Gene × Environment Interactions in Obesity: The State of the Evidence

Shafqat Ahmad^a Tibor V. Varga^a Paul W. Franks^{a–c}

^aGenetic and Molecular Epidemiology Unit, Department of Clinical Science, Lund University, Malmö, and

^bDepartment of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden; ^cDepartment of Nutrition, Harvard School of Public Health, Boston, Mass., USA

Key Words

Obesity · Gene × environment interaction · Statistical power · Systematic review

Abstract

Background/Aims: Obesity is a pervasive and highly prevalent disease that poses substantial health risks to those it affects. The rapid emergence of obesity as a global epidemic and the patterns and distributions of the condition within and between populations suggest that interactions between inherited biological factors (e.g. genes) and relevant environmental factors (e.g. diet and physical activity) may underlie the current obesity epidemic. **Methods:** We discuss the rationale for the assertion that gene × lifestyle interactions cause obesity, systematically appraise relevant literature, and consider knowledge gaps future studies might seek to bridge. **Results:** We identified >200 relevant studies, of which most are relatively small scale and few provide replication data. **Conclusion:** Although studies on gene × lifestyle interactions in obesity point toward the presence of such interactions, improved data standardization, appropriate pooling of data and resources, innovative study designs, and the application of powerful statistical methods will be required if translatable examples of gene × lifestyle interactions in obesity are to be identified. Future studies, of which

most will be observational, should ideally be accompanied by appropriate replication data and, where possible, by analogous findings from experimental settings where clinically relevant traits (e.g. weight regain and weight cycling) are outcomes.

© 2013 S. Karger AG, Basel

Introduction

Obesity is a natural consequence of chronic positive energy balance, where the caloric density of the diet exceeds the amount of energy expended through resting and non-resting energy metabolism. In simple terms, obesity predisposition should be modifiable by increasing volitional energy expenditure (through physical activity) and reducing energy intake (through diet restraint). However, there are an estimated 1.5 billion overweight or obese persons living today [1], many of whom will have unsuccessfully sought to lose weight through diet and exercise; indeed, although short-term weight loss is often achievable with lifestyle modification, lost weight is often regained in the long run. Thus, powerful biologic mechanisms are in place that cause human brains and bodies to resist weight loss and aspire toward weight regain. These dynamic characteristics are a legacy of our evolution, where an inherent

ability to lay down energy reserves in the form of adipose tissue and to protect those reserves from depletion was advantageous to survival and fecundity [2, 3]. Yet, contemporary environments contrast those in which our ancestors lived, in that only very recently have occupations been heavily automated and energy-dense foods abundant and accessible with minimal effort. Nonetheless, there is clearly considerable variability in the extent to which persons living within obesogenic environments become obese. For example, migration studies [4, 5], other comparisons of high-risk ethnic groups living in starkly contrasting environmental settings [6], and randomized controlled trials of weight loss interventions [7] illustrate the powerful influence of the environment on a person's predisposition to obesity and how people vary in the extent to which they become obese in the presence of obesogenic environments. These observations have spurred speculation that common obesity is the consequence of complex interactions between genetic and environmental factors.

The aims of this paper are to discuss the concept of gene $\times$ environment interactions in obesity, to review the published evidence supporting or refuting the presence of such interactions, and to appraise how reliable evidence of such interactions might facilitate the prediction, prevention or treatment of obesity and its pathological consequences.

A Brief Overview of the Genetics of Obesity

Accumulation of excess adipose tissue is a highly heritable trait, with up to 90% of its variance in children owing to familial factors, much attributable to shared genes [8]. Indeed, although families often share environmental and social factors that predispose to obesity, studies in twins [9–25], particularly those in twins adopted into different families soon after birth [26–28], suggest that the strong, lifelong sibling resemblance in body composition is under strong genetic influence. Seeking the identities of the specific genetic loci that influence body composition has been the subject of almost 3 decades of research.

Animal and human positional cloning studies focusing on extreme early-onset obesity revealed the identities of a range of single gene defects, primarily in hypothalamically expressed genes regulating appetite, taste preference, and satiety [29–31]. The search for genetic variants causing common forms of human obesity essentially failed until 2007, when 3 genetic association studies simultaneously discovered a region on chromosome 16 located within the first intron of the fat mass and obesity-associated gene (*FTO*) [32–34], which encodes the pro-

tein α -ketoglutarate-dependent dioxygenase [35]. The discovery of this region marked the beginning of a new era in obesity genetics, where emphasis shifted from relatively small, hypothesis-driven studies, to very large-scale, 'hypothesis-free' genome-wide associations studies (GWAS), where millions of tests are performed in parallel to identify genetic association signals.

GWAS hinge on the use of high-throughput massively parallel genetic microarrays to measure millions of variants in large population-based cohort collections and case-control studies. The obvious caveat to the GWAS approach is that it is prone to false discovery, owing to the impact of multiple testing on type 1 error rates. To overcome this problem, very large sample collections are now used to ensure studies are adequately powered to detect small effects at a conservative genome-wide probability threshold ($p = 5 \times 10^{-8}$). In 2006, the first GWAS for obesity was published that reported an association with body mass index (BMI) in 694 participants from the Framingham Heart Study for a variant at the *INSIG2* locus [36], but this finding has never been convincingly replicated. For example, the effect of this single nucleotide polymorphism (SNP) on BMI in 123,748 individuals included in the Genetic Investigation of Anthropometric Traits (GIANT) meta-analysis is not statistically significant ($p_{\text{nominal}} = 0.13$), although the direction of effect is the same as in the original report [37]. Notwithstanding the false start, GWAS has evolved into a widely applied, powerful approach for the detection of genetic association signals for common complex traits, and with this the standards and scale of GWAS have also improved to the point where GWAS results from large consortia such as GIANT are generally very reliable.

Despite the relative success of GWAS in terms of discovering reliable genetic association signals for a range of obesity-associated traits [37–39], the identified GWAS variants explain <5% of the overall heritable component of these traits [37, 40]. Moreover, genetic variants discovered using GWAS for complex metabolic diseases such as type 2 diabetes and obesity have proven ineffective for the prediction of these traits [41–43], rendering these discoveries of little or no direct clinical value. However, GWAS have provided valuable information about the biologic pathways and networks that underlie many complex traits, which may lead to the development of drug targets for treating morbid obesity and its pathologic consequences. Many of the top-ranked obesity-associated variants from GWAS reside within neuronal pathways; the loci proximal to some of these variants encode centrally expressed genes, several of which are known to control

addictive behaviors [39], appetite, satiation, and taste preference [39, 44]. For example, *Sh2b1*-deficient mice are hyperphagic and, albeit to an insufficient extent to prevent obesity, hyperenergetic [45, 46]. Another extensively investigated obesity susceptibility locus is *MC4R*, which is expressed in a range of human tissues including pituitary cells, skin cells, and hypothalamic neurons and mediates the effects of melanocortins on feeding behavior [47]. *MC4R* deficiency is the most common cause of monogenic early-onset obesity [48], which predisposes affected persons to hyperphagia and early-onset morbid obesity [49]. Knockdown studies show that some of these loci are likely to regulate adipocyte regulation. Moreover, expression of several of these loci changes during adipogenesis and is sensitive to insulin and dexamethasone (a synthetic glucocorticoid) [50].

GWAS has also revealed a distinct sex dimorphism in the genetics of obesity, with some variants conveying statistically different effects between men and women and others appearing to convey effects that are evident in only one of the sexes [51]. For example, Renström et al. [43] were amongst the first to report sex-dimorphic effects for GWAS-derived loci; the authors reported a significant gene $\times$ sex interaction for an *MC4R* variant ($p_{\text{interaction}} = 0.02$), with the minor allele being associated with BMI in women ($\beta = 0.41/\text{minor allele}$; $p = 0.0034$), but not in men ($\beta = 0.03/\text{minor allele}$; $p = 0.83$). Comparable effects were reported for total ($p_{\text{interaction}} = 0.024$) and abdominal ($p_{\text{interaction}} = 0.046$) adipose fat. Elsewhere, Fox et al. [52] reported a strong association between the rs1659258 *THNSL2* variant and visceral fat in women ($p = 1.6 \times 10^{-08}$) that was not evident in men ($p = 0.75$).

What Is a Gene $\times$ Environment Interaction?

The term gene $\times$ environment interaction is used widely and often with different meanings. In epidemiology, interactions are usually defined statistically by estimating whether the degree of risk attributable to the joint effects of a genotype and an environmental factor on an outcome (e.g. disease/no disease) is greater or less than would be expected if these joint effects were additive. An alternative, yet compatible description of gene $\times$ environment interaction is one where the risk conveyed by specific genotypes depends on the level of one or more environmental exposures. This definition is useful when thinking in terms of interventions designed to offset genetic risk where the environmental exposures can be intervened upon, as is the case with lifestyle factors such as

diet, physical activity, and smoking. Pharmacogenetics is a special area of gene $\times$ environment interaction research, where the ‘environmental’ exposure is medication. Here, the objective is often to determine whether genetic factors impact the efficacy of specific drugs or the extent to which a drug’s side effects occur. We do not address this topic in depth here and instead refer the reader to 2 excellent reviews on this topic [53, 54].

It is sometimes assumed that gene $\times$ environment interactions only occur between genetic and environmental factors that convey statistically significant marginal effects (i.e. the genetic and environmental factors are directly associated with the outcome in univariate analyses), but this is not always the case as there are plausible scenarios (i.e. with certain types of crossover interactions) where neither the genetic nor environmental exposures convey detectable risk, yet together they convey a strong interaction effect. Indeed, crossover interactions are likely to be the most clinically relevant type of gene $\times$ environment interaction, as someone at high genetic risk of disease given one environmental setting may be genetically protected from disease if the environmental exposures are improved. Nevertheless, crossover interactions are likely to be rare and searching for interactions between genetic and environmental exposures that convey detectable marginal effects on disease predisposition may be an effective data reduction strategy.

A related false assumption is that if a genotype and environmental exposure both convey strong independent effects, they are good candidates for interaction. However, the most clinically relevant statistical interactions are those that reflect causal (biologic) processes. Thus, in interaction studies, environmental exposures that perturb the activity of the selected gene(s) are those most likely to engage in causal interactions. It is also important to bear in mind that most reported examples of gene $\times$ environment interactions are unlikely to reflect causal processes; alternative explanations for observed interactions are many, but include scale dependency, confounding, reporting biases, and methodological errors (see [55] for causal inference criteria for studies of gene $\times$ environment interaction). Indeed, studies of gene $\times$ environment interaction are prone to many of the confounding factors and biases that affect the interpretation of data in conventional genetic and non-genetic association studies. Those that relate to the genetic component of interaction studies include population stratification (where latent subgroups within a study population differ in their allele and disease frequencies from other subgroups) and synthetic association/confounding by linkage disequilibrium (where the

observed locus does not directly affect gene function but is correlated with a functional locus). Other sources of bias common in epidemiological studies of gene $\times$ environment interaction are discussed in detail elsewhere [56] and include:

- *survival* bias (where higher-risk individuals have died prior to the study and thus cannot be included in the study population);
- *respondent* bias (where rates of refusal or non-response between cases and non-cases differ);
- *diagnosis* bias (where the intensity and outcome of the diagnostic process are enhanced because the participant's exposure to a relevant risk factor is known);
- *referral* or *admission rate* bias (where cases are more likely than non-cases to receive medical interventions);
- *surveillance* bias (where subclinical disease is more frequently detected in persons who are under medical surveillance than in those who are not);
- *measurement* and *outcome* biases (where measurements or outcomes are measured more accurately in cases or non-cases);
- *recall* bias (where questions about risk factors and exposures are either asked differently or responded to more accurately by cases or non-cases);
- *family information* bias (where information about a disease is more likely to be shared within a family if one of its members already has the disease), and
- *exposure suspicion* bias (where the search for a causal exposure is likely to be more intense and yield a different outcome in persons known to have a disease than in those who are not).

Systematic Review Methods

To facilitate a structured and comprehensive appraisal of the literature on gene $\times$ environment (diet and exercise) interaction in obesity, we conducted a systematic review (online supplementary table 1; see www.karger.com/doi/10.1159/000351070) through which we identified 210 studies that met our inclusion criteria. We built our search on 2 earlier systematic reviews conducted by us [57, 58] and then applied a systematic search strategy to the major scientific literature databases (search strategy shown below for PubMed and adapted for the Cochrane Library, the Web of Science, and EMBASE). The search was restricted to the period from January 1, 1995, through September 1, 2012: ['gene environment interaction' AND (obesity OR body weight OR BMI OR waist OR adiposity)] OR ['gene lifestyle interaction' AND

(obesity OR body weight OR BMI OR waist OR adiposity)] OR ['gene diet interaction' AND (obesity OR body weight OR BMI OR waist OR adiposity)] OR ['gene nutrient interaction' AND (obesity OR body weight OR BMI OR waist OR adiposity)] OR ['gene physical activity interactions' AND (obesity OR body weight OR BMI OR waist OR adiposity)] OR ['nutrition gene interactions' AND (obesity OR body weight OR BMI OR waist OR adiposity)] OR ['gene exercise interactions' AND (obesity OR body weight OR BMI OR waist OR adiposity)]. In addition to the general search strategy outlined above, we also queried the Cochrane Library, The Web of Science, and EMBASE using the following search terms: gene, environment, lifestyle, diet, nutrient, physical activity, nutrition, exercise, obesity, BMI, body weight, and waist.

Contact was made with content experts and authors of selected articles to maximize the inclusion in this review of relevant publications that were not captured in the earlier reviews [57, 58] or via our systematic search strategy. We included all the published studies designed to assess the combined relationship of genetic and lifestyle factors in obesity and related traits in children, adolescents, and adults. Studies not conducted in human populations and not published in English language journals were excluded. The automated search identified 856 articles matching the criteria outlined above. Two reviewers (S.A. and T.V.V.) worked in parallel to review publication titles and meta-data to exclude 438 articles that, although being captured by the automated search, were not relevant to the topic of this review article (typically because the papers did not describe original research or were not reporting studies on gene $\times$ environment interactions). Of the remaining 418 papers, a further 150 articles were excluded by the reviewers after careful inspection of the articles' abstracts. The 268 remaining articles were then obtained and read in full by S.A. and T.V.V. to further evaluate their relevance to this review, abstract relevant meta-data, and classify into study type (summarized in online suppl. table 1). This last step resulted in the exclusion of 58 articles, leaving a final total of 210 papers for inclusion in this review. A third reviewer (P.W.F.) arbitrated over the process outlined above to resolve inconsistencies emerging from the independent reviews conducted by S.A. and T.V.V.

State of the Evidence

The most widely cited and replicated example of a gene $\times$ environment interaction in obesity is for the *FTO* (rs9939609) locus and self-reported physical activity. The

initial study included 5,138 Danish adults and reported that the established effect of the rs9939609 variant on BMI was present in physically inactive participants, but in those who reported moderate to high levels of physical activity, the genetic effect was substantially reduced ($p_{\text{interaction}} = 0.007$) [59]. Soon after the publication of the Danish study, a second small epidemiological study ($n = 704$), conducted within a population of Amish adults from Pennsylvania in which physical activity had been objectively assessed [60], reported a directionally consistent and nominally statistically significant interaction effect ($p_{\text{interaction}} = 0.01$), albeit for an *FTO* variant (rs1861868) which in Whites is in low linkage disequilibrium with rs9939609. A third analysis of clinical trial data from the Diabetes Prevention Program (DPP) [61], a randomized controlled trial of intensive lifestyle modification [7], found no evidence of gene $\times$ lifestyle (vs. placebo control) interactions on 1-year change in BMI, but found nominal evidence of an interaction on changes in subcutaneous adipose accumulation ($p_{\text{interaction}} = 0.05$). In 2009, many replication studies were published, with approximately half of these studies reporting no evidence of an interaction effect and the other half reporting directionally consistent results to the original studies (online suppl. table 1). Because of our involvement in the DPP publication [61] and 1 of the replication studies [62], as well as our general interest in this field, we initiated a prospective, systematic meta-analysis with the intention of confirming or refuting the interaction between *FTO* and physical activity in obesity [63]. Although it is not uncommon to undertake retrospective meta-analyses of published results on gene $\times$ environment interactions, there are major caveats to such studies, which means that the results of such may not be valid (we refer the reader to Palla et al. [64] for an eloquent description of these caveats).

Our meta-analysis included data from 45 adult cohorts and 9 pediatric cohorts; standardized analyses were performed in each cohort, and these results were meta-analyzed, yielding a statistically significant interaction effect that was directionally consistent with, but of much smaller magnitude than, the original study's findings [59]; moreover, the interaction effect was only evident in the North American cohorts included in the meta-analysis, despite data from many more European ($n = 164,307$) than North American ($n = 47,938$) participants being included in the study. Although our meta-analysis is generally taken as confirmation of the Danish study's [59] findings, it remains unclear whether this interaction effect is causal [65]. Thus, much work remains to be conducted in

the area of *FTO* $\times$ lifestyle interactions, if these data are ever to be of use in clinical practice.

Even with the uncertainties surrounding the interpretation of the *FTO* $\times$ physical activity interaction results described above, it remains the most well-replicated example to date of a gene $\times$ environment interaction in obesity; thus, for the purpose of this review, we consider it appropriate to view the published interaction effect size reported by Andreasen et al. [59] as representative of the upper end of the range of interaction effect sizes that could be expected for an interaction between a common gene variant and a common lifestyle exposure in obesity. We assume that, as with the discovery of obesity-associated gene variants [37], the discovery of interaction effects will follow an asymptotic pattern, where the effect sizes of newly discovered interactions will diminish in magnitude over time, providing hypothesis-free methods continue to predominate and sample sizes continue to grow, which seem likely. Thus, we anticipate that interaction effects in obesity for common variants and common environmental exposures will generally be of smaller magnitude than for the *FTO* interaction effect reported by Andreasen et al. [59].

The minimum sample size required to afford at least 80% power to detect the interaction effect reported by Andreasen et al. [59], assuming comparable cohort characteristics, is roughly 5,000 participants, and smaller magnitude interaction effects will demand considerably larger sample collections to ensure adequately powered analyses. This is especially true when replication analyses are performed in multiple cohorts with diverse characteristics and those results are meta-analysed [93], as is often done these days.

Figure 1 shows the sample sizes of all studies included in our systematic review, the majority of which report statistically significant interaction effects. The sample sizes of these studies range from 34 to 237,434, with more than half including fewer than 1,000 participants. Indeed, it was not until 2008 that the average size of gene $\times$ environment interaction studies in obesity exceeded 5,000 participants, with only 25 of 210 studies (12% of all published studies included in online suppl. table 1) including >5,000 participants (most of which sought to replicate the interaction between *FTO* and physical activity). Whilst small studies will occasionally yield true-positive results (particularly when testing well-informed a priori hypotheses) and large studies will sometimes yield false-positive (and false-negative) findings, our review of the published literature illustrates that many studies may have been underpowered owing to small sample sizes and subjective

Fig. 1. Relationship between sample sizes of studies of gene $\times$ lifestyle interactions in obesity ($n = 210$ studies) published since 1995 and power to detect the interaction effect reported by Andreassen et al. [59] for *FTO* (rs9960939) $\times$ physical activity on BMI (dashed curve).

assessments of physical activity (documented in online suppl. table 1). The absence of replication data for most reports of gene $\times$ lifestyle interactions in obesity may indicate a preponderance of false-positive findings amongst published studies.

Two of the most exciting reports of gene $\times$ environment interactions in obesity were published recently by Qi et al. [67, 68]. These studies focused on testing interactions for a 32-SNP genetic risk score with either sugar-sweetened beverage intake [68] or physical activity [67]. In the first study, the authors reported that the association between sugar-sweetened beverage consumption and obesity was much stronger in persons with a high genetic burden, whereas the second study reported a similar interaction effect for physical inactivity, where the association of TV viewing and obesity was strongest in those persons with a high genetic risk score.

The largest study published to date related to gene $\times$ physical activity interactions in obesity that did not focus on the *FTO* locus was a quantitative analysis con-

ducted using data from the National Runners' Health Study ($n = 47,691$ participants). In this cross-sectional study, Williams [69] used data from self-administered questionnaires in which participants were asked to record their body composition, physical activity habits, and their parents' body composition. The study showed a strong positive relationship between BMI and waist circumference in the participants and their parents in participants who ran <3 km/day, but this relationship diminished in strength and magnitude as self-reported distance run increased. This study suggests the presence of gene $\times$ physical activity interactions in obesity; however, genetic and non-genetic factors influence the coalescence of obesity in families as demonstrated in studies of biologically unrelated social networks where a person's body composition is influenced by the body composition of her or his peers [70]. Moreover, all data in Williams's study [69] were obtained through a questionnaire administered by mail to the participant, and one cannot exclude the possibility that recall of body

composition (the participant's and his or her parents') is related to the activity level and/or body composition of the participant.

In a study of 20,249 adults from the UK, Kilpeläinen et al. [71] used the population-based EPIC-Norfolk cohort to seek replication of the associations between variants (rs6232 and rs6235) at *PCSK1* and obesity. The authors were unable to confirm the original study's findings and also found no evidence that physical activity or sex modifies the associations of rs6232 and rs6235 with obesity. However, the authors observed a gene $\times$ age interaction ($p_{\text{interaction}} = 0.021$), where the association of the rs6232 variant and obesity was detectable in younger (40–59 years) but not in older (>59 years) participants.

In the same cohort, Li et al. [66] performed analyses of 12 established obesity-associated variants and their interaction with physical activity in obesity. The authors reported that the association between a genetic risk score (comprised of the 12 variants) and obesity was diminished in persons reporting higher levels of physical activity. A limitation of this study is that these results were not statistically significant when people with prevalent cardiovascular disease and cancer were excluded from the analysis. It is possible that excluding these individuals, which amounts to more than one quarter of the study population ($n = 1,128$ for cardiovascular disease and $n = 4,534$ for cancer), diminished the study's statistical power such that the interaction effect was no longer detectable; it may also be that disease labeling bias (which can occur when people diagnosed with lifestyle-associated diseases misreport lifestyle behaviors) or other sources of bias associated with inclusion (e.g. survival bias and competing risk bias [72]) may have influenced the results.

Andreasen et al. [73] tested gene $\times$ environment (physical activity) and gene $\times$ gene interactions in a collection of 4 cohorts comprising 18,014 Danish participants. Variants at the *INSIG2* (rs7566605), *PFKP* (rs6602024), *FTO* (rs9939609), and *MC4R* (rs17782313) loci were assessed, and obesity was the outcome. The authors reported a statistically significant interaction between *INSIG2* rs7566605 and self-reported physical activity. No other gene $\times$ physical activity or gene $\times$ gene interactions were observed. As described above, the *INSIG2* locus is of interest owing to it being the first locus reported for obesity from a GWAS [36], yet no convincing replication of these results has been described, which may be owing to interactions with lifestyle factors. Indeed, meta-analyses of the effect of *INSIG2* rs7566605 across multiple populations illustrate that the magnitude of the genetic effect on BMI

varies considerably [74, 75], which may be attributable to gene $\times$ environment interactions. A separate analysis in the DPP showed that the minor C allele at the rs7566605 genotype modified the effects of the DPP lifestyle intervention (vs. placebo control) on change in body weight ($p_{\text{interaction}} = 0.02$), subcutaneous fat ($p_{\text{interaction}} = 0.01$), and visceral fat ($p_{\text{interaction}} = 0.02$) [61].

Several studies have reported on interactions of *FTO* variants and dietary factors in obesity. Studies in participants from Sweden [76, 77] have reported interactions with dietary factors (dietary, fiber, protein, carbohydrate, and fats). In these studies, low carbohydrate intake and high fat intake appeared to augment the effects of the *FTO* variant on BMI. No interactions were detected for fiber or protein intake. Future studies on this topic might consider accounting for the interdependency between the studied dietary factors (dietary carbohydrate and fat intake were moderately to strongly correlated with each other in this study) or between dietary factors and physical activity (which has been established in this population [77] and elsewhere [63] to interact with the *FTO* locus).

A number of clinical trial analyses of gene $\times$ lifestyle interactions have emerged in the literature in recent years [61, 78–83]. We have reported several such examples from the DPP. The advantages of assessing gene $\times$ lifestyle interactions in randomized clinical trials are that the environmental exposures (lifestyle intervention) are relatively well controlled and monitored, although few trials have adequately assessed participants' exposure to environmental factors outside the hours of the intervention. Well-conducted randomized controlled trials also have the advantages of being prospective in design, mitigating the possibility of reverse causation, and are less prone to confounding than epidemiological studies. The main disadvantages of clinical trials are that, even in relatively large trials such as the DPP, they lack statistical power to detect interactions [84, 85], and because most trials involve specific interventions, finding appropriate replication studies is often difficult. Nevertheless, our analyses in the DPP have revealed several interesting results on gene $\times$ lifestyle interactions; for example, we recently reported data on genetic predictors of weight gain following intentional weight loss with lifestyle intervention [42]. In this study, we showed that persons with a greater genetic burden (defined on the basis of a genetic risk score comprised of confirmed obesity susceptibility loci) were more likely to regain weight after initial weight loss following lifestyle intervention than were those participants with a lesser genetic burden.

Where Are the Knowledge Gaps?

Almost all studies of gene $\times$ environment interactions in obesity have focused on either biologic candidates (most studies prior to 2008) or variants identified through recent GWAS meta-analyses focused on associations with obesity traits (most studies since 2008). In general, the biologic candidate gene approach has been unsuccessful for the discovery of common variants for complex traits, as illustrated by the depressing lack of replicated genetic association studies using this method. The general failure of the biologic candidate gene approach is largely attributable to the misplaced faith we as authors, journal reviewers, and journal editors placed in the strength of the *a priori* evidence drawn upon for genetic association studies. The GWAS approach by contrast is not inhibited in this way, but the manner in which associated variants are usually discovered in GWAS may mean that they are unlikely to be involved in gene $\times$ environment interactions (see [86] for further discussion of this point). Thus, genome-wide analyses that focus on explicit or inferential tests of interaction are likely to prove most successful when seeking out novel examples of gene $\times$ environment interactions. Conducting conventional pairwise tests of gene $\times$ environment interactions on a genome-wide scale is impracticable, owing to the enormous sample sizes required to afford sufficient statistical power; hence, other relatively powerful inferential approaches may prove more successful (reviewed in detail in [86]). The joint meta-analysis method developed by Manning et al. [87], which builds on the 2 degree of freedom model popularized by Kraft et al. [88], has been applied with some success for the assessment of gene $\times$ BMI interactions in glycemic traits [89]. Two-stage models have also been proposed, where sub-cohorts are selected from the tails of the environmental exposure distribution in order to leverage the discovery of interaction effects [90]. Variance prioritization methods, which focus on determining the extent to which phenotypic variances vary across each genotype at a given (usually SNP) locus, have also yielded encouraging results; Pare et al. [91], for example, used a Leaven's-based variance prioritization approach to identify gene $\times$ BMI and gene $\times$ smoking interactions in relation to leptin and ICAM1 levels. A similar approach has also been applied by Yang et al. [92] to identify loci likely to modify the effects of obesogenic environmental factors; their study yielded one genome-wide result, for the *FTO* locus, which – given other data implicating this locus in gene $\times$ environment interactions (see above) – provides reassur-

ance that this method works. However, with only a single locus detected using this approach, despite the study's formidable sample size, other more powerful coordinated approaches are likely necessary.

As scientists interested in gene $\times$ environment interactions, we should also turn our attention to clinically relevant traits. Although cross-sectional studies of BMI are relatively convenient to perform, identifying genetic predictors of dynamic phenotypes such as weight regain following intentional weight loss is likely to yield some of the most clinically translatable findings. Thus, the future success of gene $\times$ environment interaction studies in obesity (and other complex traits) will depend on: (1) the development of new statistical approaches; (2) the continued cooperation of scientists to share resources and data for initial analyses; (3) journal editors mandating that when reports of novel gene $\times$ environment interactions are submitted for publication, these are accompanied by data that supports the veracity of the original findings (ideally in the form of replication studies), and (4) results are confirmed in diverse populations and/or experimental settings.

Summary and Conclusions

More than 200 studies have been reported in the literature purporting to have discovered examples of gene $\times$ lifestyle interactions in obesity (online suppl. table 1); the predominance of small studies that lack replication suggests that many published studies on this topic may be false positives. The field of gene $\times$ environment interaction research has advanced greatly in recent years, with the adoption of more rigorous and powerful approaches for the detection of interactions yielding at least two reliable examples of gene $\times$ lifestyle interactions in obesity [63, 68]. However, neither of these examples appears to be of direct clinical relevance, as the magnitude of the interaction effect estimates is fairly small. Despite the absence of replication data for other gene $\times$ environment interaction studies, some are no doubt true positive. Indeed, it is possible that large-magnitude interaction effects that are cohort specific and which may be detectable in relatively small cohort collections exist, and that some of the studies described in this paper report interactions of this nature. Thus, further exploration of interaction results from epidemiological studies in clinical trials may help determine which are likely to be causal and which are owing to other factors such as reverse causality and confounding.

References

- 1 International Association of the Study of Obesity. www.iaso.org/policy/aboutobesity/ (accessed December 30, 2012).
- 2 Bellisari A: Evolutionary origins of obesity. *Obes Rev* 2008;9:165–180.
- 3 Ulijaszek SJ: Seven models of population obesity. *Angiology* 2008;59(2 suppl):34S–8S.
- 4 Antecol H, Bedard K: Unhealthy assimilation: why do immigrants converge to American health status levels? *Demography* 2006;43:337–360.
- 5 Ebrahim S, et al: The effect of rural-to-urban migration on obesity and diabetes in India: a cross-sectional study. *PLoS Med* 2010;7:e1000268.
- 6 Schulz LO, et al: Effects of traditional and western environments on prevalence of type 2 diabetes in Pima Indians in Mexico and the US. *Diabetes Care* 2006;29:1866–1871.
- 7 Knowler WC, et al: Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002;346:393–403.
- 8 Franks PW, et al: Habitual physical activity in children: the role of genes and the environment. *Am J Clin Nutr* 2005;82:901–908.
- 9 Heitmann BL, et al: Are genetic determinants of weight gain modified by leisure-time physical activity? A prospective study of Finnish twins. *Am J Clin Nutr* 1997;66:672–678.
- 10 Schur EA, Heckbert SR, Goldberg JH: The association of restrained eating with weight change over time in a community-based sample of twins. *Obesity* 2010;18:1146–1152.
- 11 Williams PT: Attenuated inheritance of body weight by running in monozygotic twins. *Med Sci Sports Exerc* 2012;44:98–103.
- 12 Rebollo-Mesa I, Ordonana JR: Childbirth moderates the genetic and environmental influences on BMI in adult twins. *Obesity (Silver Spring)* 2009;17:1646–1647.
- 13 Graff M, et al: The combined influence of genetic factors and sedentary activity on body mass changes from adolescence to young adulthood: the National Longitudinal Adolescent Health Study. *Diabetes Metab Res Rev* 2011;27:63–69.
- 14 Johnson W, et al: Education modifies genetic and environmental influences on BMI. *PLoS One* 2011;6:e16290.
- 15 McCaffery JM, et al: Gene $\times$ environment interaction of vigorous exercise and body mass index among male Vietnam-era twins. *Am J Clin Nutr* 2009;89:1011–1018.
- 16 Samaras K, et al: Genes versus environment. The relationship between dietary fat and total and central abdominal fat. *Diabetes Care* 1998;21:2069–2076.
- 17 Dubois L, et al: Genetic and environmental contributions to weight, height, and BMI from birth to 19 years of age: an international study of over 12,000 twin pairs. *PLoS One* 2012;7:e30153.
- 18 Silventoinen K, Kaprio J, Yokoyama Y: Genetic regulation of pre-pubertal development of body mass index: a longitudinal study of Japanese twin boys and girls. *Behav Genet* 2011;41:234–241.
- 19 Rokholm B, et al: Increased genetic variance of BMI with a higher prevalence of obesity. *PLoS One* 2011;6:e20816.
- 20 Greenfield JR, et al: Moderate alcohol consumption, dietary fat composition, and abdominal obesity in women: evidence for gene-environment interaction. *J Clin Endocrinol Metab* 2003;88:5381–5386.
- 21 Silventoinen K, et al: Modification effects of physical activity and protein intake on heritability of body size and composition. *Am J Clin Nutr* 2009;90:1096–1103.
- 22 Mustelin L, et al: Physical activity reduces the influence of genetic effects on BMI and waist circumference: a study in young adult twins. *Int J Obes* 2009;33:29–36.
- 23 Karnehed N, et al: Physical activity, diet and gene-environment interactions in relation to body mass index and waist circumference: the Swedish Young Male Twins Study. *Public Health Nutr* 2006;9:851–858.
- 24 Cornes BK, et al: Replication of the association of common rs9939609 variant of FTO with increased BMI in an Australian adult twin population but no evidence for gene by environment (G $\times$ E) interaction. *Int J Obes* 2009;33:75–79.
- 25 Watson NF, et al: Sleep duration and body mass index in twins: a gene-environment interaction. *Sleep* 2012;35:597–603.
- 26 Biron P, Mongeau JG, Bertrand D: Familial resemblance of body weight and weight/height in 374 homes with adopted children. *J Pediatr* 1977;91:555–558.
- 27 Sorensen TI, Holst C, Stunkard AJ: Childhood body mass index – genetic and familial environmental influences assessed in a longitudinal adoption study. *Int J Obes Relat Metab Disord* 1992;16:705–714.
- 28 Stunkard AJ, et al: The body-mass index of twins who have been reared apart. *N Engl J Med* 1990;322:1483–1487.
- 29 Bell CG, et al: Association of melanin-concentrating hormone receptor 1 5' polymorphism with early-onset extreme obesity. *Diabetes* 2005;54:3049–3055.
- 30 Hinney A, et al: Systematic mutation screening of the pro-opiomelanocortin gene: identification of several genetic variants including three different insertions, one nonsense and two missense point mutations in probands of different weight extremes. *J Clin Endocrinol Metab* 1998;83:3737–3741.
- 31 Loos RJ: The genetic epidemiology of melanocortin 4 receptor variants. *Eur J Pharmacol* 2011;660:156–164.
- 32 Dina C, et al: Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat Genet* 2007;39:724–726.
- 33 Frayling TM, et al: A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007;316:889–894.
- 34 Scuteri A, et al: Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet* 2007;3:e115.
- 35 Gerken T, et al: The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science* 2007;318:1469–1472.
- 36 Herbert A, et al: A common genetic variant is associated with adult and childhood obesity. *Science* 2006;312:279–283.
- 37 Speliotes EK, et al: Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 2010;42:937–948.
- 38 Heard-Costa NL, et al: NRXN3 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. *PLoS Genet* 2009;5:e1000539.
- 39 Willer CJ, et al: Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat Genet* 2009;41:25–34.
- 40 Yang J, et al: Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet* 2011;43:519–525.
- 41 de Miguel-Yanes JM, et al: Genetic risk reclassification for type 2 diabetes by age below or above 50 years using 40 type 2 diabetes risk single nucleotide polymorphisms. *Diabetes Care* 2011;34:121–125.
- 42 Delahanty LM, et al: Genetic predictors of weight loss and weight regain after intensive lifestyle modification, metformin treatment, or standard care in the Diabetes Prevention Program. *Diabetes Care* 2012;35:363–366.
- 43 Renström F, et al: Replication and extension of genome-wide association study results for obesity in 4923 adults from northern Sweden. *Hum Mol Genet* 2009;18:1489–1496.
- 44 Berthoud HR, Morrison C: The brain, appetite, and obesity. *Annu Rev Psychol* 2008;59:55–92.
- 45 Mizuno TM, et al: Transgenic neuronal expression of proopiomelanocortin attenuates hyperphagic response to fasting and reverses metabolic impairments in leptin-deficient obese mice. *Diabetes* 2003;52:2675–2683.
- 46 Ren D, et al: Neuronal SH2B1 is essential for controlling energy and glucose homeostasis. *J Clin Invest* 2007;117:397–406.
- 47 Yeo GS, et al: The role of melanocortin signaling in the control of body weight: evidence from human and murine genetic models. *QJM* 2000;93:7–14.
- 48 Farooqi IS, et al: Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. *N Engl J Med* 2003;348:1085–1095.
- 49 Branson R, et al: Binge eating as a major phenotype of melanocortin 4 receptor gene mutations. *N Engl J Med* 2003;348:1096–1103.

- 50 Bernhard F, et al: Functional relevance of genes implicated by obesity genome-wide association study signals for human adipocyte biology. *Diabetologia* 2013;56:311–322.
- 51 Heid IM, et al: Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat Genet* 2010;42:949–960.
- 52 Fox CS, et al: Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. *PLoS Genet* 2012;8:e1002695.
- 53 Huang C, Florez JC: Pharmacogenetics in type 2 diabetes: potential implications for clinical practice. *Genome Med* 2011;3:76.
- 54 Pearson ER: Pharmacogenetics in diabetes. *Curr Diab Rep* 2009;9:172–181.
- 55 Franks PW, Nettleton JA: Invited commentary: gene $\times$ lifestyle interactions and complex disease traits – inferring cause and effect from observational data, sine qua non. *Am J Epidemiol* 2010;172:992–997; discussion 998–999.
- 56 Manolio TA, Bailey-Wilson JE, Collins FS: Genes, environment and the value of prospective cohort studies. *Nat Rev Genet* 2006;7:812–820.
- 57 Franks PW, Shungin D: The interplay of lifestyle and genetic susceptibility in type 2 diabetes risk. *Diabetes Manag* 2011;1:299–307.
- 58 Franks PW, Poveda A: Gene-lifestyle and gene-pharmacotherapy interactions in obesity and its cardiovascular consequences. *Curr Vasc Pharmacol* 2011;9:401–456.
- 59 Andreasen CH, et al: Low physical activity accentuates the effect of the FTO rs9939609 polymorphism on body fat accumulation. *Diabetes* 2008;57:95–101.
- 60 Rampersaud E, et al: Physical activity and the association of common FTO gene variants with body mass index and obesity. *Arch Intern Med* 2008;168:1791–1797.
- 61 Franks PW, et al: Assessing gene-treatment interactions at the FTO and INSIG2 loci on obesity-related traits in the Diabetes Prevention Program. *Diabetologia* 2008;51:2214–2223.
- 62 Jonsson A, et al: Assessing the effect of interaction between an FTO variant (rs9939609) and physical activity on obesity in 15,925 Swedish and 2,511 Finnish adults. *Diabetologia* 2009;52:1334–1338.
- 63 Kilpeläinen TO, et al: Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS Med* 2011;8:e1001116.
- 64 Palla L, et al: Challenges in the use of literature-based meta-analysis to examine gene-environment interactions. *Am J Epidemiol* 2010;171:1225–1232.
- 65 Franks PW, Pearson ER, Florez JC: Gene-environment and gene-treatment interactions in type 2 diabetes: progress, pitfalls, and prospects. *Diabetes Care* 2013;36:1413–1421.
- 66 Li S, et al: Physical activity attenuates the genetic predisposition to obesity in 20,000 men and women from EPIC-Norfolk prospective population study. *PLoS Med* 2010;7:e1000332.
- 67 Qi Q, et al: Television watching, leisure time physical activity, and the genetic predisposition in relation to body mass index in women and men. *Circulation* 2012;126:1821–1827.
- 68 Qi Q, et al: Sugar-sweetened beverages and genetic risk of obesity. *N Engl J Med* 2012;367:1387–1396.
- 69 Williams PT: Attenuating effect of vigorous physical activity on the risk for inherited obesity: a study of 47,691 runners. *PLoS One* 2012;7:e31436.
- 70 Christakis NA, Fowler JH: The spread of obesity in a large social network over 32 years. *N Engl J Med* 2007;357:370–379.
- 71 Kilpeläinen TO, et al: Association of variants in the PCSK1 gene with obesity in the EPIC-Norfolk study. *Hum Mol Genet* 2009;18:3496–3501.
- 72 Schatzkin A, Slud E: Competing risks bias arising from an omitted risk factor. *Am J Epidemiol* 1989;129:850–856.
- 73 Andreasen CH, et al: Non-replication of genome-wide based associations between common variants in INSIG2 and PFKP and obesity in studies of 18,014 Danes. *PLoS One* 2008;3:e2872.
- 74 Lyon HN, et al: The association of a SNP upstream of INSIG2 with body mass index is reproduced in several but not all cohorts. *PLoS Genet* 2007;3:e61.
- 75 Heid IM, et al: Meta-analysis of the INSIG2 association with obesity including 74,345 individuals: does heterogeneity of estimates relate to study design? *PLoS Genet* 2009;5:e1000694.
- 76 Sonestedt E, et al: Association between fat intake, physical activity and mortality depending on genetic variation in FTO. *Int J Obes* 2011;35:1041–1049.
- 77 Sonestedt E, et al: Fat and carbohydrate intake modify the association between genetic variation in the FTO genotype and obesity. *Am J Clin Nutr* 2009;90:1418–1425.
- 78 Peter I, et al: Association of type 2 diabetes susceptibility loci with one-year weight loss in the look AHEAD clinical trial. *Obesity* 2012;20:1675–1682.
- 79 Qi Q, et al: Insulin receptor substrate 1 gene variation modifies insulin resistance response to weight-loss diets in a 2-year randomized trial: the Preventing Overweight Using Novel Dietary Strategies (POUNDS LOST) trial. *Circulation* 2011;124:563–571.
- 80 Kilpeläinen TO, et al: Interaction of single nucleotide polymorphisms in ADRB2, ADRB3, TNF, IL6, IGF1R, LIPC, LEPR, and GHRL with physical activity on the risk of type 2 diabetes mellitus and changes in characteristics of the metabolic syndrome: the Finnish Diabetes Prevention Study. *Metabolism* 2008;57:428–436.
- 81 Bea JW, et al: Lifestyle modifies the relationship between body composition and adrenergic receptor genetic polymorphisms, ADRB2, ADRB3 and ADRA2B: a secondary analysis of a randomized controlled trial of physical activity among postmenopausal women. *Behav Genet* 2010;40:649–659.
- 82 Corella D, et al: Obese subjects carrying the 11482G>A polymorphism at the perilipin locus are resistant to weight loss after dietary energy restriction. *J Clin Endocrinol Metab* 2005;90:5121–5126.
- 83 Franks PW, et al: The Pro12Ala variant at the peroxisome proliferator-activated receptor gamma gene and change in obesity-related traits in the Diabetes Prevention Program. *Diabetologia* 2007;50:2451–2460.
- 84 Jablonski KA, et al: Common variants in 40 genes assessed for diabetes incidence and response to metformin and lifestyle intervention in the diabetes prevention program. *Diabetes* 2010;59:2672–2681.
- 85 Moore AF, et al: Extension of type 2 diabetes genome-wide association scan results in the Diabetes Prevention Program. *Diabetes* 2008;57:2503–2510.
- 86 Franks PW: Gene $\times$ environment interactions in type 2 diabetes. *Curr Diab Rep* 2011;11:552–561.
- 87 Manning AK, et al: Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP $\times$ environment regression coefficients. *Genet Epidemiol* 2011;35:11–18.
- 88 Kraft P, et al: Exploiting gene-environment interaction to detect genetic associations. *Hum Hered* 2007;63:111–119.
- 89 Manning AK, et al: A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet* 2012;44:659–669.
- 90 Boks MP, et al: Investigating gene environment interaction in complex diseases: increasing power by selective sampling for environmental exposure. *Int J Epidemiol* 2007;36:1363–1369.
- 91 Pare G, et al: On the use of variance per genotype as a tool to identify quantitative trait interaction effects: a report from the Women's Genome Health Study. *PLoS Genet* 2010;6:e1000981.
- 92 Yang J, et al: FTO genotype is associated with phenotypic variability of body mass index. *Nature* 2012;490:267–272.
- 93 Ahmad S, et al: Gene $\times$ physical activity interactions in obesity: combined analysis of 111,421 individuals of European ancestry. *PLoS Genet* 2013;9:e1003607.