

Strategies and prospects for finding insulin resistance genes

Michael P. Stern

Division of Clinical Epidemiology, Department of Medicine, University of Texas Health Science Center, 7703 Floyd Curl Drive, San Antonio, Texas 78229-3900, USA.
Phone: (210) 567-4737; Fax: (210) 567-6955; E-mail: stern@uthscsa.edu.

In this paper I review the evidence that human sensitivity to insulin is influenced by heredity, in particular by one or more genes with relatively large effects. Even granting the existence of such genes, however, the best route toward identifying them — and the related genes that are presumed to affect an individual's risk of obesity or type 2 diabetes — remains unclear. Thus, the phenotype of insulin resistance can be assessed either directly, with techniques such as the euglycemic hyperinsulinemic clamp, or indirectly, using various proxy measures such as serum insulin concentrations. In addition, several different genotyping strategies might be employed. These alternatives raise both practical and theoretical questions: Which of the available phenotypic traits offers the most promising approach to identifying insulin resistance genes? Which methodologies are most powerful, and which best exploit the anticipated continued successes of the Human Genome Project? Finally, is it more realistic to search for genes influencing “intermediate” phenotypes such as insulin sensitivity, or to seek the related disease genes directly?

Do genes influence insulin resistance?

Many studies suggest that genetic factors influence insulin resistance, but nearly all of these have assessed family resemblance, which can result from shared environmental influences rather than the effects of genes per se. For example, persistent maternal smoking could cause siblings to resemble one another in having below average birth weights, and, since low birth weight is associated with insulin resistance in adult life (1), such resemblance could lead to sibling resemblance in insulin resistance as adults. Such a shared environmental influence would simulate the action of shared genes (1). Nevertheless, taken in aggregate, the studies of family resemblance have generally been interpreted as supporting a genetic basis for the observed resemblance in various indicators of insulin resistance.

The studies of family resemblance in insulin resistance have relied on a number of indicators of insulin resistance. These include insulinemia (either fasting or following oral glucose loading), as well as direct measures, such as the euglycemic hyperinsulinemic clamp technique or the frequently sampled intravenous glu-

cose tolerance test (FSIVGTT). In addition, several alternative study designs have been used to investigate this topic, including studying rare monogenic forms of insulin resistance, assessing insulin resistance in the relatives of diabetic individuals, comparing the differential concordance of insulin resistance between monozygotic and dizygotic twins, and studying nuclear families or extended pedigrees (reviewed in detail in ref. 2).

Many mutations in the *Insulin Receptor* gene are known that affect receptor function (3). Some of these mutations lead to impaired insulin receptor biosynthesis, while others affect posttranslational modification of the receptor molecule and thus reduce the transport of the receptor to the cell surface. Still others impair the binding of insulin to its receptor or the activation of the receptor tyrosine kinase, the first step in insulin signaling. Lastly, some mutations cause accelerated receptor degradation (3). These mutations typically behave in a recessive manner and can be associated with severe insulin resistance if they are present in the homozygous or, more frequently, in the compound heterozygous state. They have also been associated with a number of rare syndromes including type A insulin resistance (the triad of insulin resistance, acanthosis nigricans, and hyperandrogenicity), leprechaunism, and others. Although these mutations are of considerable scientific interest, they are unlikely to be responsible for more than a small fraction of the variation in insulin sensitivity seen in the general population.

Studies involving different ethnic groups — e.g., Caucasians, Mexican-Americans, and Chinese — have demonstrated insulin resistance in nondiabetic relatives of type 2 diabetic subjects. For example, Haffner et al. have shown that fasting insulinemia in Mexican-Americans increases in a stepwise fashion in nondiabetic offspring having zero, one, or two type 2 diabetic parents (4). Since type 2 diabetic subjects are almost invariably insulin-resistant, these findings suggest that insulin resistance occurs more commonly in the offspring of insulin resistant parents. Gulli et al. (5) reported similar findings using the euglycemic hyperinsulinemic clamp technique, a direct measure of insulin resistance, and they confirmed that glucose disposal was impaired in the offspring of two type 2 diabetic parents, compared with matched controls. This

impairment was found to be due primarily to a defect in nonoxidative glucose disposal, rather than to a defect in glucose oxidation. First- and second-phase insulin secretion were also increased in the offspring, suggesting compensatory hypersecretion of insulin in response to insulin resistance in these individuals. Major reductions in nonoxidative glucose disposal with only modest reductions in glucose oxidation have also been observed in Caucasian nondiabetic first-degree relatives of diabetic subjects (6).

Family clustering of insulin resistance measurements offers another method of demonstrating the effects of genes on these phenotypes. Such clustering has been demonstrated in Pima Indians, using the euglycemic hyperinsulinemic clamp technique (7), and in Caucasians, using the FSIVGTT (8). The range of insulin resistance values found within families was considerably narrower than the range of values across families, indicating that family members are more likely to resemble one another in their degree of insulin resistance than they are to resemble unrelated individuals.

Finally, a number of studies have estimated the heritability of insulin resistance. Heritability is defined as the proportion of variance in a trait attributable to the additive effect of genes. Heritability of insulin resistance as estimated from twin studies ranges from 47 to 66% (2). Heritability of fasting insulin and insulin concentration 2 hours after an oral glucose load have also been estimated in a study of extended pedigrees and have been found to be 35% and 13%, respectively (9).

Is insulin resistance subject to major gene effects?

Given that genetic factors influence insulin action, the question arises whether these influences are exerted by a few “major genes,” each with a relatively large effect, or by a large number of “polygenes,” each with a relatively minor effect. Presumably, in the former case, the genes in question would both be easier to identify and have greater relevance to public health. Complex segregation analyses can be used to study the mode of inheritance of a trait and, in particular, to infer whether the distribution of the trait within pedigrees is compatible with the action of a major gene. The nominal major gene effect can be simultaneously adjusted for covariates such as age, sex, and polygenic background. Using this type of analysis, Schumacher et al. (10) presented evidence suggesting that a single major gene accounted for 33% of the total variance in fasting insulin concentration and 48% of the variance in insulin concentration 1 hour after an oral glucose load, respectively, in Caucasian pedigrees. Polygenes accounted for an additional 11% and 4% of the variance in the former and latter traits, respectively. Similar results were reported for Mexican-Americans by Mitchell et al. (11), who found a major gene accounting for 31% of the population variance in insulin concentration 2 hours after an oral glucose load.

To date, no segregation analyses have been performed using direct measures of insulin action. In Pima Indians, however, a trimodal distribution has been reported for insulin sensitivity, as measured by the euglycemic hyperinsulinemic clamp (12). Such a distribution could have been produced by a single major gene having a codominant effect. Despite all of these reports, the evidence is not yet conclusive that variation in insulin action in the general population is influenced by major genes. Final proof will probably have to await the actual identification of mutations in the relevant genes.

Defining the phenotypes of insulin resistance

Because the statistical power of a genetic study depends in part on the number of relative pairs analyzed, surrogate measures of insulin resistance offer one clear advantage over direct measures. The former are easier, less invasive, and cheaper to employ and can therefore be applied more readily to large numbers of subjects. On the other hand, the correlations between these surrogates and definitive measures of insulin action (such as are obtained from the euglycemic hyperinsulinemic clamp or FSIVGTT) are less than perfect. For example, in subjects with normal glucose tolerance, fasting insulin, the simplest surrogate, is only moderately correlated with insulin resistance as measured by the euglycemic clamp ($r^2 = 0.45$; see ref. 13). Howard et al. compared a number of surrogates, such as insulinemia, the homeostasis model (HOMA), the insulin-to-glucose ratio, and others, with direct measures of insulin sensitivity made by the FSIVGTT (14). Except for the insulin-to-glucose ratio, which was poorly correlated with insulin sensitivity, the correlations for the surrogates ranged from moderate to good (r^2 ranging from 0.4 to 0.9; see ref. 14), although they varied somewhat by ethnic group.

It remains to be seen whether the enhanced statistical power obtained by using surrogate measures that can be more readily applied to relatively large samples outweighs the loss of information resulting from their less than perfect correlation with more definitive measures of insulin resistance. Three ongoing family studies are combining whole genome scans with definitive measures of insulin resistance (FSIVGTT): the Finland-United States Investigation of NIDDM Genetics (FUSION) Study (15); the Insulin Resistance Atherosclerosis (IRAS) Family Study (14); and the Genetics of Non-Insulin Dependent Diabetes Mellitus (GENNID) Study (16).

It is also possible that insulin resistance genes will be identified by studying phenotypes associated with the insulin resistance syndrome (IRS), such as dyslipidemia, obesity (especially abdominal obesity), hypertension, and glucose intolerance. A number of studies have reported linkage to one or more of these phenotypes. Factor analysis and related techniques can be used to define combined traits that can themselves be tested for heritability and examined for genetic linkage.

Several studies (17–19) have used this analytic approach, and all have identified three independent factors underlying the insulin resistance syndrome: one reflecting predominantly glucose and insulin concentrations, the second reflecting predominantly a lipid component, and the third reflecting predominantly blood pressure. The contribution of obesity to these “latent phenotypes” is variable, since it associates with different factors in each of the three studies. Edwards et al. reported that these factors are at least moderately heritable in female Caucasian twins (20), and we have reported heritabilities ranging from 45 to 57% for the three factors in Mexican-American participants in the San Antonio Family Diabetes Study (19).

Genotyping strategies

Beyond the questions of whether to employ direct or indirect measures of insulin resistance and whether to model simple or composite phenotypes, there is the matter of genotyping strategies. The study of candidate genes has proven to be a popular approach. A candidate gene is usually defined as a gene whose protein product, based on its biological activity, can plausibly be assumed to influence the trait under consideration. The main problem with this approach is that the number of such proteins is legion. In the case of insulin resistance the list would presumably include proteins that influence insulin binding to its receptor, those involved in the insulin signaling pathway, those influencing glucose uptake and cell metabolism, and no doubt countless others, both known and unknown. Variation in any one of these proteins could potentially affect insulin sensitivity. Moreover, as more is learned about these pathways, the list of potential candidate genes, already dauntingly large, grows seemingly without limit. The chances of a “lucky hit” would seem to be remote, and indeed, studies with candidate genes have thus far proved to be disappointing.

Before investing major resources in studying a potential candidate gene, it would seem that a stronger prior hypothesis implicating the proposed candidate is needed, rather than just a general sense that, based on its biology, it could influence insulin action. Efforts to implicate candidate genes often take the form of association studies, which can be performed in either related or unrelated individuals. Population stratification — that is, the existence of more than one ancestral source of a population’s gene pool — represents a major limitation of this study design. If, as is likely, the various ancestral sources differ both in their susceptibility to various diseases and in the frequency of various genetic markers, spurious associations may be observed between genetic markers and various phenotypes. In some cases, the existence of population stratification is well appreciated, as in the case of Mexican-Americans, whose gene pool derives from both European and Native American sources. In other cases, the basis of

population stratification is more obscure. This problem can be ameliorated to some extent by selecting appropriate controls, especially family-based controls, which help assure that whatever population stratification affects the cases also affects the controls. Still, it is likely that population stratification is at least partly responsible for the large number of reports of association between candidate gene polymorphisms and diseases or conditions that have later proven to be nonreplicable.

With the increasing availability of large numbers of highly informative genetic markers, particularly microsatellite markers that now span the entire genome, the strategy of whole-genome scanning for linkage to phenotypes of interest has become feasible. Controversy still exists over what lod scores should be taken as evidence of linkage, but, based on simulation studies, widely accepted criteria for “suggestive” and “significant” linkage are lod scores above 2.0 and above 3.0, respectively (21). A candidate gene located in a chromosomal region that has been found to be linked to a phenotype of interest is referred to as a “positional” candidate. Sometimes a gene whose protein product might not otherwise strongly recommend itself as a candidate by virtue of its biological activity may suddenly be considered a plausible candidate by virtue of its presence in a linked region. One should be cautious, however, since linkages detected in whole-genome scans with microsatellite markers typically implicate relatively large chromosomal regions extending over perhaps 20–30 centimorgans (approximately 20–30 megabases). Such regions typically harbor hundreds of genes.

If a genetic variant in a positional candidate is later found to be associated with a phenotype, the prior evidence of linkage provides at least some degree of reassurance that the association is not an artifact of population stratification. For example, the gene for membrane glycoprotein PC-1 is considered to be a candidate for insulin resistance, since this protein has been shown to inhibit insulin receptor tyrosine kinase activity in cultured fibroblasts (22). We have recently reported genetic linkage between the chromosomal region where this gene is located and fasting serum insulin concentration (lod = 3.9) in nondiabetic San Antonio Family Diabetes Study participants (23). Clearly, this evidence of linkage enhances this gene’s status as a candidate gene. On the other hand, given the relatively broad chromosomal region implicated in the linkage analysis, it is important not to become overly enamored of this or any other positional candidate merely because, of all the genes in the linked region, its function happens to be known.

Just as the presence of detectable linkage does not anoint a positional candidate, neither does the absence of linkage exclude a candidate’s possible significance. Even if a given candidate gene does not play a major role in the overall distribution of insulin sensitivity, it may still be responsible for minor variations in the

phenotype in the general population, or it may exert major effects on a small number of individuals in the population. Thus, mutations in the *Insulin Receptor* gene clearly produce insulin resistance in various relatively rare conditions (3), but linkage studies indicate that variation in this gene accounts for, at most, a minor proportion of the variation in insulin sensitivity in the general population.

Future challenges in the search for insulin resistance genes

The recent publication of a first draft of the human genomic sequence will clearly facilitate the search for insulin resistance genes, but the obstacles remain formidable. First, it is not obvious from simple inspection which stretches of DNA sequence represent genes. In addition, the identification of positional candidates still depends on uncovering evidence of linkage between a chromosomal region and some phenotype related to insulin resistance. Finally, it is not sufficient merely to show an association between a variant in a candidate gene and a relevant phenotype, since such an association may merely reflect linkage disequilibrium. The variant must also be shown to have altered function.

Once a region has been found to be linked to insulin resistance, it will be necessary to find a gene in the linked region, variants of which are causally associated with variation in insulin sensitivity. In many cases, it will be necessary to narrow the region of linkage to a more manageable size for DNA sequencing or other mutation detection techniques. One approach to accomplish this is to perform disequilibrium mapping using single-nucleotide polymorphisms (SNPs). SNP discovery is proceeding at a rapid rate, and many (but by no means all) of the newly discovered SNPs are being made publicly available. However, it may still prove necessary to develop new SNPs that are adequately informative in the population under study. After finding promising SNPs that are in strong linkage disequilibrium with the phenotype of interest comes the challenge of identifying true functional sequence variants — those that directly influence insulin sensitivity. Even a favorable variant, such as a missense mutation, must still be shown to confer altered biological function in animal models or in cell systems. For putative mutations in regulatory regions, such a demonstration will be all the more crucial.

In view of these difficulties, it is clear that gene discovery requires the commitment of substantial resources. It is conceivable, therefore, that society will choose to invest in the search for genes that influence the risk of type 2 diabetes or obesity but balk at investing similar resources into finding genes that affect intermediate phenotypes, such as insulin resistance. Although it is often held that such genes will be easier to find, because the causal chain between them and the phenotype is shorter, it should be acknowledged that

at present there are no examples of this technique having successfully led to the discovery of a bona fide disease gene. Large pharmaceutical companies may be willing to take on the risk of identifying genes for intermediate phenotypes, since such genes may point to novel drug targets. Nevertheless, it seems possible that genes for intermediate phenotypes will be identified, not as stepping stones that lead to diabetes or obesity genes, but as byproducts of a direct search for these more medically relevant genes.

Conclusions

There is considerable evidence that variation in insulin action in the general population is influenced by genetic variation and that at least some insulin resistance genes have relatively large effects. The most promising strategy for identifying these genes involves examining positional candidates, i.e., genes located in chromosomal regions for which there is evidence of linkage to various insulin resistance phenotypes. When a biological function related to insulin action can be ascribed to such a gene, its claim to candidate status is clearly enhanced. However, the genome scan approach is more likely to identify anonymous genes, the biological action of which will need to be established to reinforce their candidate status. In either case, it will be necessary to prove in animal models or cell systems that any variants in the candidate gene actually alter function with respect to insulin resistance. In view of the difficulties in gene identification and the major resources that will need to be committed, it is possible that insulin resistance genes will be found as byproducts of the search for genes that have more direct public health relevance.

Acknowledgments

I would like to thank Robin Leach (University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA) and Peter O'Connell (Baylor College of Medicine, Houston, Texas, USA) for reading this manuscript and making helpful suggestions.

1. Stern, M.P., Bartley, M., Duggirala, R., and Bradshaw, B. 2000. Birth weight and the metabolic syndrome: thrifty phenotype or thrifty genotype? *Diabetes Metab. Res. Rev.* **16**:88–93.
2. Stern, M.P., and Mitchell, B.D. 1999. Genetics of insulin resistance. In *Contemporary endocrinology. Insulin resistance: the metabolic syndrome X*. G.M. Reaven and A. Laws, editors. Humana Press Inc. Totowa, New Jersey, USA. 3–18.
3. Taylor, S.I. 1992. Lilly Lecture: molecular mechanisms of insulin resistance. Lessons from patients with mutations in the insulin-receptor gene. *Diabetes.* **41**:1473–1490.
4. Haffner, S.M., Stern, M.P., Hazuda, H.P., Mitchell, B.D., and Patterson, J.K. 1988. Increased insulin concentrations in non-diabetic offspring of diabetic parents. *N. Engl. J. Med.* **319**:1297–1301.
5. Gulli, G., Ferrannini, E., Stern, M., Haffner, S., and DeFronzo, R.A. 1992. The metabolic profile of NIDDM is fully established in glucose-tolerant offspring of two Mexican-American NIDDM parents. *Diabetes.* **41**:1575–1586.
6. Eriksson, J., et al. 1989. Early metabolic defects in persons at increased risk for non-insulin-dependent diabetes mellitus. *N. Engl. J. Med.* **321**:337–343.
7. Lillioja, S., et al. 1987. In vivo insulin action is familial characteristic in nondiabetic Pima Indians. *Diabetes.* **36**:1329–1335.

8. Martin, B.C., et al. 1992. Family clustering of insulin sensitivity. *Diabetes*. **41**:850–854.
9. Mitchell, B.D., et al. 1996. Genetic and environmental contributions to cardiovascular risk factors in Mexican Americans: the San Antonio Family Heart Study. *Circulation*. **94**:2159–2170.
10. Schumacher, M.C., Hasstedt, S.J., Hunt, S.C., Williams, R.R., and Elbein, S.C. 1992. Major gene effect for insulin levels in familial NIDDM pedigrees. *Diabetes*. **41**:416–423.
11. Mitchell, B.D., et al. 1995. Evidence for a major gene affecting post-challenge insulin levels in Mexican-Americans. *Diabetes*. **44**:284–289.
12. Bogardus, C., et al. 1989. Distribution of in vivo insulin action in Pima Indians as mixture of three normal distributions. *Diabetes*. **38**:1423–1432.
13. Laakso, M. 1993. How good a marker is insulin level for insulin resistance? *Am. J. Epidemiol.* **137**:959–965.
14. Howard, G., et al. 1998. Ability of alternative indices of insulin sensitivity to predict cardiovascular risk: comparison with the “minimal model”. *Ann. Epidemiol.* **8**:358–369.
15. Watanabe, R.M., et al. 1999. Familiarity of quantitative metabolic traits in Finnish families with non-insulin-dependent diabetes mellitus. Finland-United States Investigation of NIDDM Genetics (FUSION) Study. *Hum. Hered.* **49**:159–168.
16. Raffel, L.J., et al. 1996. The GENNID study. A resource for mapping the genes that cause NIDDM. *Diabetes Care*. **19**:864–872.
17. Edwards, K.L., et al. 1994. Multivariate analysis of the insulin resistance syndrome in women. *Arterioscler. Thromb.* **14**:1940–1945.
18. Meigs, J.B., et al. 1997. Risk variable clustering in the insulin resistance syndrome. The Framingham Offspring Study. *Diabetes*. **46**:1594–1600.
19. Duggirala, R., Williams, K., Arya, R., Blangero, J., and Stern, M.P. 2000. Genetics of the factors underlying the Insulin Resistance Syndrome in nondiabetic Mexican Americans. *American Journal of Human Biology*. **12**:274. (Abstr.)
20. Edwards, K.L., et al. 1997. Heritability of factors of the insulin resistance syndrome in women twins. *Genet. Epidemiol.* **14**:241–253.
21. Lander, E., and Kruglyak, L. 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11**:241–247.
22. Goldfine, I.D., et al. 1998. Membrane glycoprotein PC-1 and insulin resistance. *Mol. Cell. Biochem.* **182**:177–184.
23. Duggirala, R., Blangero, J., Almasy, L., O’Connell, P., and Stern, M.P. 1999. Insulin resistance-related phenotypes are linked to genetic locations on chromosomes 4 and 6 in non-diabetic Mexican Americans. *Diabetes*. **48**(Suppl. 1):A47.