Genetic variation and nutrition[1,2]

Barton Childs, MD

Introduction

It is a central aim of biology to discover the molecular basis for cellular structure and function. It is presumed at the start that these purposes are mediated by proteins, each of which is specified and controlled by particular strands of DNA. These strands, or genes, vary in the sequence of their nucleotide units in ways that are reflected both in the sequences of amino acids in the proteins they specify as well as in their amounts and the times they appear in development. Much of this variation is bad; some is lethal, eg, it is accepted that perhaps 75% of all conceptuses never make it to term. Some is expressed as normal variation and some is conditional—potentially harmful depending upon circumstances. Among the many conditions to which the human organism must adapt is the nutritional environment. It is simply unthinkable that the whole species is equally adapted to the whole range of nutritional variation; so, the questions at issue are, How variable genetically is the human species? How variable is the environment which the human species inhabits? and How congruent is the genetic variation with that environmental diversity, whether nutritional or other?

The extent of human genetic variation

Let us begin with the question of how genetically variable the human species is. That depends upon how the variation is detected. When we look directly at the DNA, there is a great deal, eg, on average every 200–300 nucleotides there is a substitution that distinguishes from 1 to 50% of the population. So given 3×10^9 nucleotides, that is something like 6 000 000 options of which any one individual has quite a few. When we look at the level of protein diversity—say variation in sequence of amino acids—we find much less. No surprises there; much of the DNA is not exposed to the strict scrutiny of selection. So most of the DNA variety is in the flanking sequences and introns where it is less likely to be reflected in any physiological phenotype. In contrast, mutational variation of the proteins is far more likely to be expressed in just such physiological differences so it may be more important to know its extent. The question has been exhaustively examined in a very wide range of organisms including homo sapiens (1). It appears that in man—and in a surprising proportion of all organisms examined—perhaps 30% of loci have polymorphic variants, which is defined as two or more alleles with frequencies of at least 1%, often much more. That is not to say that the remaining 70% have no variants; a determined search is likely to turn up a variant for any protein but that 70% represents our humanness. It is the reason why we all resemble one another in so many particulars and why when we differ, we do not differ very much. Therefore, the 30% of loci is where the ordinary, frequent variation comes from; it is that which saves us from being a too-homogeneous species with serious risks of extinction.

Of course, there are, in addition, infrequent mutants of all kinds, including chromosomal aberrations, deletions of various sizes (a result often of unequal crossing over), nucleotide substitutions, frameshifts, and all sorts of mutations that alter the mechanisms of transcription and translation (2). Variation is, on the one hand, frequent and usually more or less tolerable or, on the other, infrequent and accounting for effects going all the way from lethality to making no difference at all. From all this variety in the gene pool, each person draws a unique aliquot. The genetic differences between people are based on the specificities of the genes drawn from the pool and particular combinations of those genes. By virtue of opportunities for gene interaction, the latter source (ie, the combinations of genes) contributes as much to variability as the former. Thus the genes provide each individual with a set of options: ranges of reactions to all potential experiences. How those options turn out, obviously, depends upon the ranges of events experienced.

Variation in experience

The extent of variation in experience is far less easily systematized. A gene is inherited or it is not. It can be counted and its distribution in populations and its presence or absence in individuals can be qualitatively as-

[1] From the Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD.

[2] Address reprint requests to Childs, Professor Emeritus, Pediatrics, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205.

Received July 26, 1988.

Accepted for publication August 2, 1988.

Am J Clin Nutr 1988;48:1500–4. Printed in USA.

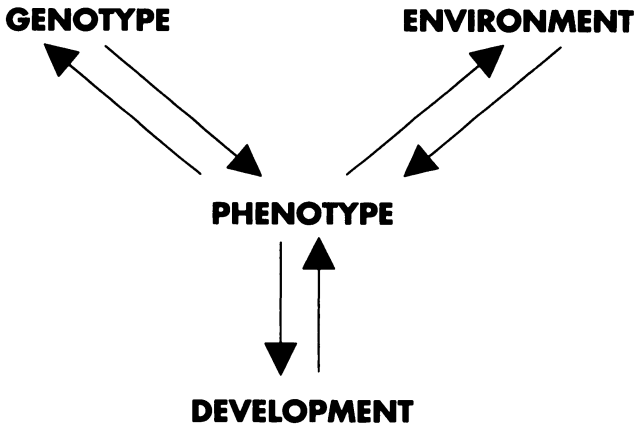

FIG 1. Relationships among genes, environment, and development are dynamic.

sessed. Descriptions of experiences are more vague and are difficult to express qualitatively or quantitatively; for example, we may wish to know how many people in a population eat animal fat, what kinds, how much and for how long, and in what relation to vegetable or fish oils. Still, however difficult to obtain, it is the business of epidemiologists to collect such data, to correlate them with serum lipids, and to assign risks associated with them. However, the meaning to a particular person of relative risks based on such population data is not always clear.

Also not clear are the rules governing the transmission of the habits, learning styles, and qualities of behavior that inform experiences as straightforward as the laws of inheritance. Cavalli-Sforza and Feldman made an attempt to impose some coherence on cultural inheritance (3). They argue that although natural selection works on phenotypes and their genes, cultural selection chooses ideas, languages, beliefs, rules, tools, and so on. Mutant forms, eg, new beliefs, new ideas, new tools, etc, come into existence and are transmitted in diverse ways, eg, within families from parent to child but also tranversely from child to child both at home and at school. Transmission may be also from one to many as in the case of rock stars or other cultural leaders, including advertisers, to their broad constituencies or many to one as with peer pressures or social mores. Obviously, the nutritional environment, specifically nutritional habits, is strongly shaped by these cultural influences and some systematic knowledge of their impact, flow, geographic and ethnic distributions, and transmission should be immensely helpful in obtaining an answer to our third question, that of congruence and incongruence between genetic and environmental variability.

The developmental matrix

The question of gene-environment interaction cannot be examined rationally in the absence of a third factor that represents the matrix within which such interactions take place, that is, development. Development, of course, is itself a product of genes acting in the context of experiences, which means that it is a historical process, not programmatic. Genes do not provide an unalterable blueprint but merely a set of options, each more or less conditional and to be taken up according to what is being experienced as well as what has been experienced. What happens next is always conditioned by what happened last and this process continues throughout life. If we are made uneasy by thinking of aging as development, we may at least accept that we all continue to evolve as long as we live. No doubt the options become fewer and narrower with age but they continue to exist. Thus, gene-environment interaction must be described in the context of development (**Fig 1**) and in a dynamic, dialectical mode (4).

Clearly, genes and experiences affect phenotypes, and the developmental contexts in which the gene-environment transactions take place affect them too but there are reverse relationships. For example, there is no question that phenotypes affect the environment; a case of PKU affects the whole family and in turn the dietary treatment affects the phenotype. Similarly, the mental retardation of untreated PKU is a property of the prolonged delay in the development of the human brain; the same metabolic error in an animal born (as many are) with a far more advanced neural development might be more of a variation than a disease. Mental retardation becomes a part of the phenotype with effects on the family and the community at large too. Finally, since it is phenotypes that are selected, reproductive outcomes affect the gene pool.

What we wish to have is a way of systematizing gene actions and their effects and to be able to do so for each individual case. For example, it is said that the cause of type 1 diabetes is infection of the islet cells followed by an autoimmune response but the causes for any one individual are another matter. **Table 1** lists the elements that go into a phenotype. They are arranged in the same three groups as in Figure 1, with the genes on the left, the developmental matrix in the center, and the sources of experiences on the right. All sorts of interactions among these three sources of variation are possible. Type 1 diabetes

TABLE 1
The elements of individuality

Genes	Constitutional factors	The environment
Major genes	Age	Time
Modifiers	Sex	Geography
	Developmental stage	Climate
	Homeostatic specificities	Socioeconomic status
	Biochemical	Occupation
	Immunological	Education
	Physiological	Dietary habits
	Maternal factors	Other habits
	Cognitive qualities	Other diseases
	Temperament	
	Ethnic group	

TABLE 2
Factors involved in type 1 diabetes

Genes	Developmental factors	Environmental factors
HLA, DR4, DR3, DR2 (5)	Age	Infection (6)
	Maternal age (8)	Season (9)
GM types (7)	Sex of affected parent (11, 12)	Location (13)
C4 BQ0 (10)		Year (13)
B1f1 (15)	Pregnancy (14)	Breast-feeding (16)

provides a good example (**Table 2**). On the left in Table 2 are several groups of genes of the immune system (including the HLA [5], the Gm groups [7], and the complement pathways [10, 15]) that have been implicated either in cause or as markers of unknown genes presumed to be involved in cause. In the middle column are several developmental factors and on the right are some experiences, which have been shown to be associated with varying degrees of certainty with some cases. For example, the onset of overt disease is more likely in the winter months (9) and the incidence of the disease varies significantly from year to year and place to place (13). There is an unconfirmed report of an inverse relationship over the years between the frequency of breast-feeding and the annual incidence of diabetes in Sweden (16). The implication is that, in some way, breast-feeding reduces the susceptibility to diabetes, perhaps by protecting against a viral infection. What needs to be done is to put all these things together for each individual. We know already that early age at onset is more likely to be associated with ketosis and coma (17), a concentration of the relevant genes (18), a positive family history (19), and increased concordance of MZ twin pairs (20). We also know that when the affected parent is the father, there are more affected sibs (11). This is due to a segregation distortion in which the father transmits preferentially his DR 4 gene (12). However, the kind of information we lack is, for example, whether diabetic children born to older mothers are more or less likely to have the DR alleles or to have early onset and so on. One might guess not; rather the maternal age effect might be bringing into the data cases that might otherwise escape. Perhaps they are more likely to have onset in the winter or, when the incidence rises or drops from one year to another, the cases might differ in the representation of any of the genetic or developmental qualities. For example, if the breast-feeding relationship were to turn out to be real, one might suppose that under conditions of widespread breast-feeding, the remaining cases of diabetes would be those with the heaviest concentration of genes, that is, the most vulnerable. The least vulnerable would have been protected by whatever is in breastmilk, if anything, or the presence or absence of some other experience may account for the rise and fall, eg, the local prevalence of some virus. So, what is needed is a combined genetic-epidemiologic attack which assigns genes, experiences, and developmental characteristics to individuals.

Incongruence as disease

Let us return now to the question of congruence between genetic variety and environmental diversity. Obviously, congruence represents health whereas incongruence represents disease and it follows, as the night the day, that the greater the incongruence the worse the disease, that is, the heavier the burden as measured in shortened life, curtailed reproduction, and social handicap. Nature abhors incongruence and so will eliminate those who express it as speedily as the disharmony is made known. (In the Western world, we may read morbidity for mortality.) There should be some sort of selective gradient in time along which cases will be eliminated (or have onset of disease) according to degree of dissonance; the worst go first. Empirical observation bears this out. Distributions of mortality rates by age show declining losses early in life and rising losses in adulthood with a nadir during the optimal reproductive years. Such distributions, however, lack one important detail; they cannot include intrauterine life. There is no census taking there.

Gestation is no less a part of life than infancy, adolescence, or old age and so must be included. In fact, the loss of life just after conception is very great, perhaps as much as 75%. Most of the early losses that we are able to observe are associated with the most extensive discordance—deletions and other abberations often involving whole chromosomes, mutations simply incompatible with development and life (21). We must suppose that there are, in addition, single gene lethal defects. The genetic storm spends itself early in gestation so that the peace of later phases is broken mainly by the less frequent intrusion of disruptive influences of the maternal environment: alcohol, smoking, drugs, abnormal placentation, birth injury, and so on.

Next there is a postnatal rise in mortality, consisting of developmental anomalies that cannot make it in extrauterine life as well as inborn errors in all their baroque variety associated with rare variant genes and challenging, one supposes, every element of every homeostatic system. Many of these may be perceived legitimately as nutritional disorders but of a special kind. They represent lack of, intolerance to, or failure to respond to essential elements of nutrition. The special clinics of pediatrics departments are populated by the least lethal of these examples of human molecular, biochemical, and immunological imperfection. The onsets of these diseases too are aggregated into the early years so that mortality in later childhood and adolescence is associated with poisons, mayhem, and homicide (22).

Finally, in adult life, the cases composing the early part of the climbing limb of the distribution of mortality rates is associated with more common, often polymorphic genes whereas the very old die either of the residua of the kind of disorders that took the lives of the younger people or of old age: of what Pope described as "this long disease, my life." It makes evolutionary sense that there should be no genes whose special duties are to contribute to longevity. Rather, in the Western world, long life,

omitting random disaster, is simply due to the absence of genes likely to shorten it. Evidence is based on the observation that incidence curves for many diseases show a decline in the later years and patients with early onset of postpubertal disorders are more severely affected and more likely to have relatives who share the disease (23). The point was made compellingly in a report from Denmark, that Eden of adoption studies (24). Sorensen et al showed that early mortality (between 16 and 58 y of age) of adopted persons was more strongly correlated with early mortality (< 50 y of age) of the biological than the adoptive parents. Of particular interest was the correlation for deaths from infections. Infections are the archetype of "environmental" disease and yet Sorensen et al show a sixfold greater risk of dying from infections for adopted-biological parent pairs than for adopted-adoptive parent pairs.

A moment's thought about the possibilities for mutational variation among the molecules of the complex human defense system as well as those of the microbial mechanisms for virulence suggests that such a correlation is inevitable. There was a similar correlation for deaths from vascular causes but that from cancer went the other way; the adopted person was more likely (relative risk 5.16) to resemble his adoptive parent. However, that is not surprising. Although the childhood and early adult cancers show strong germ-line heredity, later cancers are associated with somatic mutations presumed to be a result of environmental provocation, certainly including the nutritional environment. That is just what we expect; environmental variation (here mediated by somatic mutations) plays a more salient role in diseases with onset late in life.

There does seem to be a gradient of selective intensity that wanes throughout life but nonlinearly so that the latter part of each developmental phase is characterized by the onset of less disease with less of it from genetic diversity. There are some lessons here which I would like to illustrate, as so much of life can be illustrated, by a nursery rhyme, and one of topical interest to those interested in nutrition. You will remember that Jack Sprat could eat no fat and his wife could eat no lean and so between the two of them they licked the platter clean. The *Oxford Dictionary of Nursery Rhymes* is silent on why the Sprats had such divergent tastes but it might be that Jack has some form of hypercholesterolemia and his wife has a mild case of some protein intolerance. He might have one of a number of the rather common genetic variants of the several lipoproteins engaged in cholesterol transport and disposition whereas hers is a rare genetic defect in the homeostasis of some essential amino acid. What are the implications of these disorders for the lives of Mr and Mrs Sprat? Jack may well have had an uneventful life for his first 50 odd years. True, it may be brought to a sudden and untimely end, but because his children have grown up, he has completed his biological mission so that although Mrs Sprat, their children, and their devoted friends would miss him sorely, the species would not. On the other hand, his life may be prolonged by observing some sensible dietary and other routines. Alternatively, because the risk figures he has been given have to do with populations not with individuals, he may be much less at risk than he thinks. That is, there may be other factors that could intervene to reduce the risk accountable to his lipoprotein variation and he may be due for a long life, only to suffer the early loss of Mrs Sprat who may already have stretched her life beyond what might have been possible under other conditions, dietary and otherwise. In contrast to Jack, who has a condition in which his homeostasis is overtaxed by a life of abundance, she has a homeostasis that was defective from the start. She represents a mild version of the diseases of early life, strongly affected by genetic variation and under intense selective pressure, whereas he represents a moderately severe version of the diseases of later life, associated still with genetic diversity but with more prominent participation of experiences. In both disorders both genes and experiences play their parts. In her case the genes are more disruptive, less frequent, and the environment is universal—we all require the essential amino acids. His more common genes are less damaging, perhaps under some circumstances not damaging at all; they have allowed him a healthy life into middle age; and his environmental provocation is nonessential and restricted and can be changed although not necessarily easily. Old habits die hard.

Conclusion

I said at the start that it is a central aim of biology to discover the molecular basis for cell structure and function. That is a central aim of medicine too, where we wish to elaborate the mechanisms of pathogenesis in molecular terms. What has been neglected until recently is that all these proteins that mediate cell structure and function and that are the focus of investigations into pathogenesis are specified or otherwise controlled by genes that are more or less equally prone to mutation. In time such genes will be susceptible to examination by recombinant DNA techniques. The energy, fervor, and zeal with which investigators are exposing these genes is such as to lead us to suppose that all the genes we want to see at close range will have been examined in detail in a generation or two. This means that the genetic causes of disease will have been reduced to the elements, with inevitable benefits in unraveling pathogenesis, even in the absence of precise knowledge of provocative experiences.

One important outcome of such an elemental grasp of genetics is the breakdown of the limiting typologies of genetic or nongenetic (environmental) disease. Chromosomal, monogenic, and multifactorial phenotypes are all a product of the actions of genes in environments (nutritional among others) against the developmental backdrop. All are inborn errors of metabolism and all should be susceptible to a similar analysis. The language we use now, which contributes to these typological distinctions,

is a heritage of work in the flyroom at Columbia where it was elaborated 70–80 y ago by observations of morphological variations in drosophila. Some embellishments were added by microbial geneticists but those too were based on phenotypes not actual gene actions. As more and more genes are discovered and their products are described, our focus will shift from gene-phenotype relations observable at a distance to how the products of each gene works or how several or many genes conspire together to homeostatic ends. Then the difference between single gene and multifactorial phenotypes will disappear and with it the need for such concepts as dominance, penetrance, and so on. The aim of investigation will be simply to determine how some gene or set of genes made an individual vulnerable to the adverse effects of experiences. Perhaps when enough of the human genome is known, it may be possible to distinguish genetic contributions to differences in normal qualities, blood pressure or glucose metabolism, for example, and no doubt, still later, to cognitive differences. However, these need not concern us now. We have enough questions of treatment and prevention and of the formation of public policy in regard to genetic variation. Other papers in the symposium will attend to these.

References

1. Lewontin R. Population genetics. Annu Rev Genet 1985;19:81–102.
2. Kazazian HH, Antonarakis SE. The varieties of mutation. In: Childs B, Holtzman NA, Kazazian HH, Valle DL, eds. Molecular genetics in medicine. Progress in medical genetics. Vol VII. New York: Elsevier, 1987:43–67.
3. Cavalli-Sforza LL, Feldman MW. Cultural evolution. Am Zool 1986;26:845–55.
4. Lewontin RC. Gene, organism and environment. In: Bendall DS, ed. Evolution from molecules to man. Cambridge: Cambridge University Press, 1983.
5. Eisenbarth, GS. Genes, generator of diversity, glycoconjugates, and autoimmune B-cell insufficiency in type I diabetes. Diabetes 1987;36:355–64.
6. Pak CY, McArthur RG, Eun H-M, Yoon J-W. Association of cytomegalovirus infection with autoimmune type-1 diabetes. Lancet 1988;2:1–4.
7. Rotter JI, Vadheim CM, Raffel LJ, Rimoin DL, Riley WJ, MacLaren NK. Genetic etiologies of diabetes. Pediatr Adolescent Endocrinol 1986;15:1–11.
8. Wagener DK, LaPorte RE, Orchard TJ, Cavender D, Kuller LH, Drash AL. The Pittsburgh diabetes study. III An increased prevalence with older maternal age. Diabetologia 1983;25:82–5.
9. Fishbein HA, LaPorte RE, Orchard TJ, Drash AL, Kuller LH, Wagener DK. The Pittsburgh insulin-dependent diabetes mellitus registry: seasonal incidence. Diabetologia 1982;23:83–5.
10. Marcelli-Barge A, Poirier JC, Schmid M, et al. Genetic polymorphism of the 4th component of complement and type I diabetes. Diabetologia 1984;27:116–7.
11. Warram JH, Krolewski AS, Gottlieb MS, Kahn CR. Differences in risk of insulin-dependent diabetes in offspring of diabetic mothers and diabetic fathers. N Engl J Med 1984;311:149–52.
12. Vadheim CM, Rotter JI, MacLaren NK, Riley WJ, Anderson CE. Preferential transmission of diabetic alleles within the HLA gene complex. N Engl J Med 1986;315:1314–8.
13. Rewers M, LaPorte RE, Walczak M, Dinchowski K, Bogaczynska E. Apparent epidemic of insulin-dependent diabetes mellitus in midwestern Poland. Diabetes 1987;36:106–13.
14. Buschard K, Buch I, Mølsted-Petersen L, Hougaard P, Kuhl C. Increased incidence of true type I diabetes acquired during pregnancy. Br Med J 1987;294:275–9.
15. Schernthaner G, Mayr WR. Immunoglobulin allotype markers and HLA genes in type I diabetes mellitus. Metabolism 1983;33:833–9.
16. Borch-Johnsen K, Mandrup-Poulsen T, Zachau-Christiansen B, Jones G, Christy M, Kastrup K. Relation between breast-feeding and incidence rates of insulin-dependent diabetes mellitus. Lancet 1984;2:1083–6.
17. Japan Childhood Diabetes Research Group, Pittsburgh Childhood Diabetes Research Group. Coma at the onset of young insulin-dependent diabetes in Japan. Diabetes 1985;34:1241–6.
18. Cavender DE, Wagener DK, Rabin BS, et al. The Pittsburgh insulin-dependent diabetes mellitus (IDDM) study. J Chronic Dis 1984;37:555–68.
19. Chern MM, Anderson VE, Barbosa J. Empirical risk for insulin-independent diabetes (IDD) in sibs. Diabetes 1982;31:1115–8.
20. Johnston C, Pyke DA, Cudworth AG, Wolf E. HLA-DR typing in identical twins with insulin-dependent diabetes: difference between concordant and discordant pairs. Br Med J 1983;286:253–5.
21. Boue A, Boue J, Gropp A. Cytogenetics of pregnancy wastage. Adv Hum Genet 1985;14:1–57.
22. Costa T, Scriver CR, Childs B. The effect of mendelian disease on human health: a measurement. Am J Med Genet 1985;21:231–42.
23. Childs B, Scriver CR. Age at onset and causes of disease. Perspect Biol Med 1986;29:437–60.
24. Sorensen TIA, Nielsen GG, Anderson PK, Teasdale TW. Genetic and environmental influences on premature death in adult adoptees. N Engl J Med 1988;318:727–32.