis chips. *Anal Chem.* 1997;69:2181–2186.

33. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet.* 1993;52:506–516.
34. Thomson G. Mapping disease genes: family-based association studies. *Am J Hum Genet.* 1995;57:487–498.
35. Clarkson TB, Lofland HB Jr, Bullock BC, Goodman HO. Genetic control of plasma cholesterol: studies on squirrel monkeys. *Arch Pathol.* 1971;92:37–45.
36. McGill HC Jr, Kushwaha RS. Individuality of lipemic responses to diet. *Can J Cardiol.* 1995;11 (suppl G): 15G-27G.
37. Kushwaha RS, McGill HC Jr. Mechanisms controlling lipemic responses to dietary lipids. *World Rev Nutr Diet.* 1997;80:82–125.
38. Poorman JA, Buck RA, Smith SA, Overturf ML, Loose-Mitchell DS. Bile acid excretion and cholesterol 7 α-hydroxylase expression in hypercholesterolemia-resistant rabbits. *J Lipid Res.* 1993;34:1675–1685.
39. Kushwaha RS, Guntupalli B, Rice KS, Carey KD, McGill HC Jr. Effect of dietary cholesterol and fat on the expression of hepatic sterol 27-hydroxylase and other hepatic cholesterol-responsive genes in baboons (*Papio* species). *Arterioscler Thromb Vasc Biol.* 1995;15:1404–1411.
40. Rudel L, Deckelman C, Wilson M, Scobey M, Anderson R. Dietary cholesterol and downregulation of cholesterol 7 α-hydroxylase and cholesterol absorption in African green monkeys. *J Clin Invest.* 1994;93:2463–2472.
41. Jacobs DR Jr, Anderson JT, Hannan P, Keys A, Blackburn H. Variability in individual serum cholesterol response to change in diet. *Arteriosclerosis.* 1983;3:349–356.
42. Keys A, Anderson JT, Grande F. Serum cholesterol response to changes in the diet. III: differences among individuals. *Metabolism.* 1965;14:766–775.
43. Mistry P, Miller NE, Laker M, Hazzard WR, Lewis B. Individual variation in the effects of dietary cholesterol on plasma lipoproteins and cellular cholesterol homeostasis in man: studies of low-density lipoprotein receptor activity and 3-hydroxy-3-methylglutaryl coenzyme A reductase activity in blood mononuclear cells. *J Clin Invest.* 1981;67:493–502.
44. Beynen AC, Katan MB. Reproducibility of the variations between humans in the response of serum cholesterol to cessation of egg consumption. *Atherosclerosis.* 1985;57:19–31.
45. McNamara DJ, Kolb R, Parker TS, Batwin H, Samuel P, Brown CD, Ahrens EH Jr. Heterogeneity of cholesterol homeostasis in man: response to changes in dietary fat quality and cholesterol quantity. *J Clin Invest.* 1987;79: 1729–1739.
46. Zanni EE, Zannis VI, Blum CB, Herbert PN, Breslow JL. Effect of egg cholesterol and dietary fats on plasma lipids, lipoproteins, and apoproteins of normal women consuming natural diets. *J Lipid Res.* 1987;28:518–527.
47. Martin LJ, Connelly PW, Nancoo D, Wood N, Zhang ZJ, Maguire G, Quinet E, Tall AR, Marcel YL, McPherson R. Cholesteryl ester transfer protein and high-density lipoprotein responses to cholesterol feeding in men: relationship to apolipoprotein E genotype. *J Lipid Res.* 1993;34:437–446.
48. Katan MB, Beynen AC. Characteristics of human hypo- and hyperresponders to dietary cholesterol. *Am J Epidemiol.* 1987;125:387–399.
49. Nestel PJ, Poyser A. Changes in cholesterol synthesis and excretion when cholesterol intake is increased. *Metabolism.* 1976;25:1591–1599.
50. Miettinen TA, Kesaniemi YA. Cholesterol absorption: regulation of cholesterol synthesis and elimination and within-population variations of serum cholesterol levels. *Am J Clin Nutr.* 1989;49:629–635.
51. Denke MA, Grundy SM. Individual responses to a cholesterol-lowering diet in 50 men with moderate hypercholesterolemia. *Arch Intern Med.* 1994;154:317–325.
52. Dietschy JM, Turley SD, Spady DK. Role of liver in the maintenance of cholesterol and low-density lipoprotein homeostasis in different animal species, including humans. *J Lipid Res.* 1993;34:1637–1659.
53. Ellsworth DL, Hallman DM, Boerwinkle E. Impact of the human genome project on epidemiologic research. *Epidemiol Rev.* 1997;19:3–13.
54. Menzel H-J, Kladetzky R-G, Assmann G. Apolipoprotein E polymorphism and coronary artery disease. *Arteriosclerosis.* 1983;3:310–315.
55. Kuusi T, Nieminen MS, Ehnholm C, Yki-Järvinen H, Valle M, Nikkilä EA, Taskinen M-R. Apoprotein E polymorphism and coronary artery disease: increased prevalence of apolipoprotein E-4 in angiographically verified coronary patients. *Arteriosclerosis.* 1989;9: 237–241.
56. Hixson JE. Pathobiological Determinants of Atherosclerosis in Youth Research Group. Apolipoprotein E polymorphisms affect atherosclerosis in young males. *Arterioscler Thromb.* 1991;11:1237–1244.
57. Mahley RW. Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science.* 1988;240:622–630.
58. Sing CF, Davignon J. Role of the apolipoprotein E polymorphism in determining normal plasma lipid and lipoprotein variation. *Am J Hum Genet.* 1985;37:268–285.
59. Gregg RE, Zech LA, Schaefer EJ, Stark D, Wilson D, Brewer HB Jr. Abnormal *in vivo* metabolism of apolipoprotein E_4 in humans. *J Clin Invest.* 1986;78:815–821.
60. Sharrett AR, Patsch W, Sorlie PD, Heiss G, Bond MG, Davis CE. Associations of lipoprotein cholesterols, apolipoproteins A-I and B, and triglycerides with carotid atherosclerosis and coronary heart disease. *Arterioscler Thromb.* 1994;14:1098–1104.
61. Wilson PW, Myers RH, Larson MG, Ordovas JM, Wolf PA, Schaefer EJ. Apolipoprotein E alleles, dyslipidemia,

and coronary heart disease: the Framingham Offspring Study. *JAMA.* 1994;272:1666–1671.
62. Boerwinkle E, Utermann G. Simultaneous effects of the apolipoprotein E polymorphism on apolipoprotein E, apolipoprotein B, and cholesterol metabolism. *Am J Hum Genet.* 1988;42:104–112.
63. Davignon J, Gregg RE, Sing CF. Apolipoprotein E polymorphism and atherosclerosis. *Arteriosclerosis.* 1988;8:1–21.
64. Manttari M, Koskinen P, Ehnholm C, Huttunen JK, Manninen V. Apolipoprotein E polymorphism influences the serum cholesterol response to dietary intervention. *Metabolism.* 1991;40:217–221.
65. Gylling H, Miettinen TA. Cholesterol absorption and synthesis related to low-density lipoprotein metabolism during varying cholesterol intake in men with different apoE phenotypes. *J Lipid Res.* 1992;33:1361–1371.
66. Lopez-Miranda J, Ordovas JM, Mata P, Lichtenstein AH, Clevidence B, Judd JT, Schaefer EJ. Effect of apolipoprotein E phenotype on diet-induced lowering of plasma low-density lipoprotein cholesterol. *J Lipid Res.* 1994;35:1965–1975.
67. Kesaniemi YA, Ehnholm C, Miettinen TA. Intestinal cholesterol absorption efficiency in man is related to apoprotein E phenotype. *J Clin Invest.* 1987;80:578–581.
68. Lefevre M, Ginsberg HN, Kris-Etherton PM, Elmer PJ, Stewart PW, Ershow A, Pearson TA, Roheim PS, Ramakrishnan R, Derr J, Gordon DJ, Reed R. ApoE genotype does not predict lipid response to changes in dietary saturated fatty acids in a heterogeneous normolipidemic population: the DELTA Research Group: Dietary Effects on Lipoproteins and Thrombogenic Activity. *Aterioscler Thromb Vasc Biol.* 1997;17(11):2914–2923.
69. Friedlander Y, Berry EM, Eisenberg S, Stein Y, Leitersdorf E. Plasma lipids and lipoproteins response to a dietary challenge: analysis of four candidate genes. *Clin Genet.* 1995;47:1–12.
70. Zambon D, Ros E, Casals E, Sanllehy C, Bertomeu A, Campero I. Effect of apolipoprotein E polymorphism on the serum lipid response to a hypolipidemic diet rich in monounsaturated fatty acids in patients with hypercholesterolemia and combined hyperlipidemia. *Am J Clin Nutr.* 1995;61:141–148.
71. Lohse P, Kindt MR, Rader DJ, Brewer HB Jr. Genetic polymorphism of human plasma apolipoprotein A-IV is due to nucleotide substitutions in the apolipoprotein A-IV gene. *J Biol Chem.* 1990;265:10061–10064.
72. Menzel HJ, Sigurdsson G, Boerwinkle E, Schrangl-Will S, Dieplinger H, Utermann G. Frequency and effect of human apolipoprotein A-IV polymorphism on lipid and lipoprotein levels in an Icelandic population. *Hum Genet.* 1990;84:344–346.
73. McCombs RJ, Marcadis DE, Ellis J, Weinberg RB. Attenuated hypercholesterolemic response to a high-cholesterol diet in subjects heterozygous for the apolipoprotein A-IV-2 allele. *N Engl J Med.* 1994;331:706–710.
74. Law A, Wallis SC, Powell LM, Pease RJ, Brunt H, Priestley LM, Knott TJ, Scott J, Altman DG, Miller GJ, Rajput J, Miller NE. Common DNA polymorphism within coding sequence of apolipoprotein B gene associated with altered lipid levels. *Lancet.* 1986;1:1301–1303.
75. Rajput-Williams J, Knott TJ, Wallis SC, Sweetnam P, Yarnell J, Cox N, Bell GI, Miller NE, Scott J. Variation of apolipoprotein-B gene is associated with obesity, high blood cholesterol levels, and increased risk of coronary heart disease. *Lancet.* 1988;2:1442–1446.
76. Demant T, Houlston RS, Caslake MJ, Series JJ, Shepherd J, Packard CJ, Humphries SE. Catabolic rate of low density lipoprotein is influenced by variation in the apolipoprotein B gene. *J Clin Invest.* 1988;82:797–802.
77. Corsini A, Fantappie S, Granata A, Bernini F, Catapano AL, Fumagalli R, Romano L, Romano C. Binding-defective low-density lipoprotein in family with hypercholesterolemia. *Lancet.* 1989;1:623.
78. Fisher RM, Mailly F, Peacock RE, Hamsten A, Seed M, Yudkin JS, Beisiegel U, Feussner G, Miller G, Humphries SE, Talmud PJ. Interaction of the lipoprotein lipase asparagine 291 → serine mutation with body mass index determines elevated plasma triacylglycerol concentrations: a study in hyperlipidemic subjects, myocardial infarction survivors, and healthy adults. *J Lipid Res.* 1995;36:2104–2112.
79. McGue M, Rao DC, Iselius L, Russell JM. Resolution of genetic and cultural inheritance in twin families by path analysis: application to HDL-cholesterol. *Am J Hum Genet.* 1985;37:998–1014.
80. Austin MA, King M-C, Bawol RD, Hulley SB, Friedman GD. Risk factors for coronary heart disease in adult female twins. Genetic heritability and shared environmental influences. *Am J Epidemiol.* 1987;125:308–318.
81. De Oliveira e Silva ER, Arnberg R, Seidman CE. Environmental and genetic considerations in plasma lipoprotein levels in identical twins. *Clin Res.* 1992; 40:140A.
82. Wade J, Milner J, and Krondl M. Evidence for a physiological regulation of food selection and nutrient intake in twins. *Am J Clin Nutr.* 1981;34:143–147.
83. Kleinbaum DG, Kupper LL. *Applied Regression Analysis and Other Multivariable Methods.* North Situate, Mass: Duxbury Press; 1978.
84. De Oliveira e Silva ER, Brinton EA, Cundey KB, Breslow JL. Heritability of HDL turnover parameters. *Circulation.* 1991;84(supp. II):339.
85. Nichols RC, Bilbro WC Jr. The diagnosis of twin zygosity. *Acta Genet Stat Med.* 1966;16:265–275.
86. Carmelli D, Swan GE, Robinette D, Fabsitz RR. Heritability of substance use in the NAS-NRC Twin Registry. *Acta Genet Med Gemellol (Roma).* 1990;39:91–98.

87. Lykken DT, Bouchard TJ Jr, McGue M, Tellegen A. The Minnesota Twin Family Registry: some initial findings. *Acta Genet Med Gemellol (Roma).* 1990;39:35–70.
88. Cohen DJ, Dibble E, Grawe JM, Pollin W. Reliably separating identical from fraternal twins. *Arch Gen Psychiatry.* 1975;32:1371–1375.
89. Alford RL, Caskey CT. DNA analysis in forensics, disease and animal/plant identification. *Curr Opin Biotechnol.* 1994;5:29–33.
90. Clayton EW, Steinberg KK, Khoury MJ, Thomson E, Andrews L, Kahn MJE, Kopelman LM, Weiss JO. Informed consent for genetic research on stored tissue samples. *JAMA.* 1995;274:1786–1792.
91. The American Society of Human Genetics. ASHG report: statement on informed consent for genetic research. *Am J Hum Genet.* 1996;59:471–474.
92. Knoppers BM, Laberge CM. Research and stored tissues: persons as sources, samples as persons? *JAMA.* 1995;274:1806–1807.
93. Durfy SJ. Ethics and the Human Genome Project. *Arch Pathol Lab Med.* 1993;117:466–469.
94. McGuire RE, Daiger SP, Green ED. Localization and characterization of the human ADP-ribosylation factor 5 (ARF5) gene. *Genomics.* 1997;41:481–484.