Racialized Genetics and the Study of Complex Diseases

the thrifty genotype revisited

Yin C. Paradies,* Michael J. Montoya,† and Stephanie M. Fullerton‡

ABSTRACT Current debate on the use of population genetic data for complex disease research is driven by the laudable goals of disease prevention and harm reduction for all, especially dispossessed, formerly enslaved, or colonized populations. This article examines one of the oldest gene-based theories of complex disease causation: the thrifty genotype hypothesis (THG). This hypothesis is emblematic of the way in which genetic research into complex disease attracts a high investment of scientific resources while contributing little to our capacity to understand these diseases and perpetuating problematic conceptions of human variation. Although there are compelling
reasons to regard the high prevalence of type 2 diabetes mellitus as a by-product of our biological incapacity to cope with modern affluent and sedentary lifestyles, there is at present no consistent evidence to suggest that minority populations are especially genetically susceptible. Nor is it clear why such genetic differences would be expected, given the original pan-species orientation of the TGH. The limitations inherent in current applications of the TGH demonstrate that genetic research into complex disease demands careful attention to key environmental, social, and genetic risk factors operating within and between groups, not the simplistic attribution of between-group differences to racialized genetics. A robust interdisciplinary approach to genetic epidemiological research is proposed.

This millennium has seen a resurgence of debate on the use of ethnorace in genetics research (e.g., Bamshad 2005; Berg et al. 2005; Cooper, Kaufman, and Ward 2003; Risch et al. 2002). We contend that the complex topic of genetics and its relation to ethnoracial identity is best elucidated through specific case studies in association with more abstract debate that incorporates interdisciplinary perspectives (e.g., Fee 2006; Shields et al. 2005). This article examines the continued misuse of ethnorace in human genetics, and the implications of this use for public health, through a discussion of the genetic, epidemiological, and sociocultural aspects of the thrifty genotype hypothesis (TGH). The TGH, which posits an ancient genetic adaptation to cycles of feast and famine, has been employed repeatedly in research seeking to explain why type 2 diabetes mellitus (T2DM) disproportionately afflicts a range of disadvantaged ethnoracial groups. We begin by outlining an interdisciplinary position on human variation in order to situate the critique of the TGH that follows. An overview of the epidemiology of T2DM is then followed by a detailed description of the TGH, a critical review of its central premises, and the epidemiological and genetic evidence that has been collected in its wake. We conclude by examining the prospects for rehabilitating the genetic study of ethnoracial health disparities in light of the concerns highlighted.

**Ethnorace, Racialization, and Health**

Human beings are remarkably genetically similar, sharing 99.9% of their genomes in common (Li and Sadler 1991). Of this relatively small intra-species diversity, approximately 85–95% occurs within, rather than between, groups of individuals classified as belonging to different ethnoraces (Kittles and Weiss 2003). The vast majority of the observed genetic variation (whether within or between groups) is selectively neutral and has no effect on physiology. A currently unknown, but even smaller, proportion of between-group differences is directly relevant to disease susceptibility (Bamshad 2005). Single gene variants

---

1The term *ethnorace* is used to emphasize the ineluctably intertwined concepts of race and ethnicity as utilized in medico-scientific practices (Montoya 2003).
with clear health consequences, such as the point mutation that leads, in homo-
ygous dose, to sickle-cell anemia, are known to vary in frequency across eth-
noracial populations (Kittles and Weiss 2003), but this class of genetic diseases ac-
count for a small percentage of ethnoracial health disparities. Thus, only a very
small proportion of total genetic variation is even potentially relevant to disease
prevalence, or other health-related differences, between groups of individuals
classified by ethnorace.

Despite these observations, ethnoracial concepts have a currency and rele-
vance in everyday discourse that is independent of any definable genetic dis-
tinction (Ossorio and Duster 2005). In contemporary society, as in the past, the
attribution of ethnorace to individuals often depends on the presence of partic-
ular morphological characteristics, such as skin color, as well as known or
assumed lines of descent from ancestors with particular morphological charac-
teristics or defined geographic ancestry (Berg et al. 2005). The continuously
varying character of such features in human populations, combined with the his-
torically and culturally contingent nature of attaching social meaning to specific
morphologies, explains the marked variation in the number, nature, and type of
ethnoraces identified over both time and place (Nobles 2000).

Although ethnorace is essentially a social system of classification laid atop
complex geographic patterns of population genetic variation, it is now possible
to assign individuals to likely ethnoracial (or, sometimes, “ancestral”) back-
grounds via a cumulative examination of differences in the frequencies of com-
mon genetic variants as well as more geographically restricted (population-spe-
cific) polymorphisms (Rosenberg et al. 2002). Importantly, these efforts do not
demonstrate that ethnoraces are naturally aggregating genetic clusters. Instead,
they underline that individuals sampled from geographically disparate regions
will often differ to a small degree genetically, and in a manner predicted by stan-
dard isolation-by-distance models (Weiss and Fullerton 2005). These efforts also
do not say anything about the ethnoracial distribution of variants relevant to dis-
case risk (although it is often assumed that if clustering is achieved on the basis
of a neutral set of genetic markers, disease-linked variation will be similarly geo-
graphically distributed).

These well-established pitfalls of conflating ethnorace and genetics have been
overlooked or misunderstood in many quarters, and such categories continue to
be utilized by scientists, government, media, and lay persons under the misplaced
assumption that individuals with different ethnoracial identities are genetically
distinct and that this has relevance to ethnoracial health inequalities (Braun
2006; Stevens 2003). A telling example of such misconceptions is the ranking by
the U.S. National Library of Medicine of genetic research as the top investiga-
tive priority for health disparities (including ethnoracial disparities), ahead of re-
search into the environmental and social determinants of ethnoracial inequalities
in health (Sankar et al. 2004).

The two key misapprehensions that drive the continued overemphasis on
genetics in explaining ethnoracial health disparities are (1) that genetic variation explains much or all of the health differentials among individuals and populations, and (2) that there is more genetic variation with relevance to disease between, rather than within, ethnoracial population groups. While the former assumption is a variety of geneticization (the notion that genes are the primary determinants of biology and behavior), the latter represents a form of racialization, which, in this case, entails the exaggerated salience of ethnorace in genetic investigation (Goodman 2000).

Geneticization is the process by which priority is given to differences between individuals based on their genetic distinctiveness (Lippman 1993). Geneticization occurs conceptually, when genetic terminology is used to define problems; institutionally, when genetic expertise is enlisted to deal with specific problems; culturally, when genetic knowledge and technology lead to changes in individual and social attitudes toward, for instance, the prevention and control of disease; and philosophically, when genetic imagery produces particular views on human identity, interpersonal relationships, and individual responsibility (ten Have 2001).

Extreme forms of geneticization become a kind of “genetic fatalism” in which genes and genetic variants are considered to be both completely determining and unalterable (Alper and Beckwith 1995). Geneticization appeals to dominant ideas of individuality and related notions of personal responsibility for health. Through geneticization, the public health gaze is misfocused onto individual biophysical risk factors and away from social, environmental, and ecological factors (Ellison and Jones 2002). Despite the complexity involved in attributing causality to specific genetic determinants, genes are routinely accorded exaggerated etiological significance with scientists as well as the media, overstating the certainty and implications of genetic research (Sankar et al. 2004; Shields et al. 2005).

The incredible popularity of genetic knowledge would be more comprehensible if its allure were based upon evidence of sustained and incontrovertible public health benefit. Rather, to date, genetics has been sold largely on its incredible potential for benefit, with genetic explanations being customarily accepted in the absence of therapies based upon genetic insights (Conrad and Gabe 1999; Stevens 2003). Despite many millions of dollars and decades spent searching for genes that contribute to common complex diseases, few convincingly replicable findings have been reported, with the preponderance of null findings explained largely as a function of inadequate methodology or sample size (Buchanan, Weiss, and Fullerton 2006; Cruickshank et al. 2001; Hirshchhorn et al. 2002). This reliance upon the genetic paradigm appears to go hand in hand with the use of ethnorace as a marker of genetic difference.

Racialization is the process of ascribing somatic essentialized innate difference based on phenotype, ancestry, or culture to ethnoracial groups (Miles and Brown 2003). Although ethnoracial labels have been shown time and again to be historically and situationally determined, unreliable, and imprecise proxies for hu-
man genetic differences, ethnoracial inequalities in morbidity and mortality continue to be explained or hypothesized as a condition of innate and fixed genetic differences, with the existence of sociopolitical inequalities between human groups effectively denied.

Although most health-related genetic research has been conducted on populations of European ancestry, it is customarily indigenous, immigrant, minority, and other non-European groups who are considered to be genetically susceptible to a variety of diseases (Cooper, Kaufman, and Ward 2003). For such ethnoracial groups, genetic explanations for ill health are proposed when no such data are available, on the basis of nonsignificant findings, or even when genetic variation is unrelated to the disease under study. The association between unsubstantiated genetic attributions and the social status of an ethnoracial group demonstrates the potential harm that may result from the continuing (and increasing) use of ethnoracial categories in genetic research (Ellison and Jones 2002). Below we examine the intersection of racialization and geneticization in the context of research on T2DM through the lens of the TGH, beginning with an overview of T2DM itself.

**The Epidemiology of Type 2 Diabetes Mellitus**

Currently there are several different recognized types of diabetes, including type 1, type 2, gestational diabetes, and maturity onset diabetes of the young. Type 1 diabetes is characterized by absolute insulin deficiency, while other types of diabetes are characterized by insulin resistance or abnormal insulin secretion. By far the most common type of diabetes is T2DM, which accounts for over 90% of cases worldwide (Zimmet, Alberti, and Shaw 2001). T2DM is typically experienced as a late-onset chronic disease (although a younger age of onset is becoming increasingly common), and unlike the autoimmune-related type 1 diabetes, it is associated with risk factors such as increased obesity, dietary fat intake, smoking, and low physical activity (Quinn 2003). Additional risk factors such as low socioeconomic status, stress, and racism have also been associated with T2DM (Everson et al. 2002; Rosmond 2003; Tull and Chambers 2001). In addition, low birth weight and exposure to maternal diabetes in utero have been associated with T2DM in later life (Carter, Pugh, and Monterrosa 1996; Hales and Barker 2001).

Several different sources of evidence also suggest that genetic factors may play a role in T2DM susceptibility. Studies of siblings have found greater concordance for T2DM status among identical (monozygotic) as compared to non-identical (dizygotic) twins (Barroso 2005). Evidence of familial aggregation, particularly work demonstrating an increase in risk with an increase in the number of affected parents, also supports the importance of genetic susceptibility (Meigs, Cupples, and Wilson 2000). While a handful of segregation studies have hinted at the influence of one or more loci of major effect, the analysis of patterns of
Inheritance in families has been largely inconclusive, suggesting a more complex, polygenic mode of inheritance instead (McIntyre and Walker 2002; Schumacher et al. 1992).

The prevalence of T2DM in adults aged 20 years and over is on the increase worldwide, with a projected rise from 4% in 1995 to 5.4% in 2025. Over this time period the prevalence is expected to increase from 6% to 7.6% in developed nations, and from 3.3% to 4.9% in developing nations (King, Aubert, and Herman 1998). It is unlikely that such large changes in prevalence over the course of a single generation are related solely to variation in genetic susceptibility. Indeed, it has become clear that both obesity and T2DM are a consistent response to the voluntary or involuntary adoption of a high-consumption lifestyle. This is particularly clear from transnational comparative studies, and from studies that document a sharp increase in the incidence of T2DM following urbanization as well as migration from countries with low prevalence to those with a high prevalence of T2DM (Aschner 2002; Ravussin et al. 1994; Schulz et al. 2006; Taufa, and Benjamin 2004; van Dam 2003).

Since Kelly West’s (1974) seminal paper, it has been known that T2DM disproportionately affects fourth-world indigenous populations around the world. In 1997 the crude prevalence of diabetes (all types) for First Nation Canadian men and women was 3.6 and 5.3 times higher, respectively, than diabetes prevalence in the general Canadian population (Young et al. 2000). Similarly, the age-standardized prevalence of self-reported diabetes among Indigenous Australians in 2004–2005 was three times that of the non-indigenous population (Australian Bureau of Statistics 2006). In the United States, the prevalence of T2DM for non-Hispanic blacks, Hispanic/Latino Americans, and American Indians/Alaskan Natives (who receive care from the Indian Health Service) aged 20 years or older was 1.6 times, 1.5 times, and 2.2 times the rate for non-Hispanic whites in 2002, respectively (National Center for Chronic Disease Prevention 2006). Likewise, in 2000 the prevalence of diabetes for Maori and Pacific Islanders in New Zealand was nearly three times that of white New Zealanders (New Zealand Ministry of Health 2000).

Although Inuit populations have historically had one of the lowest rates of diabetes in the world, this is now changing, with the overall age-adjusted prevalence among Alaskan Inuits increasing from 8.8 per 1,000 people in 1985 to 12.1 per 1,000 people in 1993, with a prevalence of 6.6% reported among three Inuit groups in 1994 (Ebbesson et al. 1998). In the Pacific, T2DM has gone from being virtually nonexistent to a major health problem in recent decades (Foliaki and Pearce 2003).

Overall, the epidemiology of T2DM suggests a complex etiological picture,

---

2“Fourth-world indigenous peoples” are defined as minority indigenous populations living in first-world nations wherein institutionalized power is held by a colonizing, subordinating majority (O’Neil 1986).
with social and environmental risk factors being clearly implicated, while twin studies and related analyses simultaneously support a plausible genetic component. While genetic predispositions may, in part, explain inter-individual variation in T2DM, the racialized form of the TGH proposes that genetic predispositions also explain prevalence differences between populations—a contention that will be considered below.

**The Thrifty Genotype Hypothesis**

The thrifty genotype was proposed by James Neel in 1962 as an explanation for the increased prevalence of obesity, and by extension diabetes, in contemporary society. Neel argued that during 99% of human existence we existed as hunter-gatherers who experienced frequent cycles of alternating feast and famine. As a result we developed a genotype exceptionally efficient in the absorption, storage, or utilization of nutrients, which has now become maladaptive in a first-world context of sustained energy surplus. The initial form of the TGH was not racialized but was, instead, an attempt to explain the sudden rise in diabetes throughout the modern world.

One of the earliest references to the TGH was in an article by Johnson and McNutt (1964), which suggested that a high prevalence of T2DM among the Alabama-Coushatta Indians was caused by a thrifty genotype thought to be associated with T2DM in all humans, but which was much more common among Alabama-Coushatta Indians. Another early study suggested that the high prevalence of T2DM among Seneca Indians from western New York was “compatible with the presence of an inherited susceptibility to diabetes in this ethnic group” (Doeblin, Evans, and Ingall 1969, p. 625). In 1976, the TGH was proposed as an explanation for hyperglycemia that was more common in “full blood obese compared with part blood obese Aborigines” in Australia (Wise et al. 1976, p. 194). This paper noted, however, that the environment was still likely to be the “dominant . . . causative factor” in T2DM prevalence (p. 195).

Zimmet (1979) utilized the TGH to suggest that “Micronesians and Polynesians have a hereditary susceptibility to diabetes (i.e. diabetic genotype)” and that “The genotype for [T2DM] may render an individual susceptible to such factors as obesity, dietary abuse, stress, certain viruses or other as yet undefined environmental factors” (p. 147). In a study of young Aboriginal men in northeastern Arnhem Land, O’Dea, Spargo, and Nestel (1982) argued that insulin resistance in Aboriginal Australians, despite their extreme leaness, regular physical activity, and traditional diet, may indicate a susceptibility to T2DM if they were to westernize further.

Neel’s second article on the TGH (1982) explicitly discussed the high prevalence of diabetes in a range of ethnoracial groups, suggesting that this could be seen “as a ‘telescoped’ example of the type of genotype-environmental interaction which populations of Western European extraction had sustained earlier.” In
this article, Neel considered diabetes as a genetic disease, suggesting that the heterogeneity of diabetes requires "not one but several different thrifty genotypes" (p. 285). Although Neel discussed ethnorace as a factor in the TGH, the hypothesis did not involve racialized genetics, as no particular susceptibility was proposed for one ethnoracial group over another but only differences in the timing of environmental changes.

Neel suggested that there could be as many types of thrifty genes as types of T2DM, but that if there turned out to be too many (say 100 or more), then this could be explained by simple mutation pressure and "the need for the thrifty genotype would disappear" (p. 290). In a mix of cautious science and flamboyant intellectualism he concluded: "All these speculations may be utterly demolished the moment the precise etiologies of [T2DM] become known. Until that time, however, devising fanciful hypotheses based on evolutionary principles offers an intellectual sweepstakes in which I invite you all to join" (p. 290).

Over the coming years, many did indeed join this "intellectual sweepstakes." For example, McGarvey (1994) suggested that, for Polynesians, evolutionary pressure to produce a thrifty genotype occurred through the hardship of interisland colonization in the Pacific, likely difficulties in establishing agricultural crops upon arrival, and the destruction of such crops after cyclones. Wendorf (1992) hypothesized that the thrifty genotype was selected for during the initial colonization of North America, Australia, and some Pacific Islands due to the common experience of adapting to unfamiliar biotic systems, relying on foods that became extinct or were severely depleted, and frequent food shortages that continued for several generations. Wendorf suggested that this putative genotype did not occur in Inuits or certain South American tribes because these groups utilized a range of nutritional sources and did not suffer severe environmental changes or extinction of food sources.

These speculations notwithstanding, in his last publication on the TGH, Neel (1999) wrote that he found "no support to the notion that high frequency of [T2DM] in reservation Amerindians might be due simply to an ethnic predisposition—rather, it must predominantly reflect lifestyle changes" (p. S3). He went on to discuss how the original TGH "presented an overly simplistic view of the physiological adjustments involved in the transition from the lifestyle of our ancestors to life in the high-tech fast lane" (p. S4) and highlighted the role of phenotypic (developmental) rather than genotypic changes implicated in the much higher contemporary prevalence of diabetes.

Despite Neel himself refuting many aspects of the TGH, however, the racialized incarnation of this hypothesis continues to outlive its progenitor as it continues to be reiterated and researched in relation to Indigenous Australians, Native Americans, and First Nation Canadians (Busfield et al. 2002; Hegele et al. 2000; Wolford et al. 2001). In addition to indigenous peoples, the racialized TGH

---

2James V. Neel died on February 1, 2000, at the age of 84.
has also been applied to populations as diverse as Mexican Americans, African Americans, and Asians (Fujimoto 1992; Horikawa et al. 2000; Osei 1999).

The racialized TGH also retains currency in popular scientific circles, as well as in fields as diverse as epidemiology, anthropology, psychology, and genetics (Allen and Cheer 1996; Bindon and Baker 1997; Cacioppo et al. 2000; Carulli et al. 2005; Diamond 2003; R. C. Williams et al. 2000). In addition, the TGH continues to be examined in a range of ongoing research (Busfield et al. 2002; Clark et al. 2005; Poulsen et al. 2003; Ruiz-Narvaez 2005; Vander Molen et al. 2005). One of two key research themes in a proposed Centre for Aboriginal Genetic Epidemiology in Australia, for example, will be the genetic epidemiology of T2DM among Indigenous Australians. The founding director of this center cited the interplay between a genetic propensity for T2DM among Indigenous people due to their history as hunter-gathers and modern lifestyle changes (i.e., the racialized TGH) as the reason for such a focus (Blackwell 2005). This persistent recourse to a racialized TGH is perplexing, not only in the light of Neel’s subsequent disavowal, but also because of the many conceptual limitations of the hypothesis.

**Limitations of the Thrifty Genotype Hypothesis**

Modern humans are believed to have evolved from a shared ancestor no earlier than 200,000 years ago. In an evolutionary context, this is a relatively short period of time compared to the 4 million years that hominids have walked erect on the Earth (Kittles and Weiss 2003). The primary mode of subsistence during this vast period of time was hunting and gathering. Although it has been argued that famines are incapable of producing the differential mortality required to select for a thrifty genotype, it is, nonetheless, plausible that metabolic efficiency, as well as energy balance and storage, would have been a selective advantage in most human populations during the last few million years (McGarvey 1994; Speakman 2006). Hence, under the racialized TGH, there is a need for an intense environmental impetus to reverse this selective pressure for Europeans first (or only), in order to explain why this ethnoracial group does not now suffer from T2DM at the same rate as other ethnoracial groups. It has been suggested that agriculture, which became a dominant mode of subsistence 10,000 years ago for some human societies, could be this impetus (Jackson 1991).

Underlying the proposal of agriculture as this environmental impetus is the assumption that its advent was associated with an abundant food supply (De Courten, Bennett, and Tuomilehto 1997). There is little evidence, however, that

---

4 There have also been several variant hypotheses proposed with the salt-sensitive slavery hypothesis, arguably, the most prominent. Although beyond the scope of this paper, a detailed critique of this hypothesis can be found in Kaufman and Hall (2003).
agricultural societies were free of feast-and-famine cycles (Turner, Levy, and Clark 1993). In fact, it is likely that agricultural societies were at least as susceptible to feast-and-famine cycles as hunter-gatherers, with undernourishment rampant in the Greco-Roman empire and Europe both before and during the Industrial Revolution (Allen and Cheer 1996; Benyshek and Watson 2006; Speakman 2006). Even during the 20th century, famine was common during the First World War, the Great Depression, and the Second World War (Rock 2003). Thus, the overabundance of food associated with chronic disease, which disproportionately affects those from disadvantaged ethnoracial groups, is very much a phenomenon restricted to first-world nations in the second half of the 20th century. Furthermore, some hunter-gather societies were free of feast-and-famine cycles altogether (Winterhalder and Smith 1981). For instance, Nauruans and other Pacific populations, who now suffer from a very high prevalence of T2DM, formerly lived in thinly populated tropical islands that boasted a generous food supply all year round (Baschetti 1999).

To complicate this simple evolutionary story further, a number of disadvantaged ethnoracial groups, who are thought to now have the putative thrifty genotype, have been using agriculture for many thousands of years. For example, the Pima Indians, the leitmotif of the TGH, have been farming intensively for over 2,000 years (Schulz and Weidensee 1993). This historical perspective also reveals limitations in the hypotheses of McGarvey (1994) and Wendorf (1992), who fail to describe how the particular nutritional hardships suffered by indigenous peoples differed from those suffered by colonists of various human groups throughout history.

A further problematic assumption of the TGH is the notion that a disease such as T2DM could affect either survival to reproductive age or fertility during reproductive years of life (Lev-Ran 1999). Although conditions such as polycystic ovary syndrome (which are associated with T2DM) can have a marked effect on fertility, such an effect is unlikely to have evolutionary significance given the relatively short lifespan of Europeans (less than 50 years on average) during their agricultural history and the fact that T2DM has a post-reproductive age of onset (Weiss and Terwilliger 2000). Although it is plausible that putative genes associated with T2DM may have effects earlier in the lifespan, it is not at all clear that such effects would result in the selective pressure required by the racialized TGH.

Finally, one of the proposed explanations for why those of predominantly European ancestry seem to have a “resistance” to T2DM is that this group has the most genetic diversity and hence its members are less likely to harbor recessive diabetogenes (Lev-Ran 1999). Numerous studies have shown, however, that sub-Saharan African populations have the highest genetic diversity in the world, sometimes almost twice that of non-African populations (Jorde et al. 2000). Despite the conceptual difficulties detailed above, the racialized TGH remains one of the most widely cited hypotheses in all of genetic epidemiology. Below, we review the evidence that has accrued for and against this hypothesis.
The Thrifty Genotype Hypothesis: Genetic and Epidemiological Evidence

Before reviewing the evidence that has accrued for the TGH, it is useful to distinguish between two orientations to research on this topic: (1) research that seeks to understand the genetic basis of susceptibility to complex disease; and (2) research that seeks to understand variations in complex disease incidence/prevalence across populations. The former type of research is referred to here as genetic epidemiology, while the latter is labeled genetic epidemiology (Montoya 2003). The principle difference between these two orientations rests in the sociocultural implications of their respective uses of ethnoracially labeled DNA. Genetic epidemiological research utilizes population DNA to understand the genetics of disease. The objects of analysis are genetic. Ethnoracial labels (indigenous, white, African American, etc.) are used to describe the study populations involved in this form of research. The main reason why ethnoracially identified study populations are utilized in genetic epidemiological research is to guard against confounding due to population structure (mistaking a chance difference in allele frequency across cases and controls as being related to disease etiology).

In genetic epidemiology, by contrast, ethnoracial labels are used in a manner that does more than simply identify from which human group the study DNA was derived. In genetic epidemiology, the objects of analysis are human groups, and the ethnoracial labels define attributes of those groups. This type of research seeks to identify genetic contributions (variants or haplotypes) to disease susceptibility in a given population, and to discover how specific susceptibility variants segregating in defined populations explain observed differences in disease incidence/prevalence. Often, these two distinct forms of research occur simultaneously, which confuses rather than clarifies evidential claims in this field.

“Genetic” Epidemiological Research: Genetic Contributions to Disease

Neel (1976) once described diabetes as a “geneticist’s nightmare,” of which the main ingredient has always been T2DM. At least 60 distinct genetic disorders are associated with glucose intolerance and clinical diabetes (Vadheim and Rotter 1992). Multiple loci are almost always involved in a complex disease such as T2DM, often with a plethora of associated alleles, as well as gene-gene interactions and interactions with gene products (proteins) and environmental signals (Permutt et al. 2005). Genetic epidemiological research undertakes the challenging task of identifying the genetic basis of complex human diseases, such as T2DM, through research involving genetic material whose donors are described ethnoracially.

A number of studies have attempted to find genes or genetic variants associated with T2DM in ethnoracial populations as diverse as Pima Indians, First Nation Canadians, Indigenous Australians, people of Indian ethnic origin in
Guadalupe, Japanese, European Caucasians, African Americans, and Mexican Americans (Baier et al. 2000; Baier and Hanson 2004; Boullu-Sanchis et al. 1999; Busfield et al. 2002; Elbein et al. 2000; Evans et al. 2001; Fingerlin et al. 2002; Hegele 2001; Hegele et al. 1999, 2000; Horikawa et al. 2000, 2003; Majer et al. 1996; Muller et al. 2005; Rasmussen et al. 2002; Thompson et al. 1995). Many of these studies either failed to find candidate loci or associated susceptibility variants or failed to replicate associations reported by other investigators (Barroso 2005). A brief review of the most compelling evidence available for the influence of specific genes or genetic regions on T2DM susceptibility is presented below, highlighting the use of socially defined ethnoracial groupings for genetic discovery.

Genetic contributions to T2DM risk have been investigated via a combination of family-based (linkage) and population-based (association) approaches. While more than 50 T2DM linkage studies have been reported (Barroso 2005), only a relatively small number of genomic regions have shown consistent evidence of linkage to T2DM: chromosomes 1q, 3q, 8p, 10q, 12q, and 20q (Permutt et al. 2005). Not only has consistent evidence of genetic linkage to T2DM been hard to identify, but few individual genome scans have shown significant evidence for linkage in a single scan (LOD score > 3.6) (Florez et al. 2003). The strongest evidence for linkage, both with respect to the significance of the linkage signal and replication in multiple populations, implicates chromosomal regions 1q21–q24 (observed in Pima Indians, American and European Caucasians, Old Order Amish, and Chinese), 12q24 (identified in American and European Caucasians as well as West Africans), and chromosome 20 (observed in American and European Caucasians, Ashkenazi Jews, West Africans, and Japanese; Barroso 2005; Florez et al. 2003).

The first T2DM susceptibility locus identified directly from linkage analysis was the calpain 10 (CAPN10) gene on chromosome 2 (Horikawa et al. 2000). In this work, the linkage of T2DM to 2q in Mexican Americans was explained by a complex susceptibility genotype involving a pair of multi-variant CAPN10 haplotypes associated with increased risk in Mexican Americans as well as two European Caucasian populations. Several studies have replicated this finding, but a number of others have failed to do so (Barroso 2005; Florez et al. 2003). Moreover, a meta-analysis of 10 published case/control studies consisting of 3,303 subjects from British, Chinese, Japanese, Finnish, South Indian, and Mexican American populations concluded that the main risk polymorphism (SNP-44) was associated with only a 1.17-fold increased risk of T2DM (95% CI: 1.02–1.34; Weedon et al. 2003). Nevertheless, independent population genetic evidence has suggested that nucleotide polymorphism in the vicinity of the CAPN10 gene has been subject to the effects of natural selection in diverse populations, making it a plausible candidate for an evolutionary hypothesis (Fullerton et al. 2002; Vander Molen et al. 2005).

Nearly 30 other candidate loci have been similarly investigated for their asso-
cation with T2DM disease risk in various ethnoracially defined groups (Barroso 2005). However, genetic association studies have been plagued by the same problems of replicability as linkage analyses, so that currently only a handful of candidate genes are established as playing a confirmed role in T2DM susceptibility. Recently, Weedon and colleagues (2006) proposed that a reproducibly associated T2DM variant be considered one in which a meta-analysis of published studies reaches genome-wide levels of significance. By this criterion, they identify only three genes, PPARG, KCNJ11, and TCF7L2, as having been reproducibly associated with T2DM (Altshuler et al. 2000; Grant et al. 2006; Hani et al. 1998). In all three cases, the populations with confirmed evidence of association are European or American Caucasians; the PPARG association has also been observed in Japanese and Japanese-American samples (Barroso 2005; Florez et al. 2003). Risk variants in each of these genes only moderately predispose carriers to T2DM, with odds ratios ranging from ~1.15 to ~1.50 (Weedon et al. 2006).

In summary, current genetic epidemiological evidence suggests a diffuse and highly heterogeneous picture of genetic susceptibility to T2DM, with a restricted number of variants at a small number of loci showing any appreciable degree of contribution to disease risk. Many of the best described gene-disease associations have been observed only in European or American Caucasian samples, and replication across different ethnoracial groups has proven problematic. To their credit, genetic epidemiological studies do not generally situate genetic susceptibility with respect to membership in a given ethnoracially described group, preferring to summarize association as it is observed across different population samples. However, finding susceptibility gene variants will remain elusive as long as these studies and their readers mistakenly presume that sociohistorical identities remain constant over a lifetime, that external human morphologic features are concordant with deeper traits, and that ethnoracial labels are reliable genetic descriptors of human groups (Goodman 2001; Hahn, Mulinare, and Teutsch 1992). This last presumption is often implied in genetics-based research (genetic epidemiology) and thus shares many of the fallacies inherent in genetic epidemiology, to which we now turn.

**Genetic “Epidemiological” Research: Disease in Populations**

Genetic epidemiological research seeks to explain differentials in disease incidence/prevalence between various ethnoracial populations via the identification of genetic differences between these populations. In comparison to genetic epidemiological research, relatively little genetic epidemiological research has been conducted. This is particularly so in contemporary research which involves more sophisticated genomic technologies.

Of the existing research literature, admixture studies are the most common approach to genetic epidemiological research on the TGH. These studies seek to
determine whether and how much the incidence of T2DM (and associated factors) varies with an individual's ethnoracial ancestry after controlling for (or, more typically, assuming the absence of) differences in social/environmental risk factors between members of the study population. An association between a higher proportion of a particular ethnoracial ancestry and an increased prevalence of T2DM is taken to indicate that one or more genes inherent to the ethnoracial group in question are associated with a predisposition to T2DM.

There are a variety of approaches to measuring admixture in this research literature. An assessment of self-reported ancestry is the most common approach, with skin color and HLA antigen status also utilized (Gardner, Stern, and Haffner 1984; King et al. 1984; Serjeanston et al. 1983; Stein et al. 1965; D. R. R. Williams et al. 1987). Taken together, admixture studies provide some evidence that the prevalence of T2DM increases with an increasing degree of ancestry among Native Americans, Nauruans, and Mexican Americans (Brosseau et al. 1979; Gardner, Stern, and Haffner 1984; Serjeanston et al. 1983; Stein et al. 1965; R. C. Williams et al. 2000). There is also some research indicating similar associations between blood glucose/insulin levels and ancestry among populations in the Pacific and Indigenous Australians (King et al. 1984; D. R. R. Williams et al. 1987).

Unfortunately, many of these studies fail to adjust for potential confounders or adjust only for factors such as age, sex, BMI, tricep skinfold thickness, or obesity, rather than for possible social/environmental influences. In fact, data on income, as well as employment status, diet, and smoking, were collected in two of these studies but were not utilized in analyses (Gardner, Stern, and Haffner 1984; D. R. R. Williams et al. 1987). Similarly, BMI was measured in one study and found to vary by admixture but was still not examined in analysis (R. C. Williams et al. 2000).

Of particular concern is the offhand manner in which a number of these studies dismiss the relevance of sociocultural variables by failing to measure but nonetheless rejecting lifestyle factors (e.g., diet and physical activity), as well as environmental and sociocultural characteristics as possible influences on study findings. Such disregard for nongenetic factors is especially surprising given evidence that the differential prevalence of T2DM (and related conditions) across population groups can be explained by non-genetic factors alone (Dowse et al. 1990; Hazuda et al. 1988; Schulz et al. 2006; Tanaka et al. 2005). Moreover, a number of studies have found that degree of ancestry is unrelated to T2DM in Native Americans, Pacific Islanders, Mexicans, and Aboriginal Australians (Guest and O’Dea 1992; Marshall, Hamman, and Baxter 1993; Stern et al. 1992; Zimmerman, Kirk, and Serjeanston 1982).

As a whole, the genetic epidemiological research suffers from a failure to measure, adjust for, and consider in interpretation the social and environmental risk factors of known relevance to T2DM etiology. A number of researchers also continue to perpetuate the misconception that there are a priori genetic differences between people of various ethnoracial identities, as well as homogeneity within
these groups (Osei 1999; Schulz and Weidensee 1993; Stern 1999; R. C. Williams et al. 2000; Zimmet, Kirk, and Serjeanston 1982). Of most concern is a tendency in this body of research to misconstrue weak, inconsistent, and nonexistent evidence as supporting the racialized TGH, and a failure to acknowledge contrary evidence as casting doubt upon this hypothesis. For example, a recent study suggests that the prevalence of T2DM among U.S. Pima Indians is higher than that of Mexican Pima Indians due to differences in obesity and physical activity between these two populations. Despite this finding, and a genetic analysis which demonstrates that these two study populations are closely genetically related, the concluding paragraph of this study still describe Pima Indians as a “genetically highly susceptible population” (Schulz et al. 2006).

Overall, there are clearly considerable limitations to the evidence from both forms of genetic epidemiological research on the TGH. There is only very preliminary and ambiguous evidence for specific thrifty genes, and both gene-specific studies and admixture studies fail to sufficiently account for known and possible social and environmental causes of T2DM. The fact that over 250 genes have been studied as possible causes of T2DM, but together these genes explain less than 1% of diabetes prevalence worldwide (Lev-Ran 1999), should give researchers—and others—pause.

Ethnorace, Thrifty Genes, and Social Inequality

The etiology of T2DM is not explained by the genetics of population differences alone—genetic epidemiology. T2DM is also not wholly explained as a function of specific genetic predispositions of individuals, from whatever group—genetic epidemiology. Given this lack of consistent empirical evidence for the racialized TGH together with the conceptual limitations inherent to the hypothesis and continuing critiques (e.g., Fee 2006; Speakman 2006), it is unclear why investigators persist in assuming a pre-existing genetic susceptibility for certain ethnoracial groups that is “amplified” or “unmasked” by rapid environmental change (Bennett 1999; O’Dea 1997). With recognized difficulties in determining the independent effects of genes and environments, and the overwhelming influence of environmental factors on recent trends in T2DM, Occam’s razor would suggest that, in the absence of reliable evidence for such a susceptibility, the TGH is superfluous. Unfortunately, a feature of adaptive scenarios such as the TGH is that they cannot be explicitly tested (McGarvey 1994). As such, until perfect knowledge of human genetics is achieved, the endless search for the thrifty genotype will continue just as long as researchers persist in conflating race and genetics and misattributing genetic differences to socially defined groups.

It is only because the populations used in T2DM research have an a priori ethnoracial identity that their genetic material can be imbued with racializing potential. In general, ethnoracial groups are specifically selected for genetic epi-
demiological research because of the high prevalence of a disease of interest. It should be evident that such increased disease prevalence is due to specific environmental, historical, social, and political pressures interacting with individual genetic predisposition in complex and varying ways.

Ethnoracial population labels, if at all informative for genetic epidemiology, reference the physiological effects of a specific population’s history rather than their inherent genetic constitution. To derive genetic explanations for such prevalence differences without controlling for environmental confounding tautologically reiterates “the social conditions from which the ethnic groups are derived,” rather than providing insight into individual or group genetics (Montoya 2003, p. 39). At the same time such research, almost paradoxically, renders invisible the social etiology of diseases such as diabetes.\(^5\) As a consequence, social inequality is embodied as a form of individual pathology, while the “hegemony of biological solutions for a chronic sociological dis-ease” is asserted (Montoya 2003, p. 40). Moreover, this use of ethnoracial labels distances researchers from the important work of locating the physiological and environmental triggers for complex diseases (Braun 2002).

Research demonstrates that scientists are often aware that the ethnoracial labels they utilize are inaccurate and politically loaded, and that they are concerned by the possible social consequences of their research (Montoya 2007). As such, it is important to recognize that racialization is a sociocultural phenomenon and not a characteristic of individual scientific projects, much less individual scientists. In addition, there are considerable epistemological barriers to imagining disease etiology as both socially and genetically mediated—barriers that reinforce professional pressures to conduct disembodied genetic epidemiology.

Ethnoracial health research can, however, incorporate genetic knowledge as a way of attending to “local biologies” of health and illness (Lock 1994). To do so will require rigorous interdisciplinary collaboration between geneticists and social scientists in order to develop new approaches to understanding the intersections between ethnorace, genetics, biology, and society (Fee 2006).

**Rehabilitating Genetic Epidemiological Research**

Before discussing the possibilities of interdisciplinary genetic epidemiological research, we shall consider the three classes of rehabilitations that have been presented as solutions to the unreflexive recapitulation of genetic distinctions among ethnoracial groups. First, there has been an interest in bypassing ethnoracial alto-

---

\(^{5}\)Interestingly, advocates of the TGH regularly recommend environmental (lifestyle and sociopolitical) change as a remedy for disparate rates of T2DM, even as they situate the disease as primarily genetic in origin (Osei 1999; Zimmet 2000).
gether, via an exclusive emphasis on the identification of disease-linked genes shared among family members. Family-based studies, which traditionally have been the preferred methodology of human molecular genetics, represent the best way to enrich one’s sample for defined genetic determinants and to provide some reduction in genetic heterogeneity relative to population-scale association studies (Khoury and Yang 1998). Moreover, family-based studies largely bypass the issues arising from population stratification. These advantages, however, must be weighed against the difficulties inherent to family-based recruitment, the greater degree of genotyping required to achieve similar statistical power, and the fact that the variants identified may be rare and family-specific, and hence contribute little to population attributable risk (Hirshchhorn et al. 2002).

A second, increasingly popular, alternative to family-based genetic identification is to employ so-called “race neutral” approaches. These methods, which fall broadly under the banner of “genomic control,” involve either the genotyping of several dozen to several hundred random genomic markers to determine the degree of sample stratification or the genotyping of a smaller number of “Ancestry Informative Markers” or AIMs, preferentially chosen on the basis of their ability to distinguish the presumed ancestral populations contributing to varying degrees of admixture among cases and controls (Devlin et al. 2001; McKeigue 2005). While such methods may appear, at first blush, an appropriate work-around, allowing genetic epidemiologists to control for confounding without any recourse to ethnoracial ascription, they misleadingly imply that genetic confounding is the only confounding of interest, that environmental exposures and other life experiences are irrelevant to the assessment of disease susceptibility, and that AIMs themselves do not construct inaccurate typologies. Research that attempts to explain disease prevalence between human groups using such techniques alone will almost always conflate the descriptors of the population with the attributes of the population. It is precisely this racialization we seek to thwart.

The third option is an approach to genetic epidemiological research that recognizes that population health differentials reflect the embodied interaction of genetic and sociopolitical conditions (Krieger 2005). Such an approach would require detailed information on genetic variation throughout the genome and detailed information on social and environmental exposures (Ellison and Jones 2002). A first step toward such socio-genetic research would be to acknowledge that the use of ethnorace as a proxy for genetic differences alone limits our understanding of the complex interactions among sociopolitical processes, lived experiences, and human biology. In keeping with this, perhaps it is appropriate for genetic epidemiological researchers consistently to include a caveat that genetic variation exists between any human groups—however defined—and thus it is erroneous to assume that such variation reflects “racial categories” (Duster 2005). Such a caveat would not preclude the study of genuine genetic variation across human populations as long as such research considers gene-envi-
environment interactions and also measures, stratifies by, and controls for social and environmental factors, including sociocultural and sociopolitical variables such as racism, stress, and class (Cooper, Kaufman, and Ward 2003; Ellison and Jones 2002; Kittles and Weiss 2003; Weiss and Terwilliger 2000).

Genetic epidemiological researchers need to move beyond the mere recognition of ethnorace as socially constructed and towards the practice of studying ethnoracial health disparities as the biological embodiment of sociopolitical processes. Research that combines the study of genetic and social etiology without reinforcing dangerous notions of racialized genetics will allow us to understand why some gene-disease associations have varying strengths across human groups. To ignore such ethnoracial differences in disease outcomes would be just as wrong as essentializing those differences as simply genetic in origin. When prioritizing research, however, it is important to consider that, at present, environmental factors are (at least technically) far more amenable to modification than genetic ones, and that they are likely to have a greater impact on public health (Merikangas and Risch 2003).

The challenge of interdisciplinary research in this area is to link population health parameters, including environmental exposures and social conditions, with both pathophysiology and genetic risk estimates. Such an endeavor will require the development of interlinked genetic, sociocultural, and ecological constructs via careful attention to the life course of individuals, families, and social groups. Methods to achieve this may include in situ participant-observation, environmental scans, informal interviews, time use and behavioral inventories, and life histories. We imagine, for example, the creation of ethnographically derived characterizations of the life conditions of human groups affected by complex disease, both to enable multivariate analyses of a range of biocultural constructs and to situate complex disease within the broader sociocultural environments and institutional arrangements through which contemporary humans traverse (Goodman and Leatherman 1998). Data and concepts derived from such research will lend untold power to biobank inventories, quantitative surveys, and other biophysical data, including genetic data, by allowing questions to be drawn simultaneously from both sociocultural and biological constructs. This approach, we believe, more accurately reflects human life conditions—especially those related to health and disease. Conversely, an absence of such ecologically oriented biocultural data will perpetuate the reductionistic disconnect between genotype and phenotype, ecology and well-being (Krieger 2005). While research into the dynamic between human life conditions and human biology is more challenging, we believe that this approach offers the best hope for understanding and addressing health inequalities.
References

Cacioppo, J. T., et al. 2000. Multilevel integrative analyses of human behavior social neu-


