T/18 Weissman lecture UCB 3/90 -- re genetic load This is really two talks --

1) To honor the public health motif of this setting. some perspectives on environmental mutagenesis as they affect the human gene pool. and

2) A more specialized consideration of the "spontaneity" of mutation in bacteria, and some reflection on what this may mean for evolutionary thinking.

Environmental Mutagenesis

In recent years this movement has turned very strongly in the direction of the somatic, namely the cancer implications of environmental mutagens. Thanks in large measure to the marvelous work led by Bruce Ames, mutagenic assays in microorganisms are a proxy for expected consequences for somatic mutagenesis: namely cancer in higher organisms. I will leave to him further discussion of the merits and problematics of that assay procedure. Meanwhile considerations of environmental influence on genetic hygiene surface from time to time, including improbable allegations of transmission of birth defects through males exposed to hormones. But they have become relatively unfashionable as part of scientific discourse.

The issues of genetic load are out of favor since they carry the freight of all the headaches of

o existing genotype and its implications for reproductive behavior,

o ethnic differences, genetic screening dilemmas,

o antenatal diagnosis and selective abortion

o in general some sense of futility about what to do with imperfect knowledge that is fraught with ideological controversy.

Carcinogenesis has few of those difficulties. There is less ideological controversy: no one defends the privacy of their own cancer cells but would rather seek to suppress them. Many people are jealous of their very own gametes regardless of what they may do to or for the lives of their children. And they certainly do not welcome anyone else taking any interest in them. On the other hand, there are publically visible external villains to point to, e.g. the chemical industry and other creatures of the establishment.

Nevertheless genetic load accounts for at least 20% of our overall morbidity with a given system of care and prophylaxis. Only a quarter of that is attributable to major 1-gene defects like Huntington's or cystic fibrosis; the rest is the genetic contribution to complex syndromes like heart disease, schizophrenia or cancer.

I should explain the allocation of variance. If juvenile diabetes or PKU were untreated then the genetic load would account for a still higher morbidity. On the other hand if we lacked vaccines and antibiotics to deal with external infection, this would predominate, and the contribution of the genetic load would be proportionally diminished (although it does of course interact with those exogenous agents).

In principle we could greatly reduce the genetic load with compensatory environmental amelioration and I advocate that as the most practical measure whenever feasible. In fact almost 30 years ago I introduced the expression euphenics as a slogan in opposition to eugenics.

Nor am I addressing species deterioration, which was of such concern to Julian Huxley and H.J. Muller and their eugenics movement. Nazism showed where that could lead us in the final extremity. In any case our current knowledge is so primitive we have only the clumsiest diagnostics; and the interventions are even cruder.

But the burden of illness for ourselves and our children does weigh heavily and we ought to be looking at the understanding we need to have for any hope of mitigating it. (I will remark later how this does or does not connect with the human genome project.)

When I refer to 20% of the burden of morbidity related to genetic causes, these have to do with polymorphisms referred to the standard normal genotype. There may be a few examples of rate genes that give an exceptional positive advantage. C.J. Glueck has reported some families with high HDL, low LDL that protected against coronary disease with no evident penalty in other arenas. For a variety of reasons this work needs to be corroborated but it is already a standard for euphenic, i.e. dietary and pharmacological interventions to achieve a similar shift of gene expression as between these two apo-lipoprotein moieties.

We should not think this is a evolutionary anomaly -- Malthusian fitness has little to do with survival beyond age 50 or 60, and longer survival is an artifact of modern civilization quite out of the domain of species evolution.

I turn now to some round numbers to described the genetic load. That genetic load is part and parcel of the evolutionary process but we should recognize that continued evolution -- whatever we think of the state to which it has brought us -- is no longer acceptable. The price is too high. The estimates I am going to give on genetic load are quite crude; and they are undoubtedly frail and oversimplified in concept as well. We have had high hopes that the new methods of molecular genetics would lend greater clarity - in some respects they have. But they have uncovered a background of individual variability at the DNA level whose significance for biological performance is bewildering. But let us at least try an introduction.

One of the discouragements of dealing with environmental mutagenesis is that it obliges us to confront the existing genetic load, which corresponds to some scores or hundreds of generations of accumulation of historical evolutionary backlog with which we are in equilibrium, paying the price of natural selection as the one means there is of reducing the load in the gene pool.

Sometime ago H.J. Muller pointed out that it requires on average "one genetic death", that is to say a reduction in Malthusian fitness equivalent to removing of one individual, to remove at equilibrium each new deleterious mutation introduced into the gene pool. This would apply no matter how slight the deleterious effect since there would be no other way of diminishing that gene's frequency. In that case 100 parents might have a 1% reduction each of relative fecundity, a diminution that by itself could never be detected. Many mutations are so nearly

neutral that they are never or very slowly eliminated and they remain in the residual gene pool, the counterpart of not having to pay the bill to account of that genetic death.

CHART:

In round numbers I will start with the estimate that we have 3 billion base pairs in our genome: allocated to about 100,000 effective gene loci; that these comprise about 1,000 base pairs of structural information and 2,000 of regulatory that will add up to about 300,000 base pairs of genic functionality. That leaves about 90% of the genome in a category that we know no better than to describe as junk. The leading hypothesis for this junk is that it serves no significant purpose for the host organism but is a reflection on the propensity of DNA itself to replicate and to propagate i.e. that it reflects parasitic or selfish DNA. Mutations in this 90% of junk are presumably are of no consequence to us except that as they reflect just this process of genic parasitism and its potentiality for spread to more vital parts of the genetic machinery.

DNA structural studies have shown that the average individual is heterozygous at something like 1 to 4 per 1,000 base pairs, namely that there is at least 1 nucleotide substitution in every gene, as part of our evolutionary backlog. Most of these nucleotide substitutions are silent, either being caught up in the redundancy of the genetic code, or resulting in amino acid substitutions that do not alter the functionality of the protein gene product. But about 10% of these, that is to say about 1 in every 10 of our genes or 10,000 of them, have changes that are phenotypically consequential (at the level for example of electrophoretic mobility or thermostability) though not necessarily impairing their routine biological function. We guess that about 1/10 of those or 1,000 loci have some bearing on biological performance in either heterozygous or homozygous condition. For the latter, these are the spread of the 4 lethal equivalents in recessive genes that I mentioned before. In a given individual (not the product of consanguineous mating) a handful, perhaps 10, of these are homozygous and 1 of these 10 recessive homozygous genes in each of us may account for a large part of our individual idiosyncrasies of health. The other 9 plus the 1,000 heterozygous loci contribute to the rest so we are indeed a highly polymorphic species: the overwhelming majority of that polymorphism contributes to our personal indivduality; it has accumulated in ways that suggest that it has not been a target of negative selection.

Fortunately, the deleterious recessive mutations are largely expressed as lethals in early development which may be no more than a missed pregnancy or, through progressive stages, a miscarriage, or worst of all surviving to term as a significant birth defect, as is consequential for about 1% of births with disease of early onset or manifestation. Deleterious dominants are rather quickly weeded out.

Turning now to the origins of the load it can be classified as

o mutational,

o segregational owing to heterozygote advantage or

o in a grabbag of violations of the usual rules of Mendelian transmission like meiotic drive.

CHART -- mutational.

The idea that we start with a backlog of 10s to 1000s of generations of ongoing mutation worth of backlog load may inspire some nonchalance about the importance of new mutations. We add that there is probably no statistically significant example of induced mutations in the human germ line, even the Hiroshima exposure having failed to give a clear cut statistic. So most of our calculation comes from extrapolation from microbes or mice. In contrast, we have the unhappy knowledge of at least a score of chemicals, smoking, etc being carcinogenic by their human consequences.

Nevertheless, we know of mutagens that, in the laboratory, can multiply mutation rates by a thousand-fold over the spontaneous background, though we have little knowledge of their penetration to human germ cells. And I do not include in this discussion chromosome breakage, or substances like colchicine that are known to impair spindle function. Also left out are agents that affect mitochondrial DNA in other species, like the acridine dyes.

o segregational owing to heterozygote advantage Most of the more prevalent genetic diseases are probably the consequence of just one or a few ancestral mutations. They are maintained not by mutation pressure, but rather by biological advantage of the heterozygotes. The prototypic example is Hb-S, the story familar to all of you ...

The cruel irony is that even in the sustained presence of malaria, the natural equilibrium would be that point where the number of genetic deaths from sickle cell disease just balances the advantage of protection from malaria. The homozygotes pay the price for the advantage, current or historical, enjoyed by the heterozygots. We certainly look for more humane ways than natural evolution either to reduce the gene frequency of Hb-S or to mitigate its impact.

These loci are probably irrelevant to issues of mutagenesis, but may be even predominant as sources of the overall genetic load.

o meiotic drive -- which sustains some "selfish genes" in other organisms has not, as yet, been observed in the human.

Finally, the real scandal in the history of mutagenesis has been the discovery, rightly connected with the name of Barbara McClintock, that a large proportion of mutational events are more of a biological than a chemical phenomenon -- they are due to "jumping genes" or transposons. They have certainly played a large role in the evolution of our genome, a host of repeated sequences have all the earmarks. DNA insertions are part of the protocol for the targetted mutagenesis that is the hope of gene therapy in somatic cells. In bacteria and in mice, viruses are known which can greatly enhance the mutation rate -- and unlike the base substitutions, with real perturbations of gene function. Many dysgenic effects of chemicals and radiation operate indirectly through the activation of built-in transposons. But we do not as yet know their significance in the human situation.

THE GENOME
and its contradictions

TOTAL DNA = 3,000 MM bp.

100,000 genes each gene having

1000 bp structural
2000 bp regulatory

3000 bp / gene

== 300 MM bp.

rest: "90% junk" -- parasitic DNA

EXISTING GENETIC LOAD
how much heterozygosity?

About **1-4 per 1000 bp. nucleotide substitutions.**
0.3 Insertion/Deletion/Rearrange

i.e. 1 in every gene == 100,000 total

most substitutions are silent

10%= 10,000 may have phenotypic consequence*
10%= 1,000 of those may be biologically significant unless already contra-selected

Other measures : 1 - 4 lethal equivalents in recessive mutation

* *open question: how much dominance*

MUTATION RATES
less well known

10 -5 per locus per generation with sign. biol. effect
10 -8 per bp. per generation*

== 1 locus mutated in every gamete, perhaps == .01-.1 L-EQ.

== 10-100 generations for equilibrium on lethals

most heterozygosity, new mutation is neutral??

** There is very little direct information on the spontaneous mutation rate at nucleotide level in man. Methodological limitations: genetic loci at which mutation has been observed may be relatively unstable. Many of these mutants are more complex (deletions, insertions,...). However, some point mutations can be devastating -- e.g. thalassemia or sickle cell disease.*