J Mol Med (2001) 79:57–70
DOI 10.1007/s001090100202

REVIEW

A.G. Comuzzie · J.T. Williams · L.J. Martin
J. Blangero

Searching for genes underlying normal variation in human adiposity

Received: 18 September 2000 / Accepted: 15 January 2001 / Published online: 16 February 2001

Abstract A primary challenge in biomedical research today is the elucidation of the underlying genetic architecture of complex conditions such as obesity. In contrast to simple Mendelian disorders that result from a mutation in a single gene, complex phenotypes are the product of the action (as well as interaction) of multiple genes and environmental factors. The genetic configuration of these genes can range from effectively polygenic (i.e., many genes each with a relatively small contribution) to oligogenic (i.e., a few genes with relatively large measurable effects often expressed on a residual additive genetic background). While the task at hand is complicated, it is not intractable; however, it does require consideration of the nature of the disease and definition of its associated phenotypes in selecting the most appropriate study design. Here we will discuss the characteristics of obesity and its related phenotypes, which must be considered in designing analyses to identify the genes involved as well as reviewing what these approaches have provided in the search for genes influencing adiposity in humans

Keywords Oligogenic · Quantitative trait loci · Genome scanning · Positional candidate genes

Abbreviations *BMI:* Body mass index · *IBD:* Identity by descent · *IBS:* Identity by state · *QTL:* Quantitative trait locus

ANTHONY COMUZZIE received his Ph.D. degree in 1993 from the University of Kansas, USA. He currently holds the position of Associate Scientist in the Department of Genetics, Southwest Foundation for Biomedical Research in San Antonio, Texas. His research interests focus on the genetics of complex diseases with emphasis on obesity and type II diabetes.

JOHN BLANGERO received his Ph.D. degree in 1987 from Case Western Reserve University, USA. He is presently Scientist in the Department of Genetics at the Southwest Foundation for Biomedical Research in San Antonio, Texas. In recent years his primary research focus has involved the development of variance component methods for linkage analyses of complex traits and diseases.

A.G. Comuzzie (✉) · J.T. Williams · L.J. Martin · J. Blangero
Department of Genetics,
Southwest Foundation for Biomedical Research,
P.O. Box 760549, San Antonio, TX 78245-0549, USA
e-mail: agcom@darwin.sfbr.org

Introduction

While it is clear that environmental factors play a significant role in the development of obesity [18, 31, 40, 67, 72], work over the past two decades has clearly demonstrated a genetic influence on adipose tissue accumulation and distribution [11, 24], with estimations of an additive genetic contribution to the expression of some obesity-related traits of 50% or greater [16, 25, 42, 63]. With the existence of a genetic component established, more recent efforts in this area have focused on the identification of the specific genes involved in the regulation of body fat. Some of this effort has focused on the identification of rare single gene defects that typically lead to extreme obesity (e.g., [54, 73]. While these efforts have provided important insights into various aspects of the biochemical pathways involved in adipose tissue regulation, they have not identified the genes influencing the common variation observed in obesity-related phenotypes. Here we focus on the approaches employed to identify common genes for complex phenotypes such as

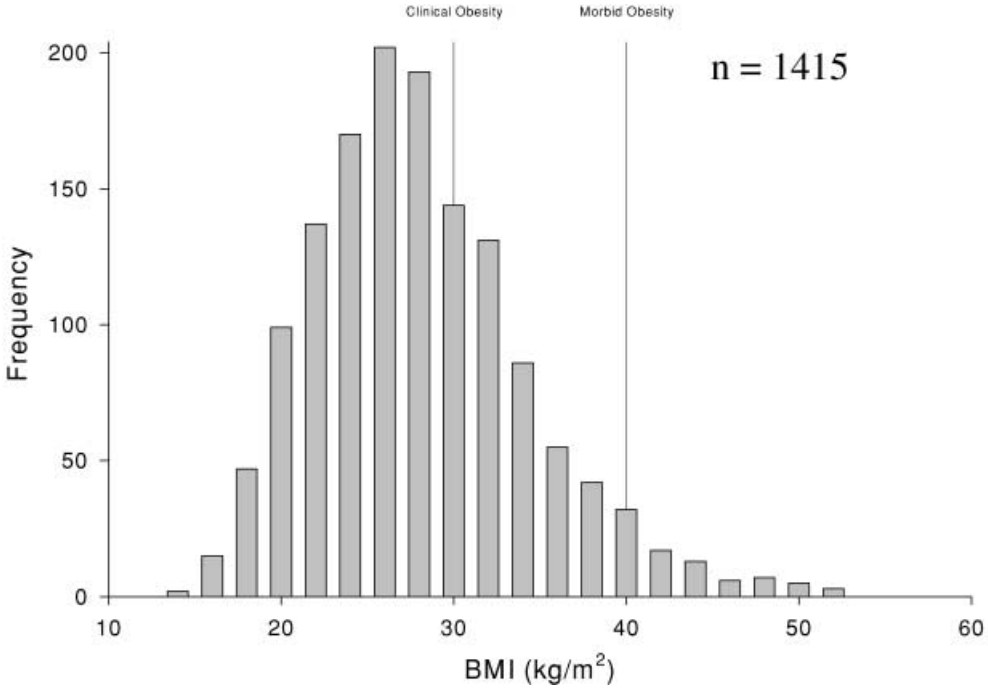

Fig. 1 Distribution of body mass index (*BMI*) in a randomly ascertained sample of Mexican-Americans

those related to obesity as well as reviewing the current results of these efforts.

In this search for common genes influencing the variation observed in human adiposity a variety of methodological approaches have been employed. Unfortunately, given the nature of obesity as a clinical condition and the characteristics of obesity-related phenotypes in general, many of these methodological approaches are inefficient or even inappropriate with regards to identifying genes with significant measurable effects for common diseases/conditions [5, 14]. As a result, there has been little consistency in findings between many of the studies to date. However, given recent developments in both molecular genetic techniques and statistical genetic methods in the past few years, there now exists some very powerful analytical tools for the identification of genes for common complex diseases such as obesity which have begun to yield replicable results across a number of studies. Specifically, these developments have now made genome-wide scanning feasible making it possible to move beyond the largely unsuccessful traditional candidate gene approach and allowing for the identification of previously unsuspected genes with measurable effects on the expression of obesity-related phenotypes.

Defining the phenotype

Obesity as a complex phenotype

A complex phenotype refers to a quantifiable characteristic of an organism that is influenced by both multiple genetic and nongenetic (i.e., environmental) factors as well as their interactions, which leads to a significant degree of variation in expression across the population (e.g., body mass index, BMI; see Fig. 1). While the term "complex" is generally associated in this context with phenotypes which are measurable on a continuous scale (i.e., body weight or waist circumference), it refers to the sources of interindividual variation for a specific trait, rather than whether a trait is continuous or discontinuous in its distribution within a population. Therefore, even when obesity is diagnosed clinically based on some "cut-point" (i.e., a BMI greater than 27) it can still be classified as a "complex" trait because variation in its age at onset or severity of other associated symptoms cannot be attributed to a single gene or environmental factor. As a result of this complexity it is generally impossible to easily discern a pattern of simple Mendelian inheritance across generations of relatives. However, with the proper sampling strategies and analytical tools (selected on the basis of the characteristics of the condition of interest), it is possible to not only elucidate the genetic architecture of such complex conditions as obesity, but to identify the specific genes responsible for the variation observed in its expression in a population [3, 11, 66].

Characterizing obesity: a continuous versus discrete phenotype

Obesity as a clinical condition is currently defined as an excess accumulation of adipose tissue resulting in a BMI greater than 30 [84]. While the diagnosis of obesity by this, and other, criteria may have clinical relevance with regards to intervention, management, and/or treatment of the condition or from an epidemiological perspective [45], it offers little utility for the study of the genetics underlying variation in body fat accumulation and distribution. For example, BMI values exhibit a normal distri-

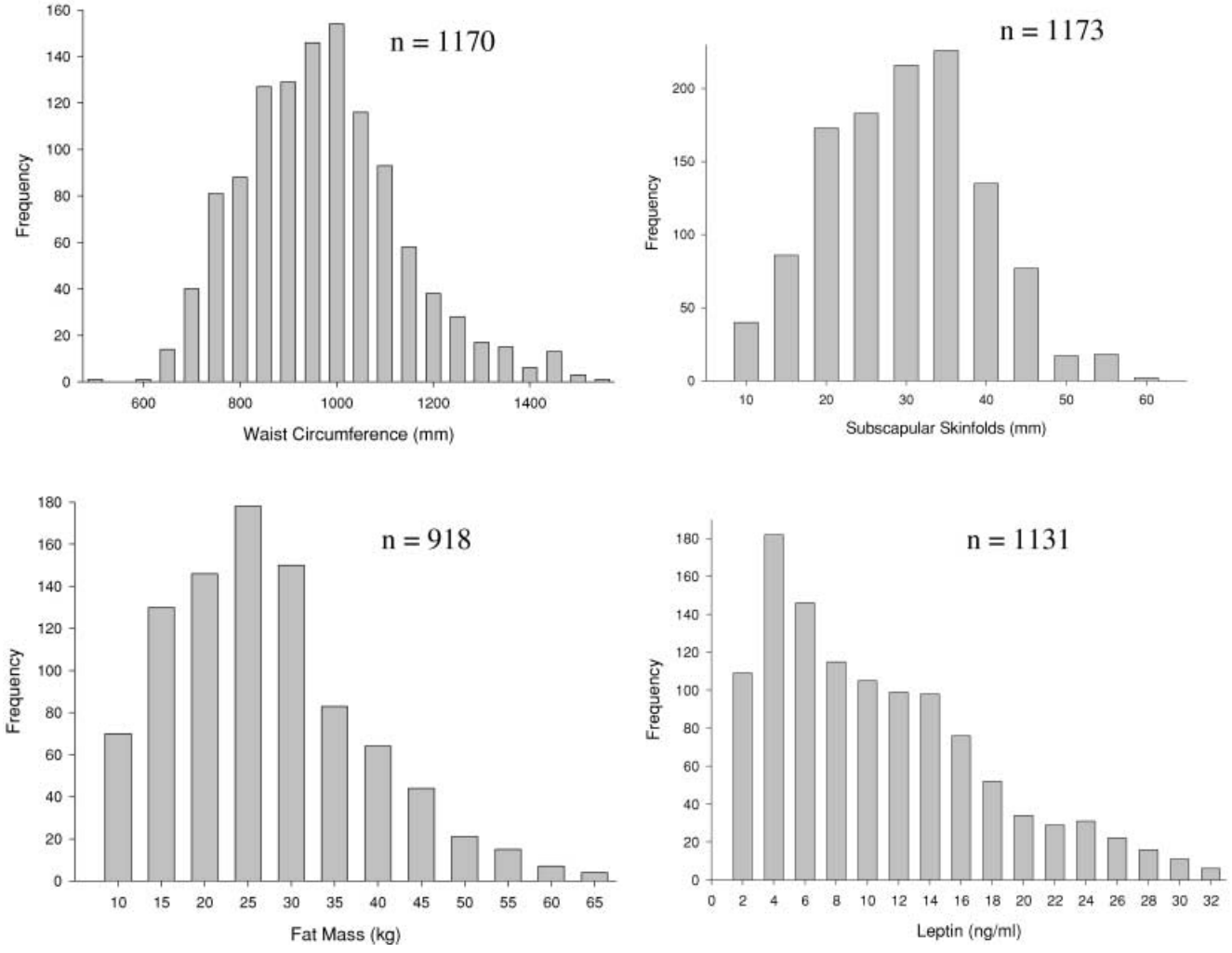

Fig. 2 Distribution of waist circumference, subscapular skinfolds, fat mass, and serum leptin levels in a randomly ascertained sample of Mexican-Americans

bution with no clear discontinuities between the "clinically obese" (BMI >30) and the nonobese. Even focusing on the morbidly obese (BMI >40), these individuals merely fall at one end of a continuum (Fig. 1). Such a pattern of continuous distribution is not restricted to merely BMI but is found in all obesity-related phenotypes, including anthropometric measures (e.g., skinfolds, and circumferences), measures of body composition (e.g., percentage body fat, fat mass), and associated endocrine levels (e.g., leptin) (Fig. 2). Given these observations, it becomes apparent that the "disease state" of obesity represents merely the extreme of one tail of the distribution for these continuously distributed phenotypes. Therefore the division of individuals into "obese" versus "nonobese" categories has a certain degree of arbitrariness that does not appear to follow any underlying biological phenomenon or natural discontinuities. As a result the basis for defining obesity as a clinical condition can vary and has varied based on changing popular opinion, rather than on marked discontinuities between lean and obese individuals. This is in sharp contrast to the use of continuous phenotypic measures whose variation is consistent with what is known concerning the mechanisms of gene action. Most importantly, however, the use of a clinical definition (which takes continuously distributed measures and converts them to such dichotomous states as "disease" vs. "no disease") actually hinder attempts to identify genes influencing complex phenotypes by greatly decreasing statistical power, since significant information is lost in the transition from a continuous to a discontinuous scale [29, 79, 83]. Therefore consideration of the distributional characteristics of obesity-related phenotypes becomes of paramount importance in selecting the appropriate method for genetic analyses of these traits [3, 14].

Selecting the analytical tools

Quantitative genetics

As we and others have demonstrated, obesity is a complex condition that is defined based on measures of clearly continuously distributed traits (e.g., BMI, fat mass, waist circumference, leptin levels). Given these

characteristics of obesity and obesity-related phenotypes, a quantitative genetic approach is required. At its most simplistic, quantitative genetic analyses are used to estimate the relative proportion of the total phenotypic variance in a complex trait that is attributable to the additive effects of genes (i.e., its heritability), and represents the first and most fundamental step in the identification of specific genes with measurable effects on the phenotype. Using basic quantitative genetic principles, it is possible to decompose the total phenotypic variance in a trait (σ^2_P) into its genetic (σ^2_G) and environmental (i.e., nongenetic; (σ^2_E) components, so that $\sigma^2_P = \sigma^2_G + \sigma^2_E$ [32, 49]. By extension, these components can be further decomposed, such that σ^2_G can be separated into components representing the variance attributable to additive genetic effects (σ^2_A) dominance (σ^2_D), and epistasis (i.e., gene–gene interaction; σ^2_I), while σ^2_E can be decomposed into components attributable to measured environmental factors (e.g., smoking and alcohol consumption) and random, unmeasured factors. The heritability (h^2) of a trait represents the proportion of the phenotypic variance due to genetic differences among individuals. The proportion of phenotypic variation attributable to *all* genetic effects (e.g., additivity, dominance, epistasis) is referred to as the "broad sense" heritability and is expressed as $h^2 = \sigma^2_G/\sigma^2_P$ while the heritability in the "narrow sense" refers to the proportion attributable to the *additive* genetic variance alone, i.e., $h^2 = \sigma^2_A/\sigma^2_P$ [32, 48]. Unless otherwise specified, we use the term "heritability" in this latter, narrow sense here. Given the additive nature of the components of the phenotypic variance, the heritability of a trait is influenced by the magnitude of the underlying genetic variance and the amount of environmentally induced variation.

From these basic quantitative genetic principles it is possible to make theoretical extensions to allow for the identification of specific genes that influence the observed phenotypic variation [2, 6, 8, 34, 68]. In theory, *any* locus that influences the variability in a quantitative phenotype (e.g., leptin levels or BMI) may appropriately be classified as a quantitative trait locus (QTL) [66]. In practice, however, there are a number of factors that impact on the ability to detect a QTL, and these factors also impact the choice of methods utilized in this search for genes [3, 14, 66].

In effect the ability to detect a QTL is a complex function of the effect size of the QTL itself (i.e., the proportion of the total phenotypic variance attributable to the QTL), the study design (i.e., data obtained from sib-pairs, sibships, nuclear families, or extended pedigrees), sample size, and the characteristics of the genetic data (i.e., number and heterozygosity of molecular genetic markers). Given these conditions, there is generally a limit to how small a QTL effect that can reliably be detected. In general, the detection and localization of QTLs are most successful for traits whose phenotypic variance is influenced primarily by a single locus (i.e., monogenic) or a few loci each with substantial effects (i.e., oligogenic) as opposed to traits whose phenotypic variation is attributable to dozens of genes with equal and individually small effects (i.e., polygenic). With the exception of the few single gene defects referred to above, obesity and its related phenotypes seem to best fit the oligogenic category of traits that is, a few genes each with measurable effects expressed on a polygenic background [21, 26, 37, 48, 55, 56, 76]. By means of the analytical methods discussed below, it is now reasonable to expect that, with appropriate sample sizes and study designs, individual genes accounting for as little as 10–15% of variation in a trait can be localized to specific chromosomal regions [80, 81, 82].

The failure of the traditional candidate gene approach

Until relatively recently efforts to identify specific genes influencing complex traits such as obesity have relied on the candidate gene concept. Candidate genes are those with generally well-known chromosomal locations and which are known to be part of a relevant biochemical pathway to the phenotype under study. This approach is appealing since the selection of genes for use in linkage and/or association studies is based upon a priori knowledge of the phenotype and the potential function of the gene involved. Based on this type of approach there are currently more than 200 genes that have been proposed as potential candidates for human obesity [22]. However, to paraphrase Thomas Huxley, the vast majority of these candidates have been "beautiful hypotheses killed by the ugly fact" that they have not been shown to have a significant contribution to the observed variation in obesity-related phenotypes. A perfect example of this can be found in the leptin gene (*Ob*) as a candidate gene for the regulation of leptin expression. While several studies have now identified genes with measurable effects of the expression of leptin [26, 37, 76], the leptin gene has never been demonstrated to have a significant effect on the normal variation in leptin expression or related measures of adiposity. In contrast, several new obesity-related candidate genes that have now been identified based on their proximity to a linkage signal would not necessarily have been suspected prior to a full genome scan (this approach is discussed in more detail below).

There are a number of potential explanations as to why the candidate gene approach has failed to yield significant results in the effort to identify genes with measurable effects on the common variation seen in adiposity among humans. The first explanation begins with the fact that the majority of these candidate genes have been generated from rodent models of obesity. While the contribution of these animal models to the study of obesity (and other complex diseases) has been invaluable in helping to delineate the biochemical pathways that influence the expression of these complex phenotypes, in general they have not identified genes with significant effects at the population level in humans. Given the evolutionary distance that separates humans from many of these animal models, it would not be surprising that not

all genes which influence adiposity in rodents are similarly influential in humans and vice versa. Another issue that could account for the lack of significant candidate gene results to date is the fact that most of these studies were carried out utilizing inefficiently collected or relatively small samples, resulting in a significant loss of statistical power to detect these genetic effects. Finally, there is also the possibility that the failure to obtain truly significant results utilizing the traditional candidate gene approach simply derives from our general inability to specify reasonable human candidate genes a priori for complex phenotypes. As a result of this general lack of positive results from candidate gene studies in human obesity, numerous researchers have begun adopting alternative approaches to gene discovery such as genome-wide scans.

The rise of genome scanning

Given the equivocal nature, at best, of the results produced from the traditional candidate gene approach (for both linkage and association studies), along with recent revolutionary advances in both molecular and statistical genetics allowing for the use of large numbers of anonymous polymorphisms, we are now seeing a move toward full genome scanning to identify genes influencing a variety of complex phenotypes including those related to obesity.

The primary point that distinguishes the genome scan from the "traditional" candidate gene approach is the fact that no a priori assumptions concerning the potential importance of genes or chromosomal regions are made before starting the scan, but rather the variation across the entire genome is examined. Therefore it is the results of the genome scan which establish candidate chromosomal regions, or in some cases positional candidate genes, which then becomes the focus of more intensive follow-up analyses (e.g., combined linkage/disequilibrium analysis [5]). A positional candidate gene differs from a "traditional" candidate gene in that it is only considered as a candidate after the establishment of its proximity to a QTL that was identified via linkage analysis in a genome screen and not merely a priori assumptions concerning its possible physiological role. This is not to say that some current "traditional" candidate genes may not also prove to be positional candidate genes as well. Thus, the genome scan approach offers the potential of identifying new and/or previously unsuspected genes influencing the phenotype of interest.

In a genome scan, linkage analysis is conducted using a series of anonymous polymorphisms, scattered across the entire genome to identify QTLs affecting the phenotype of interest. Microsatellites, or simple sequence repeat loci (e.g., single tandem repeats) have become the dominant type of genetic marker for linkage analyses [19, 77]. These units are highly susceptible to mutations that increase or decrease the number of repeats, and therefore a population accumulates a large number of alleles at these loci that differ in the number of repeat units present. The human genome map now includes more than 8,000 of these polymorphisms.

In order to maximize their information content these anonymous markers should be highly heterozygous and spaced evenly across the entire genome. The power to detect a QTL in a genome scan depends in part on the heterozygosity of the genetic markers used. Heterozygosity refers to the probability that a random individual is heterozygous for any two alleles at a locus. Markers of low heterozygosity are less useful in determining which alleles are identical by descent among a set of individuals, and therefore statistical power is reduced (see below). The average spacing among polymorphic markers is also a factor in determining statistical power, since functional genes that occur in large gaps within a map are less likely to be detected than genes more closely linked to a genotyped marker. The generally accepted density of information for initial linkage analyses is a linkage map of markers with average heterozygosity about 0.70 or higher, and average spacing less than or equal to 10 cM.

While there is general enthusiasm for the genome scan approach for identifying genes influencing complex traits in humans, there is some disagreement as to what constitutes the most efficient sampling strategies and analytical methods for the identification of genes contributing to the expression of such common conditions as obesity.

QTL mapping methods for complex phenotypes

Classical penetrance-based linkage analysis has been successfully applied to the mapping of diseases with monogenic influences (i.e., Huntington disease and achondroplasia) in human pedigrees. While classical penetrance model based methods can be extremely powerful for mapping genes, they require detailed knowledge of the underlying genetic model. Specifically, it is necessary to stipulate the prevalence of the trait, its mode of inheritance, the allele frequencies at the presumed disease gene, and the probability of being affected, or penetrance, for each genotype, and in addition, for quantitative traits, the allele frequencies, and the mean trait value for each genotype must also be specified. In the case of complex phenotypes such as those related to obesity and other common diseases, which are influenced by multiple genes with a variety of interactions, such precise specification for all relevant parameters is extremely difficult. Failure to properly specify these parameters leads to a significant loss of power to detect linkage [35, 36, 50] but can also increase the probability of falsely excluding a chromosomal region containing a QTL [65].

Given these issues concerning proper model specification, there has been a recent effort to develop more efficient linkage analysis methods specifically designed to address complex traits using a penetrance model free approach. Here we present a brief overview of some of the

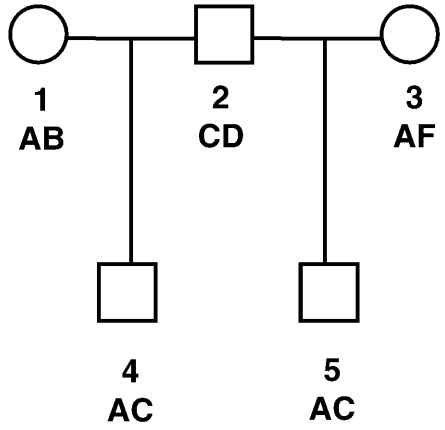

Fig. 3 An example of identity by descent (*IBD*) and identity by state (*IBS*)

most widely used methods and study designs for linkage analysis of complex phenotypes with particular emphasis on their applicability to the study of obesity. While these types of methods do not estimate parameters regarding the mode of inheritance, references to them as "nonparametric" and "model-free" linkage methods is somewhat misleading since they do involve the estimation of other parameters.

The underlying principle on which penetrance model-free methods for linkage analysis is based is the concept of identity by descent (IBD) allele sharing among relatives. For two alleles to be considered IBD they must have been inherited from the same ancestor. In addition, it is also possible to describe allele sharing in terms of identity by state (IBS). In this case two alleles are lBS if they are of the same form, regardless of their origin. These two concepts can be illustrated in the following example. In Fig. 3, individuals 4 and 5, who are half-siblings, share the allele C IBD because both received it from a common ancestor, their father. Alternatively, while they both have an A allele, these alleles are not IBD but rather IBS since they were inherited from the unrelated mothers of the two half-siblings and cannot be traced back to a common ancestor. In addition, while individuals 1 and 3 cannot share any alleles IBD, as they are unrelated, they share the A allele IBS. IBD allele sharing provides information about linkage, while IBS allele sharing provides information about population level association between the marker and the trait of interest.

In the absence of inbreeding (i.e. that is matings between close relatives), a pair of individuals can share 0, 1, or 2 alleles IBD, and the degree of IBD sharing is usually expressed as the proportion of alleles shared (π) and can take the values 0, 1/2, and 1. The expected IBD sharing over the entire genome [E(π)] for a relative pair is equal to twice the pair's kinship coefficient (2Φ). For siblings, E(π) is 1/2, whereas it is 1/4 for half siblings and 1/8 for first cousins. If a pedigree is inbred, higher degrees of IBD sharing are possible when individuals are homozygous for an allele from a particular ancestor. Calculation of IBD sharing for a genotyped marker locus is straightforward for nuclear families but becomes more complicated for extended pedigrees in which the top few generations are often unavailable for genotyping. However, a number of methods have been proposed to calculate IBD probability matrices when some marker data are unavailable [7, 27, 78].

Analytical approaches

Currently there are two basic analytical approaches to penetrance model free linkage analysis that are routinely used in genome scanning efforts: the Haseman and Elston [39] method and variance components methods [2, 3, 6, 14, 33, 34, 61, 68]. Until recently the most commonly used of these two basic approaches to quantitative trait linkage analysis was the Haseman-Elston method. This method is based on regressing the squared difference in trait values for pairs of siblings on the proportion of alleles shared IBD. The conceptual basis of this approach is rather straightforward; whenever a marker locus is closely linked to a gene that influences a phenotype, a significant negative regression coefficient should exist between number of alleles shared IBD and the squared phenotypic difference between sibs. This comes from the fact that siblings who are more similar phenotypically should also share more alleles IBD at the marker locus than sibs who are dissimilar for the phenotype under analysis. Recently the Haseman-Elston method has been revised [30] to use the product of the sibs' trait values rather than the difference. This new method addresses the phenotypic covariance between sibs and is functionally similar to variance component methods (described below), except in that it uses regression rather than maximum likelihood. Least squares regression is computationally more efficient than maximum likelihood estimation, and the new Haseman-Elston method can be more rapid than variance component methods.

While the attraction of the Haseman-Elston sibpair method is its relative simplicity of design and computation, it has been shown that for linkage analysis of quantitative traits the variance component method is more powerful [4, 62, 80]. Specifically, the variance component method explicitly deals with the nonindependence of the elements of the phenotypic covariance matrix, and since it is likelihood based, it yields a more powerful test than the Haseman-Elston method. In addition it can be expanded to deal with pedigrees of arbitrary size and complexity [2, 3, 14].

The basic idea behind the variance component linkage method is to attribute the population variance (i.e., the spread of the phenotypic values around the population mean) to a variety of genetic and nongenetic causes. The variance component method seeks to explain the correlations in phenotype among members of a family by partitioning the phenotypic variance into components due to the effect of a specific QTL linked to a genotyped marker, to other QTLs unlinked to the region under consider-

ation, to environmental factors shared among family members (e.g., diet), and to individual-specific environmental sources such as measurement error [6, 43].

In its simplest form for a trait influenced by *n* QTLs, the variance components are estimated by modeling the covariance among pedigree members as:

$$\Omega = \sum_{i=1}^{n} \hat{\Pi}_i \sigma_{qi}^2 + 2\Phi\sigma_a^2 + I\sigma_e^2$$

where $\hat{\Pi}_i$ is a matrix of estimated IBD sharing among family members at marker *i*, σ_{qi}^2 is the additive genetic variance due to a QTL linked to marker *i*, Φ is a matrix of kinship values, σ_a^2 is the residual additive genetic variance, σ_e^2 is the individual-specific environmental variance, and I is an identity matrix. One advantage of the variance component approach is that not only does it provide an estimate of the location for a QTL but also an estimate of the relative effect size for each modeled QTL as well. Following these same analytical principles, the variance component method can be extended to incorporate the consideration of oligogenic (i.e., multiple) QTLs simultaneously [2, 13]. With regards to the formal test for linkage, the model in which one or more QTL components are estimated is compared to that of a model in which these components are constrained to zero, testing the hypothesis that σ_{qi}^2 is significantly greater than zero. For a single locus analysis, the difference in the $\log_{10}$ likelihood between the two models is equivalent to the LOD score obtained from classical parametric linkage analysis.

There are several factors that have a direct effect on the power of the variance component method to detect a QTL. Most important among these factors are the magnitude of the QTL-specific heritability, the size and complexity of the relationships represented in the sample, as well as in the case of complex disease states (e.g., obese or diabetic) the population prevalence of the condition or disease of interest.

As previously noted, heritability in the broad sense is the proportion of the phenotypic variance that is attributable to all genetic effects while the QTL-specific heritability is the proportion of the phenotypic variance attributable to the additive genetic effects pf a specific QTL. Higher QTL-specific heritability yields greater power to detect linkage [81]. There are several ways to increase the QTL-specific heritability. For example, using the variance component method it is possible simultaneously to consider the effects of measured covariates, such as age, sex, and environmental exposures, on the phenotype. The inclusion of these covariate effects reduces the unexplained variance in the phenotype and increases the relative proportion of variance due to any QTLs, thereby increasing the power to detect linkage.

Another way to increase the power to localize QTLs is to allow for some of the complications of inheritance that are likely to play a role in determining quantitative variation in phenotypes. For example, genotype-by-environment interaction is likely to play an important role in the determination of some physiological phenotypes [12, 70]. The relationship between phenotype and genotype may be context dependent, such that genes act differently in the two sexes or change as a function of age [75, 86]. If such complications exist in the genotype-phenotype relationship, localization of QTLs may be more difficult. However, the variance component approach to quantitative trait linkage analysis can be easily extended to allow for such genotype-environment interactions [14, 74]. Other complications which may reduce the power to detect QTLs includes gene-gene interaction and epistasis. Such effects are very similar to those seen in genotype-environment interaction, and they too can be incorporated into the variance component framework [14, 23, 51, 71]. However, the practical impact of allowing for such complicated models of gene action remains to be seen and depends upon the frequency of such effects and their relative importance which is currently unknown.

Another approach to improve the localization of QTLs is to perform multivariate analyses. By exploiting the genetic and environmental correlations between phenotypes, such analyses can improve the power to detect linkage and provide more precise localizations of QTLs. Multivariate linkage analyses are still in their infancy, but they represent straightforward extensions of the standard variance component approach [4, 38, 83].

Another important factor influencing the power to detect a linkage signal is the structure of the familial relationships represented in the dataset as well as total sample size. In general, the larger and more complex the sampling unit, the more powerful is the variance component analysis [82], such that on a per person basis nuclear families provide more linkage power than sibpairs, and extended pedigrees provide more power than nuclear families. The availability of analytical results for the asymptotic power of affected sibpair and variance component linkage analysis allows alternative sampling designs to be readily compared with regard to their efficiency for localizing genes that influence complex common diseases. Figure 4 compares the power to detect linkage with these approaches by comparing the number of individuals required in each analysis to achieve 80% power to detect a QTL with a LOD score of 3 or better. Five study designs are illustrated in each graph. The two dashed lines show the numbers of individuals required in an affected sibpair study when the disease-associated allele is rare (P=0.1) or common (P=0.5). The three solid lines represent the required sizes of *randomly* selected samples when the variance component method is used with sibships of size 2 and 4 (labeled ‘2′ and ‘4′, respectively) and in extended pedigrees (labeled ‘P’). For the extended pedigrees, we used a typical family structure from the San Antonio Family Heart Study containing 48 individuals. The prevalence of the focal disease is 1% in Fig. 4a, and 35% in Fig. 4b. The prevalence of 1% is typical of a rare complex disease, such as schizophrenia, whereas a prevalence of 35% is seen for numerous chronic diseases in the United States, including obesity and hypertension. It is clear from Fig. 4 that each approach to linkage analysis is optimal for different disease

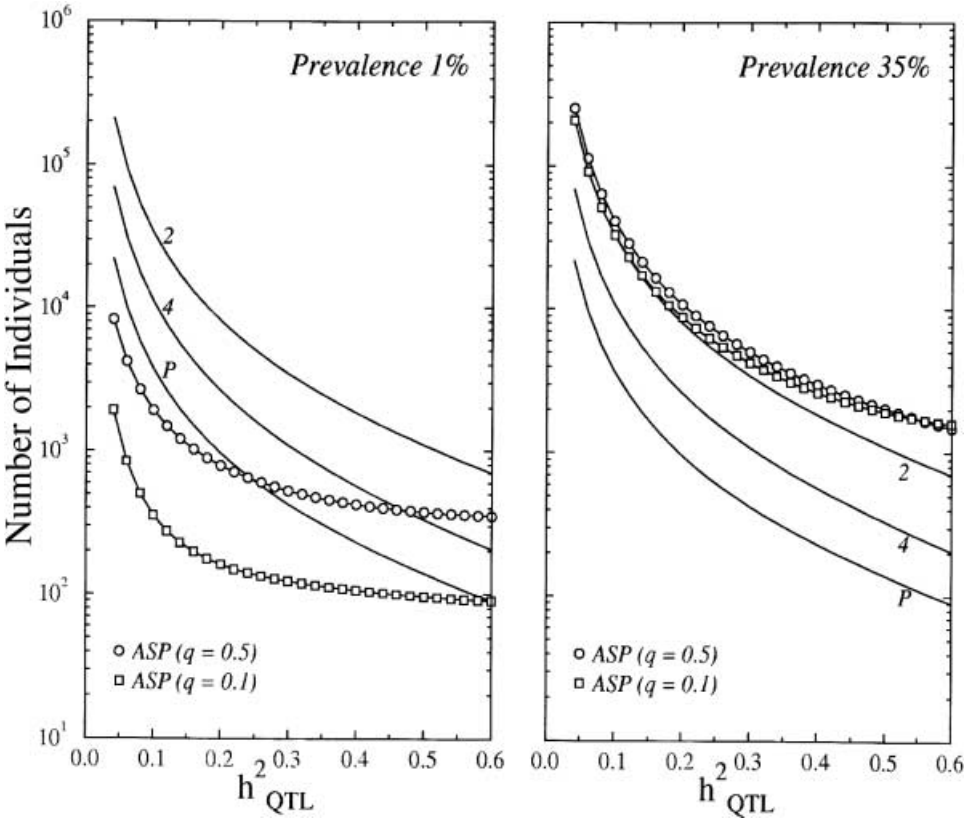

Fig. 4 Sample size requirements for 80% power to detect linkage under different sampling strategies and at different disease prevalences

prevalences, with the quantitative trait variance component approach outperforming the affected sibpair approach for common diseases. These power curves also reveal the dramatic advantage of pedigree-based designs over sibship-based designs and highlight the potential power of extended pedigrees for mapping QTLs at any disease prevalence. For example, QTLs accounting for as little as 5% of the variance can be mapped with only 10000 persons if large pedigrees are employed. This may seem to be a prohibitively large sample size, but many epidemiological studies are larger, as are some collaborative genetic studies.

Lastly, the variance component approach has its maximum power in the detection of genetic effects on normal variation in *common*, continuously varying, complex traits such as that expressed in obesity-related phenotypes [83]. There are both strengths and weaknesses to the various penetrance model free linkage methods available, and the choice of the most appropriate method depends greatly on the characteristics of the phenotype being analyzed (e.g., continuous vs. discrete distribution; common vs. rare expression) as well as the design of the data collection (e.g., simple sibpairs vs. extended pedigrees). With particular regard to obesity-related phenotypes, recent work has shown that quantitative traits measured on a continuous scale are more informative for linkage analysis than dichotomizations of the same trait [10, 29, 46, 79]. That is to say, if the trait of interest is continuously distributed (as we have already shown is the case with obesity-related phenotypes), or trait has quantitative correlates, analyzing the said trait measured on a continuous scale has more power than an analysis based on categorization of individuals into high and low or affected and unaffected classes. However, if the phenotype of interest must be measured on a discontinuous scale, the choice of sampling design and analytical method is driven by the prevalence of the condition. In such a case affected relative methods are often used, and the statistical power of affected relative pair methods is maximized when population prevalence of the disease/phenotype is low (i.e., typically less than 10% in the population). In contrast, variance component methods are most powerful when disease prevalence approaches 50%. Additionally, the collection of data from extended pedigrees within populations that are randomly ascertained with respect to phenotype offers demonstrably greater power to detect QTLs for common continuously varying complex traits.

General discussion of power

A critical concern with any method of linkage analysis is its statistical power, or the probability that the test will correctly reject a false null hypothesis of no linkage. Power studies can be difficult, however, and often make use of simulated datasets [29, 80]. Recently, however, some general findings have been presented that can be used to evaluate different strategies for linkage analysis [65, 81]. An essential lesson emerging from these studies and comparisons is that different diseases require differ-

ent study designs, which in turn require different analytical methods.

Disease prevalence offers the best basis for choosing the optimal sampling design and analytical method for a given trait. With rare complex diseases, linkage studies using only affected individuals recover most of the information regarding the genetics of the trait [64]. Unaffected individuals provide little additional information; hence it is most cost effective to target the sampling effort to the collection of the affected individuals.

Risch [64] has shown that the power to localize a disease-influencing QTL in an affected-sibpair linkage design is a monotonic function of the relative risk to siblings, λ_S. The relative risk is usually defined in terms of prevalences as $\lambda_S = K_S/K$, where K is the population prevalence and K_S is the prevalence in sibs of an affected proband. The relative risk can also be expressed in terms of the trait heritability as:

$$\lambda_S = 1 + \frac{(1-K)}{K}\left(\frac{1}{2}h_b^2 + \frac{1}{4}d_b^2\right)$$

where h_b^2 is the relative proportion of variance due to dominance effects on the binary scale. We emphasize that the additive genetic and dominance effects are on a binary scale because this scale is fundamentally inappropriate for diseases that truly have an underlying continuous liability; the discrepancy between the two scales generally leads to underestimates of the true heritability [28].

The dominant quantity in the expression for λ_S is the disease prevalence; the factor $(1–K)/K$ can take any positive value depending upon the magnitude of K, while the factor $\left(\frac{1}{2}h_b^2 + \frac{1}{4}d_b^2\right)$ has a theoretical maximum of 0.5. Therefore λ_S can become very large as the disease prevalence decreases, even when the contribution of the QTL is very small. This has the peculiar effect of rendering affected sibpair methods either more or less powerful to detect the same gene in populations that differ only in their disease prevalence.

With common complex diseases the situation is entirely different. For a common disease (e.g., obesity as based on a clinical threshold of a BMI >30), having a prevalence greater than about 10%, inclusion of unaffected individuals can markedly improve power, and studies limited only to affected individuals rapidly lose power. In contrast to the affected-pair method, the power to detect linkage using the variance component method for quantitative traits is primarily a function of the disease heritability on the quantitative scale, and is only slightly influenced by disease prevalence.

Williams and Blangero [81] investigated in general terms the power of variance component linkage analysis of quantitative traits and derived exact expressions for the sample size required to achieve a given power with various sampling structures. For example, in a linkage analysis model consisting only of a major gene effect, a residual additive genetic effect, and an individual-specific random environmental effect, the contribution by a single sibpair to the expected LOD score is:

$$ELOD = \frac{1}{2\ln 10}\frac{\left(h_q^2\right)^2\left[\left(h_T^2\right)^2 + 4\right]}{2\left[\left(h_T^2\right)^2 - 4\right]^2}$$

where h_T^2 is the total trait heritability and h_q^2 is the heritability due to the QTL. For a test of linkage having 80% power at a LOD score of 3.0, the number of sibpairs required is n=20.78/[(2ln10) ELOD] and the total number of individuals required is $2n$.

A result for arbitrary relative pairs is also of interest and can serve as a basis for estimating the power of any given pedigree. At a given locus let k_1 denote the probability that two individuals, i and j, share one allele IBD, and let k_2 denote the probability that i and j share both alleles IBD. The contribution per relative pair to the expected LOD score is then:

$$ELOD = \frac{1}{2\ln 10}\frac{\left(h_q^2\right)^2\left[\left(h_T^2\right)^2 k_1^2 + 1\right]\left[k_1 + 4k_2(1-2k_1)\right]}{4\left[\left(h_T^2\right)^2 k_1^2 - 1\right]^2}$$

This result can be used to estimate the power to detect linkage with a pedigree of any structure by summing the expected LOD score for each relative class over the distribution of relative classes within the pedigree. This result is based, however, on the assumption that the relative pair in question exhibits nonzero variance in the number of alleles they share IBD; consequently, parent-offspring pairs and monozygotic twin pairs must be excluded. This is of no particular disadvantage, however, for investigating the power to detect linkage, since neither of these relationships exhibits any variance in IBD sharing at a QTL and cannot be informative for linkage.

Current results of genome scans for obesity

While the use of genome scans has increased over the past several years, the number of published scans with obesity-related phenotypes as the primary focus is still rather small. However, there is a large amount of literature looking at linkage and association with traditional candidate gene loci and which is covered in a number of other reviews [17, 20, 22, 58, 58]. For the reasons outlined in the earlier part of this review, however, we focus only on those studies based on full genome scans. To date the results of full genome scans for obesity-related traits have been published on five different populations and include Mexican-Americans [26, 53], Native Americans [38, 56, 56, 76], French [37], a United States sample of individuals of predominantly white ancestry [48], and French Canadians [21], with several other studies in progress or set to begin soon. Despite the relatively small number of scans thus far published there have already been several significant findings, some which have now been replicated across several populations.

Mexican-Americans (the San Antonio Family Heart Study)

At present the strongest evidence for a quantitative trait locus influencing obesity-related phenotypes in humans comes from the San Antonio Family Heart Study [26, 41]. This study utilized a sample of 459 Mexican-Americans distributed in ten families (this sample represents 5667 relative pairs ranging from parent-offspring to double second cousins) using a variance component linkage approach [2] to genomic screening. Results were initially reported based on a 20-cM map of microsatellite markers, and identified significant linkage on chromosomes 2 (LOD=4.95) for leptin levels and fat mass (LOD 2.75) [26], and following the typing of additional markers to create a 10-cM map, the multipoint LOD score for leptin on chromosome 2 increased to 7.46 (Fig. 2), the largest LOD score yet published for a QTL in humans [41]. In addition, significant linkage has also been detected on chromosome 8 with both leptin (LOD=2.2) [26] and BMI (LOD=3.2) [53].

In the case of both the chromosome 2 and chromosome 8 linkage results, strong positional candidate genes for obesity have been found within the chromosomal regions identified. The 95% confidence interval surrounding the chromosome 2 QTL contains the *POMC* locus, which codes for the prohormone pro-opiomelanocortin. *POMC* was identified as a candidate based on its location [26] and is further supported by recent studies detailing its potential physiological involvement in obesity [47, 85]. Recently completed work has now identified polymorphisms in *POMC* that can be used in formal association analyses. With a haplotype generated using two common polymorphisms in *POMC* (one located in exon 3 and the other in the 5′ untranslated region) we have detected significant association (P=0.001) between molecular variation in the *POMC* locus and variation in serum levels of leptin among Mexican-Americans [41]. Since neither of these two polymorphisms appears to be functional, work is currently focusing on identifying additional polymorphisms in the promoter region *of POMC*.

The 95% confidence interval surrounding the chromosome 8 linkage contains the structural gene for the β-3-adrenergic receptor (*ADRB3*), a strong candidate gene previously identified based on its known physiological activity [69]. Given the linkage results from the San Antonio Family Heart Study, *ADRB3* is supported as a potential contributor to observed variation on the basis of chromosomal position as well. While previous association studies testing the relationship between the Trp64Arg polymorphism of *ADRB3* and obesity-related phenotypes yielded equivocal results [1], the argument for *ADRB3* as a human obesity gene has been strengthened by follow-up analyses in this same sample of Mexican-Americans discussed above with regard to *POMC* [52]. By first accounting for the effects of the QTL identified on chromosome 2, Mitchell and colleagues [52] were then able to detect association between variation in *ADRB3* and several obesity-related phenotypes (i.e., BMI, fat mass, waist circumference).

The Pima

The Pima Indian community of Arizona has been the subject of extensive efforts to identify genes influencing diabetes and its related conditions such as obesity [44]. To date the results of the genome screening of obesity-related phenotypes has detected potential evidence of several QTLs [38, 55, 56, 76]. This work has been restricted to sibships and nuclear families ranging upwards to approximately 1000 individuals and utilizing an average map density less than 10 cM. These analyses have utilized both the Haseman-Elston sibpair approach as well as variance components. The obesity-related phenotypes for which genome scan results have been published include BMI, percentage body fat, the ratio of waist to thigh circumference, 24-h metabolic rate, sleeping metabolic rate, 24-h respiratory quotient, and leptin levels. Currently QTLs have been detected on chromosomes 1, 3, 6, 11, 18, and 20 with LOD scores ranging from 2 to 3.6 [38, 55, 56, 76]. The strongest evidence for a QTL in the Pima is on chromosome 11 (approximately 11q21-q24) with a LOD score of 3.6 for BMI and was obtained using a variance component approach [38]. Additional obesity-related phenotypes including percentage body fat (LOD=2.8) [55], 24-h energy expenditure (LOD=2.0) [56], and diabetes status (LOD=1.5) [38] have also yielded linkage signals in this same region of chromosome 11 in the Pima. A bivariate analysis of BMI and diabetes status produced a LOD score of approximately 5.0 in this same region [38]. The next best supported evidence for a QTL in the Pima is located on chromosome 20 (approximately 20q11.2) for 24-h respiratory quotient with a LOD score of 3.0. Using a variance-components approach, Norman et al. [56] detected a LOD score of 2.3 on chromosome 18 (18q21) for percentage body fat and a LOD score of 2.8 for 24RQ on chromosome 1 (1p31-p21). Walder and colleagues [76], also using a variance component approach, have recently reported a QTL on chromosome 6 (6p) with a LOD score of 2.1, as well as several other smaller signals (1<LODs<2) on chromosomes 3, 11, 13, 15, and 16 in the Pima. However, given the marginal LOD scores for several of these signals along with the total number thus far reported, it is highly likely that some of these signals represent false positives.

French

Hager and colleagues [37] have published the results of a genome scan for obesity as a discrete trait (defined as a BMI >27) in affected sibpairs as well as quantitative variation in leptin levels also conducted in sibpairs in a sample collected from French families. This study utilized 514 individuals distributed over 158 nuclear families with each family having a proband with a BMI higher than 40 and at least one sibling with a BMI higher than 27. Using an affected sibpair approach they detected significant evidence of linkage on chromosome 10 (10p) with a LOD score of 4.9 [37]. In addition, using a

quantitative trait analysis they reported suggestive evidence of linkage for variation in serum leptin levels on chromosomes 2 (2p) and 5 (5cen-q) with LOD scores of 2.7 and 2.9, respectively [37]. While there are no readily apparent positional candidate genes in the regions of linkage on chromosomes 5 and 10 which might immediately appear to be involved in the regulation of adipose tissue, the signal they report on chromosome 2 is in the region of *POMC* as was previously reported in Mexican-Americans [26, 41].

United States whites

Using a data collection scheme focused on the recruitment of extremely obese individuals, Lee and colleagues [48] conducted a genome scan using both quantitative and qualitative obesity-related phenotypes in families of primarily white ancestry collected from across the United States. This study utilized a gene map with an average density of 10 cM and included 513 individuals from 92 nuclear families ascertained through an extremely obese proband (BMI ≥40) with obese siblings (BMI >30) along with siblings and parents of normal body weight (BMI <27) [48, 59]. The phenotypes analyzed for this genome scan included quantitative measures of BMI and percentage body fat as well as discrete traits of BMI ≥30 and percentage body fat ≥40%. While the results of this genome scan are somewhat difficult to follow given the number of analytical methods employed, overall the strongest evidence of linkage was found on chromosomes 10 and 20. The strongest signal detected was on chromosome 20q13 (LOD score=3.2) using the discrete trait of BMI of 30 or higher in an affected sibpair test, however, the other methods and phenotypes also showed some evidence of linkage in this same region. Not unsurprisingly, the strongest evidence of linkage on chromosome 10 (10p) came from the quantitative analysis of BMI (P=0.0148). However, the region of potential linkage reported for both chromosomes 10 and 20 is relatively large (as much as 83 cM for the chromosome 10 signal), making localization of the signal difficult with respect to the identification of potential positional candidate genes.

French Canadians (the Quebec Family Study)

While the genome scan recently reported by Chagnon and colleagues [21] focused on phenotypes relating to body leanness, their findings are relevant to a discussion of obesity genome scans given the nature of their findings. Utilizing data collected as part of the Quebec Family Study, they conducted a genome scan for fat-free mass in a maximum of 609 pairs of extended relatives using a gene map with an average density of approximately 12 cM. They report the detection of significant linkage signals on chromosomes 15 (15q25-q26) and 18 (18q12) with LOD scores 3.6 and 3.5, respectively, as well as suggestive evidence of linkage on chromosome 7 (7p15.3) with a LOD score of 2.7. These results are of interest not only because of their focus on a measure of body composition, but also because two of the areas of linkage identified contain good positional candidate genes for obesity-related phenotypes. The chromosome 7 linkage signal spans the region containing both the neuropeptide Y (*NPY*) and growth hormone-releasing hormone (*GHRH*) receptor genes, while the chromosome 15 signal is in the region containing the insulin-like growth factor 1 receptor (*IGF1R*) gene.

Replications

While the genome scans published to date are based on what are still relatively small sample sizes, the pattern of linkage signals which is beginning to emerge is quite fascinating. Indeed, not only are there several linkage signals with significant LOD scores, but many of these are in regions which contain very strong positional candidate genes for obesity-related phenotypes. These points not withstanding, however, it is the fact that we are already beginning to detect evidence of replication of some of these findings across multiple studies which is truly the most exciting (Table 1). In particular the linkage result from the San Antonio Family Heart Study on chromosome 2 has now been replicated in the French sample [37] as well as a sample of African-Americans [67]. These findings are even more encouraging given the close proximity of the signal across the three studies [26, 37, 41, 67]. Also of particular note is the similarity in parameter estimates between the Mexican-American and African-American samples [67]. At the same time there also looks to be good evidence of replication of linkage on chromosome 20 in both the Pima Indian study [56] and in the population studied by Lee and colleagues [48]. However, this is tempered by the breadth of the linkage signal reported in both of these studies. Finally, there is potential evidence of replication for the chromo-

Table 1 Replicated QTLs from genome scans in humans for obesity-related phenotypes

Phenotypes	Chromosome location	Original report	Replication
Serum leptin levels	2p	Comuzzie et al. [26]	Hager et al. [37], Rotimi et al. [67]
BMI	10p	Hager et al. [37]	Lee et al. [48]
24-h respiratory quotient; BMI (treated as a discrete trait; BMI >30)	20q	Norman et al. [56]	Lee et al. [48]

some 10 linkage signal reported by Hager and colleagues [37] in the French study and the signal on chromosome 10 detected by Lee and coworkers [48]. Additionally, as the number of genome scans focusing on obesity continues to increase, and their results are published, there is strong reason to suspect that additional replications will be brought to light.

Conclusion

Certainly there is no denying that obesity represents a complex biological phenomenon, but this is not to suggest that the task of dissecting out the relevant genetic contribution to its expression is intractable. However, it is abundantly clear that success in this endeavor to identify the genes affecting the expression of obesity is contingent upon the application of the most appropriate analytical tools for the job along with the collection of large informative data sets. While the number of published genome scans to date focusing on obesity is still limited and their sample sizes relatively small, there already are some intriguing findings as well as potential replication of some of these findings across populations. With the completion of the human genome project now well is sight, and the rapid development of powerful new tools such as micro-array technology, the necessary infrastructure for moving from linkage to the identification of functional genetic variants will soon be in place. Indeed, given the findings from the genome scans to date, there is reason for genuine optimism that these efforts will have a tangible pay-off in the identification of genes with significant influence on the expression of obesity and its associated phenotypes.

Acknowledgements This work was supported by National Institutes of Health grants HL45522, HL28972, and MH59490.

References

1. Allison DB, Heo M, Faith MS, Pietrobelli A (1998) Meta-analysis of the association of the Trp64Arg polymorphism in the beta3 adrenergic receptor with body mass index. Int J Obes Relat Metab Disord 22:559–566
2. Almasy L, Blangero J (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 62:1198–1211
3. Almasy L, Blangero J (1999) Linkage strategies for mapping genes for complex traits in man. In: Crusio WT, Gerlai RT (eds) Handbook of molecular-genetic techniques for brain and behavior research (Techniques in the behavioral and neural sciences, vol 13. Elsevier, Amsterdam, pp 100–112
4. Almasy L, Dyer TD, Blangero J (1997) Bivariate quantitative trait linkage analysis: pleiotropy versus coincident linkage. Genet Epidemiol 14:953–958
5. Almasy L, Williams JT, Dyer TD, Blangero J (1999) Quantitative trait locus detection using combined linkage/disequilibrium analysis. Genet Epidemiol 17 [Suppl 1]:S31–S36
6. Amos CI (1994) Robust variance-components approach for assessing genetic linkage in pedigrees. Am J Hum Genet 54: 535–543
7. Amos CI, Dawson DV, Elston RC (1990) The probabilistic determination of identity-by-descent sharing for pairs of relatives from pedigrees. Am J Hum Genet 47:842–853
8. Amos CI, Zhu DK, Boerwinkle E (1996) Assessing genetic linkage and association with robust components of variance approaches. Ann Hum Genet 60:143–160
9. Amos CI, Krushkal J, Thiel TJ, Young A, Zhu DK, Boerwinkle E, de Andrade M (1997) Comparison of model-free linkage mapping strategies for the study of a complex trait. Genet Epidemiol 14:743–748
10. Bailey JN, Almasy L (1995) A brute force dichotomization approach to quantitative trait linkage analysis. Genet Epidemiol 12:719–722
11. Barsh GS, Farooqi IS, O'Rahilly S (2000) Genetics of bodyweight regulation. Nature 404:644–651
12. Blangero J (1993) Statistical genetic approaches to human adaptability. Hum Biol 65:941–966
13. Blangero J, Almasy L (1997) Multipoint oligogenic linkage analysis of quantitative traits. Genet Epidemiol 14:959–964
14. Blangero J, Williams JT, Almasy L (2000) Quantitative trait locus mapping using human pedigrees. Hum Biol 72:35–62
15. Blangero J, Williams JT, Almasy L (2001) Variance component methods for detecting complex trait loci. Adv Genet 42: 151–181
16. Borecki IB, Blangero J, Rice T, Perusse L, Bouchard C, Rao DC (1998) Evidence for at least two major loci influencing human fatness. Am J Hum Genet 63:831–838
17. Bouchard C, Perusse L (1996) Current status of the human obesity gene map. Obes Res 4:81–90
18. Brownell KD (1999) The central role of lifestyle change in long-term weight management. Clin Cornerstone 2:43–51
19. Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. Curr Opin Genet Dev 3:939–943
20. Chagnon YC, Perusse L, Bouchard C (1998) The human obesity gene map: the 1997 update. Obes Res 6:76–92
21. Chagnon YC, Borecki IB, Perusse L, Roy S, Lacaille M, Chagnon M, Ho-Kim MA, Rice T, Province MA, Rao DC, Bouchard C (2000) Genome-wide search for genes related to the fat-free body mass in the Quebec family study. Metabolism 49:203–207
22. Chagnon YC, Perusse L, Weisnagel SJ, Rankinen T, Bouchard C (2000) The human obesity gene map: the 1999 update. Obes Res 8:89–117
23. Cloninger CR, Van Eerdewegh P, Goate A, Edenberg HJ, Blangero J, Hesselbrock V, Reich T, Nurnberger J Jr, Schuckit M, Porjesz B, Crowe R, Rice JP, Foroud T, Przybeck TR, Almasy L, Bucholz K, Wu W, Shears S, Carr K, Crose C, Willig C, Zhao J, Tischfield JA, Li TK, Conneally PM, Begleiter H (1998) Anxiety proneness linked to epistatic loci in genome scan of human personality traits. Am J Med Genet 81:313–317
24. Comuzzie AG, Allison DB (1998) The search for human obesity genes. Science 280:1374–1377
25. Comuzzie AG, Blangero J, Mahaney MC, Mitchell BD, Hixson JE, Samollow PB, Stern MP, MacCluer JW (1995) Major gene with sex-specific effects influences fat mass in Mexican Americans. Genet Epidemiol 12:475–488
26. Comuzzie AG, Hixson JE, Almasy L, Mitchell BD, Mahaney MC, Dyer TD, Stern MP, MacCluer JW, Blangero J (1997) A major quantitative trait locus determining serum leptin levels and fat mass is located on human chromosome 2. Nat Genet 15:273–276
27. Curtis D, Sham PC (1994) Using risk calculation to implement an extended relative pair analysis. Ann Hum Genet 58 (Pt 2):151–162
28. Dempster ER, Lerner MI, Robertson A (1950) Heritability of threshold characters. Genetics 35:212–236
29. Duggirala R, Williams JT, Williams-Blangero S, Blangero J (1997) A variance component approach to dichotomous trait linkage analysis using a threshold model. Genet Epidemiol 14:987–992
30. Elston RC, Buxbaum S, Jacobs KB, Olson JM (1998) Haseman and Elston revisited. Genet Epidemiol 19:1–17
31. Esparza J, Fox C, Harper IT, Bennett PH, Schulz LO, Valencia ME, Ravussin E (2000) Daily energy expenditure in

Mexican and USA Pima indians: low physical activity as a possible cause of obesity. Int J Obes Relat Metab Disord 24:55–59
32. Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th edn. Prentice-Hall, Englewood Cliffs
33. Fulker DW, Cherney SS (1996) An improved multipoint sib-pair analysis of quantitative traits. Behav Genet 26:527–532
34. Goldgar DE (1990) Multipoint analysis of human quantitative genetic variation. Am J Hum Genet 47:957–967
35. Greenberg DA (1990) Linkage analysis assuming a single-locus mode of inheritance for traits determined by two loci: inferring mode of inheritance and estimating penetrance. Genet Epidemiol 7:467–479
36. Greenberg DA, Hodge SE (1989) Linkage analysis under "random" and "genetic" reduced penetrance. Genet Epidemiol 6:259–264
37. Hager J, Dina C, Francke S, Dubois S, Houari M, Vatin V, Vaillant E, Lorentz N, Basdevant A, Clement K, Guy-Grand B, Froguel P (1998) A genome-wide scan for human obesity genes reveals a major susceptibility locus on chromosome 10. Nat Genet 20:304–308
38. Hanson RL, Ehm MG, Pettitt DJ, Prochazka M, Thompson DB, Timberlake D, Foroud T, Kobes S, Baier L, Burns DK, Almasy L, Blangero J, Garvey WT, Bennett PH, Knowler WC (1998) An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians. Am J Hum Genet 63:1130–1138
39. Haseman JK, Elston RC (1972) The investigation of linkage between a quantitative trait and a marker locus. Behav Genet 2:3–19
40. Hill JO, Peters JC (1998) Environmental contributions to the obesity epidemic. Science 280:1371–1374
41. Hixson JE, Almasy L, Cole S, Birnbaum S, Mitchell BD, Mahaney MC, Stern MP, MacCluer JW, Blangero J, Comuzzie AG (1999) Normal variation in leptin levels in associated with polymorphisms in the proopiomelanocortin gene, POMC. J Clin Endocrinol Metab 84:3187–3191
42. Hong Y, Rice T, Gagnon J, Despres JP, Nadeau A, Perusse L, Bouchard C, Leon AS, Skinner JS, Wilmore JH, Rao DC (1998) Familial clustering of insulin and abdominal visceral fat: the HERITAGE Family Study. J Clin Endocrinol Metab 83:4239–4245
43. Hopper JL, Mathews JD (1982) Extensions to multivariate normal models for pedigree analysis. Ann Hum Genet 46: 373–383
44. Knowler WC, Pettitt DJ, Saad MF, Charles MA, Nelson RG, Howard BV, Bogardus C, Bennett PH (1991) Obesity in the Pima Indians: its magnitude and relationship with diabetes. Am J Clin Nutr 53:1543S-1551S
45. Kopelman PG (2000) Obesity as a medical problem. Nature 404:635–643
46. Korczak JF, Goldstein AM (1997) Sib-pair linkage analyses of nuclear family data: quantitative versus dichotomous disease classification. Genet Epidemiol 14:827–832
47. Krude H, Biebermann H, Luck W, Horn R, Brabant G, Gruters A (1998) Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. Nat Genet 19:155–157
48. Lee JH, Reed DR, Li WD, Xu W, Joo EJ, Kilker RL, Nanthakumar E, North M, Sakul H, Bell C, Price RA (1999) Genome scan for human obesity and linkage to markers in 20q13. Am J Hum Genet 64:196–209
49. Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer, Sunderland
50. MacLean CJ, Bishop DT, Sherman SL, Diehl SR (1993) Distribution of LOD scores under uncertain mode of inheritance. Am J Hum Genet 52:354–361
51. Mitchell BD, Ghosh S, Schneider JL, Birznieks G, Blangero J (1997) Power of variance component linkage analysis to detect epistasis. Genet Epidemiol 14:1017–1022
52. Mitchell BD, Blangero J, Comuzzie AG, Almasy LA, Shuldiner AR, Silver K, Stern MP, MacCluer JW, Hixson JE (1998) A paired sibling analysis of the beta-3 adrenergic receptor and obesity in Mexican Americans. J Clin Invest 101:584–587
53. Mitchell BD, Cole SA, Comuzzie AG, Almasy L, Blangero J, MacCluer JW, Hixson JE (1999) A quantitative trait locus influencing BMI maps to the region of the beta-3 adrenergic receptor. Diabetes 48:1863–1867
54. Montague CT, Farooqi IS, Whitehead JP, Soos MA, Rau H, Wareham NJ, Sewter CP, Digby JE, Mohammed SN, Hurst JA, Cheetham CH, Earley AR, Barnett AH, Prins JB, O'Rahilly S (1997) Congenital leptin deficiency is associated with severe early-onset obesity in humans. Nature 387:903–908
55. Norman RA, Thompson DB, Foroud T, Garvey WT, Bennett PH, Bogardus C, Ravussin E (1997) Genomewide search for genes influencing percent body fat in Pima Indians: suggestive linkage at chromosome 11q21–q22. Pima Diabetes Gene Group. Am J Hum Genet 60:166–173
56. Norman RA, Tataranni PA, Pratley R, Thompson DB, Hanson RL, Prochazka M, Baier L, Ehm MG, Sakul H, Foroud T, Garvey WT, Burns D, Knowler WC, Bennett PH, Bogardus C, Ravussin E (1998) Autosomal genomic scan for loci linked to obesity and energy metabolism in Pima Indians. Am J Hum Genet 62:659–668
57. Perusse L, Chagnon YC, Dionne FT, Bouchard C (1997) The human obesity gene map: the 1996 update. Obes Res 5:49–61
58. Perusse L, Chagnon YC, Weisnagel J, Bouchard C (1999) The human obesity gene map: the 1998 update. Obes Res 7: 111–129
59. Price RA, Reed DR, Lee JH (1998) Obesity related phenotypes in families selected for extreme obesity and leanness. Int J Obes Relat Metab Disord 22:406–413
60. Poston WS, Foreyt JP (1999) Obesity is an environmental issue. Atherosclerosis 146:201–209
61. Province MA, Rice T, Borecki IB, Gu C, Rao DC (2001) A multivariate and multilocus variance components approach using structural relationships to assess quantitative trait linkage via SEGPATH. Genet Epidemiol (in press)
62. Pugh EW, Jaquish CE, Sorant AJ, Doetsch JP, Bailey-Wilson JE, Wilson AF (1997) Comparison of sib-pair and variance-components methods for genomic screening. Genet Epidemiol 14:867–872
63. Rice T, Despres JP, Perusse L, Gagnon J, Leon AS, Skinner JS, Wilmore JH, Rao DC, Bouchard C (1997) Segregation analysis of abdominal visceral fat: the HERITAGE Family Study. Obes Res 5:417–424
64. Risch N (1990) Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 46:229–241
65. Risch N, Giuffra L (1992) Model misspecification and multipoint linkage analysis. Hum Hered 42:77–92
66. Rogers J, Mahaney MC, Almasy L, Comuzzie AG, Blangero J (1999) Quantitative trait linkage mapping in anthropology. Am J Phys Anthropol 110:127–151
67. Rotimi CN, Comuzzie AG, Lowe WL, Luke A, Blangero J, Cooper RS (1999) The quantitative trait locus on chromosome 2 for serum leptin levels is confirmed in African-Americans. Diabetes 48:643–644
68. Schork NJ (1993) Extended multipoint identity-by-descent analysis of human quantitative traits: efficiency, power, and modeling considerations. Am J Hum Genet 53:1306–1319
69. Shuldiner AR, Silver K, Roth J, Walston J (1996) Beta 3-adrenoceptor gene variant in obesity and insulin resistance. Lancet 348:1584–1585
70. Sing CF, Haviland MB, Reilly SL (1996) Genetic architecture of common multifactorial diseases. Ciba Found Symp 197:211–229
71. Stern MP, Duggirala R, Mitchell BD, Reinhart LJ, Shivakumar S, Shipman PA, Uresandi OC, Benavides E, Blangero J, O'Connell P (1996) Evidence for linkage of regions on chromosomes 6 and 11 to plasma glucose concentrations in Mexican Americans. Genome Res 6:724–734
72. Strauss RS, Knight J (1999) Influence of the home environment on the development of obesity in children. Pediatrics 103:E85

73. Strobel A, Issad T, Camoin L, Ozata M, Strosberg AD (1998) A leptin missense mutation associated with hypogonadism and morbid obesity. Nat Genet 18:213–215
74. Towne B, Siervogel RM, Blangero J (1997) Effects of genotype-by-sex interaction on quantitative trait linkage analysis. Genet Epidemiol 14:1053–1058
75. Turner ST, Boerwinkle E, Sing CF (1999) Context-dependent associations of the ACE I/D polymorphism with blood pressure. Hypertension 34:773–778
76. Walder K, Hanson RL, Kobes S, Knowler WC, Ravussin E (2000) An autosomal genomic scan for loci linked to plasma leptin concentration in Pima Indians. Int J Obes Relat Metab Disord 24:559–565
77. Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396
78. Whittemore AS, Halpern J (1994) Probability of gene identity by descent: computation and applications. Biometrics 50: 109–117
79. Wijsman EM, Amos CI (1997) Genetic analysis of simulated oligogenic traits in nuclear and extended pedigrees: summary of GAW10 contributions. Genet Epidemiol 14:719–735
80. Williams JT, Blangero J (1999) Comparison of variance components and sibpair-based approaches to quantitative trait linkage analysis in unselected samples. Genet Epidemiol 16: 113–134
81. Williams JT, Blangero J (1999) Power of variance component linkage analysis to detect quantitative trait loci. Ann Hum Genet 63:545–563
82. Williams JT, Duggirala R, Blangero J (1997) Statistical properties of a variance components method for quantitative trait linkage analysis in nuclear families and extended pedigrees. Genet Epidemiol 14:1065–1070
83. Williams JT, Van Eerdewegh P, Almasy L, Blangero J (1999) Joint multipoint linkage analysis of multivariate qualitative and quantitative traits. I. Likelihood formulation and simulation results. Am J Hum Genet 65:1134–1147
84. World Health Organization (1995) Physical status: the use and interpretation of anthropometry. Report of a WHO expert committee. Technical reports series, no 854. WHO, Geneva
85. Woods SC, Seeley RJ, Porte D Jr, Schwartz MW (1998) Signals that regulate food intake and energy homeostasis. Science 280:1378–1383
86. Zerba KE, Ferrell RE, Sing CF (1996) Genotype-environment interaction: apolipoprotein E (ApoE) gene effects and age as an index of time and spatial context in the human. Genetics 143:463–478