Slide Tribes and diatribes: races, genes and diseases DYE

{Note: this lecture is based on many sources, some of which are quoted directly alongside the slides. Ask for specific sources if you need them: dyec@who.int.} My third Gresham lecture series is about the way in which health and disease depend on genes, environment and culture. I'm going to start with genes and disease, discussed in the context of racial differences. Race is divisive in any context, and health is no exception. That's always been true, so why do I want to talk about race and health now? Thanks to the Human Genome Project and its successors, the 21st century is going to be an age of genetic revolution that will help to put the concept of race on an entirely different footing with potential benefits for everyone. The essential point, from the perspective of the new genetics, is that racial differences (physical appearance) have some bearing on health, but not much. The true understanding of the genetic basis of health - for individuals and populations - must come from looking directly at genes, not from using arbitrary visual cues. Which is why I like this picture. This is clearly an image of obesity, but it's obesity without the skin colour. Rather, the skin is colour-coded by the amount of heat generated at the body surface (in thermograms). White and yellow are hot, blue and black are cold. White is least fat, black is most (Photo: Ray Clark/Mervyn Goff/Science Photo Library).



But we do have to confront the fact that race is in some ways linked to health. Obesity is linked to diabetes, and the prevalence of diabetes varies among racial groups (graph). In the USA, about 6% of the population has diabetes (mostly type 2), with 1 million new cases added yearly and more than 6 million not yet diagnosed. As the leading cause of blindness, kidney disease, and non-traumatic amputations in adults, diabetes costs the US healthcare system over \$130 million a year. accounting for 1 in 7 healthcare dollars spent. Is there a genetic basis to these differences? One idea about why we're suffering an epidemic of obesity is the "thrifty gene" hypothesis, proposed in 1962 (and subsequently rejected in 1982) by geneticist James Neel. It imagines that ancient pastoralists, subject to frequent famine, had a gene that efficiently stored fat. Some people still have that gene, and it efficiently stores fat as we overeat. Social deprivation nowadays means, not starvation, but an abundance of low-guality, "supersizing" food. If it's true, are thrifty genes more common among, say, American Indians, or is there a social explanation for the high rates of diabetes? In the context of race relations, the question arises of who is to blame. As Stephen O'Rahilly

1

(Cambridge) has put it, there are 3 kinds of political argument about cause: (1) old labour (blame society), (2) conservative (blame irresponsible individuals), (3) new labour (blame your parents).



And studies like this one in the highly-respected New England Journal of Medicine generate immediate controversy. The findings were that a medicine for heart disease (an ACE inhibitor) apparently works for whites but not blacks.

While there are clear racial differences in conditions such as diabetes, and treatments e.g. for heart disease, the question of whether to frame scientific investigations in terms of race has generated much heat, as illustrated by these two contrasting quotations from NEJM. Notice how, in the second, the use of the term "profiling" in the title already has negative connotations (like "surveillance").

Not surprisingly, the debate has escaped from the medical literature. Thus Joseph Graves' book, The Emperor's New Clothes, in which he makes his case against raced-based research. "Present day thinking about pharmacogenetics is that there is genetic variability that relates to how a drug acts in different people. But the genes for drug activity and response don't seem to be localized in the socially defined races. What is more important is to use an experimental design to find out what the genes are. We need to know which genes are responsible for a drug given to someone, and that is not going to be limited by a person's socially constructed race." More than this, racial profiling is, in his view, a hindrance to social integration, including racial integration in medical care.

## Slide Races, genes and diseases: 6 what I will talk about 1. The race-genes-health controversy

- 2. How genes vary among races
- Genes as the basis of disease
- Why we have bad genes

   Trade-offs in evolution
   Selfishness & illness
- 5. Genes, race and you

With that background to the controversy, here is what I'm going to talk about. I've already begun to make one key point: ethnicity and race have something to do with genetics, but not much. A second general point is that, while we commonly divide causes between nature and nurture, the difference is not as obvious as it seems. In the case of disease, our reaction to exposures that might cause illness, like food or tobacco, depends on our genetic makeup. In general, we can't interpret nurture except in the context of nature. Part 2 is about race and genes, parts 3 and 4 are about genes and disease, and part 5 is an attempt to bring race, genes and disease back together in the context of personalized medicine.

Slide 7



What I will not talk about: ethical and moral aspects of race, genes and disease, and of genetic manipulation.

Slide 8



First, a sketch of how genes vary among races. Or more generally, how genes vary among human populations around the world. How much genetic variation is there among races – variation that might be the basis of racial variation in disease profiles?

Slide 9



If you compare two people, you can ask how similar they are and how different. Some analyses are designed to emphasize differences. This is genetic tree shows relatedness (closer when on same branch) for 37 populations based on 80 independent genetic sites (loci). The increasing distance from populations in Africa to those in East Asia and the Americas shows how they have drifted apart geographically and genetically. Within a geographic region, the genetic clustering of populations often parallels the linguistic clustering. African populations are somewhat removed from the others, probably reflecting the initial human expansion out of Africa (Tishkoff Nature Genetics 2004).



Relations between genetic "haplotypes" in different populations European International HapMap Project Asian Asian African

Similarly, this is the genetic variation in a sample of 3000 European individuals genotyped at over half a million variable DNA sites in the human genome. Remarkably, by looking at DNA, it is possible to place 50% of individuals within 300 km of their reported origin and 90% within 700 km of their origin (Novembre, Nature 2008)

Other analyses highlight the similarities. In 2002, a group of scientists gathered in Washington DC to launch the International HapMap Project - a major new initiative to create a map of human genetic variation. The map will be based on common haplotype patterns, i.e. combinations of DNA sequence variants that are usually found together. The results of these studies show that humans are genetically very similar to one another. Any two unrelated genome sequences differ at only one position in a thousand, on average. Haplotype methods have contributed to the identification of genes for Mendelian diseases and, recently, disorders that are both common and complex in inheritance. But the HapMap has not yet provided evidence that race can act as a surrogate for genetic constitution in medicine or public health. Recent genomic surveys have shown that as few as 3-5 common haplotypes capture the bulk of variation at any specific site (locus) throughout the genome, and those haplotypes are generally represented in the populations of all continents (diagram, Cooper NEJM 2003, from Gabriel Science 2002). Therefore, if genes for diseases are located on common haplotypes, those genes must be shared by members of all populations. Consequently, it will be difficult to calculate the "genetic susceptibility score" for any particular racial category.

Slide Biologically, there are no castes and tribes, only communities" VR Rao

And here is one more very recent example of similarity, where cultural differences that identify races (in India) belie underlying genetic similarity. The M2 haplogroup is unique to India, and is a subgroup of the M haplogroup of the first migrants to India. M2 appeared in India about 50,000 years ago, and makes up nearly 10% of the modern tribes' mitochondrial DNA (map). The M2 "signatures of earliest antiquity" cut across major linguistic barriers: both the Dravidian and Austro-Asiatic language groups share these genetic markers. The traces of the earliest settlers are "not restricted to one particular language family, or one particular population", says author Satish Kumar. And according to VR Rao: "Biologically, there are no castes and tribes, there are only communities." But this interpretation is slanted towards similarity; there are clearly differences too among the tribes.



Some studies highlight the similarities, and others the differences. We need to be aware that there are specific reasons for emphasizing similarity and difference, and to understand the facts as they are. We need an objective measure of similarity and difference.

Slide 14

12

Genes as the basis of disease People from different regions of the world do show genetic differences, but on the whole these are small, and not obviously linked to markers of race. So if race is used as a basis for studies of the genetic causes of disease, we are dealing with differences that have a lot to do with culture and based only on a small proportion of our genes. So can a few genes account for racial differences in health?

Slide 15



Among these small genetic differences are some striking effects that highlight racial differences. above all skin colour. This is from a controversial Hillary Clinton advertisement: critics claimed that the picture on the right deliberately darkened Obama's skin colour.



This is the global distribution of skin colour (map http://anthro.palomar.edu/vary/). Why do these arise and what have they got to do with health?

We have about 30,000-40,000 genes, but only 5-50 control skin colour. There is a gradient of pigmentation, imparted by size and density of melanosomes. What are the health effects of these racial distinctions: African, Asian, European? From a medical perspective, inadequate protection from sunlight has a major impact on human health. In Australia, the lifetime cumulative incidence of skin cancer approaches 50%, yet the "smart tanning" industry continues to grow, and there is controversy over the extent to which different types of melanin can influence susceptibility to ultraviolet (UV) radiation. At the other end of the spectrum, inadequate exposure to sunlight, leading to vitamin D deficiency and rickets, has been mostly cured by nutritional advances made in the early 1900s (Barsh PLoS Biol 2003).

Slide 18



So a few genes can have major effects on attributes that affect health. like skin colour. Is that how all genetic disease work? We commonly think of genetic diseases as those severe conditions caused by single gene defects. There are about 6000 disease caused by single gene defects (mutations), but severe and fatal disorders that are purely hereditary are rare, for the obvious reason that people who die young do not pass on their defective genes. One example is Huntington's disease, and the picture show thinning of the brain cortex in early disease and during disease progression (from red least thinning to vellow greatest thinning). HD results from genetically programmed degeneration of brain cells (neurons). in certain areas of the brain. Illness is characterized by abnormally uncoordinated, jerky body movements called chorea, and a decline in mental abilities that affects aspects of behavior. As the disorder progresses, it can cause complications that significantly reduce life expectancy. The exact mechanism of HD is unknown and there is currently no cure. Global incidence varies, from 3 to 7 per 100,000 people of Western European descent, down to 1 per 1,000,000 of Asian and African descent.



Slid 20



The discovery of the oldest evidence yet found of a human hereditary genetic disorder was announced in 2004 by researchers at the Hebrew University of Jerusalem. The fossil is 1.5 million years old and is from a two-year-old Homo erectus (a precursor of modern man) child in Ethiopia suffering from amelogenesis imperfecta, i.e. defective enamel formation, obviously a mild disorder. The affected enamel may be thinner than normal, it may be of normal thickness but lacking strength, or it may be pitted (Zilberman J Hum Evol 2004).

Most "genetic" diseases are not due purely to gene defects, but rather to the interaction between genes and environment. To measure how much of each, we use the idea of heritability. Most diseases do not score either 0 or 1, but something in between. One way to measure heritability is by comparing twins (www.answers.com/topic/heritability). Identical twins (monozygous, MZ twins) are twice as genetically similar as fraternal twins (dizygous, DZ) and so heritability is approximately twice the difference in correlation between MZ and DZ twins. Bigger differences between MZ and DZ suggest that a condition is more heritable, i.e. it has a stronger genetic component. Compare autism and alcoholism in the diagram.

Pictures are sometimes clearer than graphs...

Slide 21





The easiest genetic effects on health to recognize are those that are (a) severe and (b) based on one or two genes, so they have simple hereditary mechanisms. But there are many other subtle effects of genes that we can now see thanks to the power of modern genetic analysis. To identify genetic risk factors for common diseases, researchers in this study scanned DNA from patients (2000 per disease) and controls (3000 shared for all 7 diseases studied) for the frequency with which they contained each of the 500,000 genetic markers, or single nucleotide polymorphisms (SNPs), from the human genome. They found that most markers showed very little difference in the frequency of their two constituent forms — or alleles — between controls and cases. However, some SNPs occurred at a greater frequency in patients. Such alleles (one is shown in red) can be considered to be genetic risk factors for a particular disease. Some of these genetic changes are hard to find because they have small effects. The small effects mean that routine testing is not be worthwhile, especially if there is no obvious course of action to improve health.



Genetic analysis is also uncovering larger effects. A major challenge in genetic analysis is to understand how genes interact with environmental exposures, like tobacco smoking. It could be that a gene appears to have a small effect, only until the right kind of exposure comes along. Lung cancer is largely caused by smoking. Tobacco is responsible for about 5 million premature deaths every year and smoking is still clearly the largest risk factor. But three genetic analyses published this year suggest that, among smokers, some people may be as much as 80% more at risk than others because of their genes. People who have never smoked might also have a slightly increased risk of developing lung cancer and similar problems, although the three studies disagree on whether this is actually the case (Thorgeirsson Hung Nature 2008, Amos Nature Genetics 2008). Would you want to know if you carry these genes?

Slide 24



The key gene variants, which are carried by around half of people with European ancestry, may not only sensitize cells to nicotine, but also make it harder for smokers to give up. But the "doublewhammy" theory is not yet proven.

> Why we have bad genes

Why do we have bad genes? If they are so detrimental to health that carriers don't have children, then surely they should be eliminated from the human population.

Slide 26



Put more specifically, why do genes, like those that cause haemophilia, keep running in families? Why do they not die out with their hapless carriers? Haemophilia is a hereditary disorder in which the blood clots slowly due to a deficiency in blood coagulation factor VIII. History's most famous carrier of the hemophilia gene was Queen Victoria (1819-1901), grandmother to most of the royalty in Europe. In 1853, Victoria gave birth to her eighth child, Leopold, Duke of Albany, who had hemophilia and died at the age of 31 from internal bleeding after a fall. Two of Queen Victoria's four daughters, Alice and Beatrice, also carried the gene for hemophilia. Alice's daughter Alexandra was also a carrier, and she transmitted the disease to her son Alexis, whose father was Czar Nicholas 11 (1868—1918) of Russia (pictured). Alexis, perhaps the most famous of the European royals with hemophilia, was heir to his father's throne. Historians like to discuss, possibly without resolution, the role Alexis's condition played in the Russian revolution of 1918 (Biophoto Associates/Science Photo Library/St Mary's hospital/Corbis). The question of why these genes persist has not been fully answered; the first job would be to calculate how long we can expect them to persist, given the mechanism of inheritance.



For some other inherited blood disorders, like sickle cell disease, there is a much clearer answer to the question of persistence. In sickle cell anaemia, red blood cells contain abnormal haemoglobin, the oxygen-carrying pigment. In areas of deoxygenated blood the haemoglobin distorts the red blood cells into a sickle shape, causing bone ache and organ damage. Carriers of one copy of the sickle cell gene are protected against malaria but do not suffer from sickle cell disease; only people with two copies of the gene suffer this form of anaemia. Contrary to popular belief, sickle cell anaemia is not a black disease nor did it originate in West Africa (map). The gene responsible for sickle cell provides protection against malaria, so it is present wherever we find malaria. That includes Greece, Yemen, India, East and West Africa and the Middle East, where it originated. The reason it is often thought of as a black disease in the Americas and Europe is because the slaves came from West Africa. If the slaves who worked the cotton fields of America had come from Yemen or Greece, then it would have been seen as a Yemeni or a Greek disease (Eye of Science/Science Photo Library).



Still on malaria, it's an apparent geographical oddity that one of the two major causes of "tertian" (paroxysms recur every third day) disease (Plasmodium vivax) is absent from much of west and central Africa (map

www.vivaxmalaria.com/template\_disease.htm). We might describe this phenomenon by saying that certain west African racial groups do not suffer from vivax malaria, but it is much better to describe it in terms of physiology and genes. The mistake we might otherwise make is to assume that it is only these racial groups that have this characteristic. P vivax exploits a protein on the surface of red blood cells, the Duffy antigen (also called DARC), in order to enter cells. Some people are Duffy-negative: they lack the Duffy antigen on their red blood cells and so P vivax can't get access.



Who was Duffy? Not the popular, racially Welsh singer (bottom right), but rather the haemophiliac patient in whom the antigen was first discovered. We've known about Duffy and P vivax since 1950, but there is a new and controversial twist to the Duffy story. US and UK scientists this year found that Duffy might increase their vulnerability of carriers to HIV infection by as much as 40% because of the way it allows the HIV virus to attach to red blood cells. The findings, If extended to all Africans, could explain an estimated 11% of the HIV burden in Africa. (Images: left, a lymphocyte immune cell full of newly manufactured HIV particles (red); centre, red blood cells invded by P falciparum malaria parasites.)

Cystic fibrosis and the white plague Cystic fibrosis 2 gene copies, all die before having children 1 gene copy (carrier), protects against TB  $\int_{0}^{2} \int_{0}^{2} \int_{0}^{2$  What about cystic fibrosis? Cystic fibrosis is the most common lethal single-gene mutation in people of European descent, with a carrier frequency upwards of 2%. The age of the most prevalent mutation,  $\Delta$ F508, has been estimated to be approximately 600 generations, or perhaps 15,000-20,000 years. Based upon molecular research, resistances in the heterozygote to cholera and typhoid fever have been proposed to explain the persistence of the mutation. But neither cholera nor typhoid fever are likely to have provided enough historical selective pressure (killed enough people) to produce the modern incidence of cystic fibrosis. By contrast, the European tuberculosis pandemic (the "white plague", but note the "black death" was also a plague of whitish people) beginning in the 17<sup>th</sup> century would have provided sufficient selective pressure (Poolman RSIF 2007).

Slide 31

Slide

30



This is nothing to do with race, but I can't resist it as an example of why we (vertebrates) have bad genes. Those of you who believe that there cannot be pleasure without pain will want to add this to your catalogue of told-you-sos. Humans have oncogenes – genes that cause cancer -and fish have them too. Why haven't they been eliminated by natural selection? André Fernandez has recently offered a possible answer: in a species of swordtail fish, *Xiphophorus cortezi,* the oncogene *Xmrk* also causes a dark pattern on male fish's tail fin that females find attractive. Indeed, fish that have developed cancer had the largest patterns and were able to mate for many months before the disease killed them (PNAS 2008).

## Genes, race and you

Slide 33	Must genetic diseases be fixed by genetic engineering?	
	Selective breeding (eugenics)     Genetic manipulation         Radiation & mutagens         Viruses         Transposable elements	Constant of the second se

- Gene targeting

- Gene therapy on germline

Cloning

Creating new genes (playing God)

That's a short account of the interrelations between race, genes and diseases. What can we do about genetic defects? I've said that: (a) simple genetic defects that cause severe disease – ones that are perhaps most easily be fixed -- are relatively rare, (b) most diseases have some sort of genetic basis, but there are often numerous genes each with small effects, where the size of the health effects depends on environmental circumstances. Race highlights the problem of understanding the link between genes and disease by grouping people on the basis of superficial physical and social characteristics. Taken together these facts suggest that genetic medicine is likely to evolve slowly. To make progress we need to be very cautious about arbitrary groupings like "race" and focus on the specific environmental and genetic determinants of ill health. Having deduced what causes disease in appropriately defined populations (by doing good epidemiology), we may get close to finding remedies for individuals with specific characteristics, so called "personalized medicine".

We want to get rid of genes that cause disease. How do we do that? There are broadly 3 ways, old and new, to manipulate genes in individuals and populations: (1) Selective breeding ("unnatural selection"), is done by choosing animals or plants with desirable traits. In humans, parents could produce a large number of fertilized eggs in vitro. Each could be grown for a while in vitro, and the ones with the desired characterstics implanted in the mother. Besides the moral and ethical objections, there are some big problems to overcome: selecting for a large number of traits will be impossible; certain desirable traits (disease free) are the result of many genes acting together; fertilized eggs have one copy of each gene from each parent, so bad genes cannot be eliminated if one parent has two copies of that gene. (2) There are many methods of genetic manipulation, of introducing new genetic material into a cell or organism, or altering the existing material: radiation and mutagenic compounds; viruses engineered to carry desirable genes; transposable elements (natural gene shuffling tools); gene targeting to replace a specific gene with another within a cell. Gene therapies usually try to replace or repair defective genes in tissues of fully-grown organisms. Gene therapy does not usually alter the "germ line" (reproductive cells), so even if gene therapy corrects a problem, the problem will still be inherited by children. With cloning technology, not even a fertilized egg is needed, just a cell that will grow in cell culture. But most of these procedures are still inefficient, and in any case most of the

functions of human genes are not yet known. (3) Ultimately we might be able to create entirely new genes and proteins, and even new organisms (www.biofact.com/cloning/#Breeding). (4) More realistically, in the near future, we can adjust environmental risks that interact with genes. Despite the difficulties, here are some examples...

Slide 34



Some forms of obesity are due to simple gene defects. As you'd predict, they are rare forms, but the solutions are nonetheless immensely satisfying, not to say miraculous. This gene makes leptin, which controls the "adipostat" (the fat controller, with thanks to Thomas the Tank Engine).

Slide 35



Leptin therapy is life-changing for these children.

Slide 36



This is another, less-specific kind of environmental manipulation. Here 30 men with early-stage prostate cancer, Dean Ornish University of California San Francisco tested the effects of a dramatic lifestyle change on gene expression in the prostate gland. Biopsies taken before and after 3 months of healthy eating, moderate exercise, stress management and psychotherapy showed a significant change in the expression of hundreds of genes. Many, including several genes involved in tumor formation, were down-regulated, or less active. Others, including some disease-fighting genes, were more active.



Real genetic engineering is happening too: not simply replacing a gene product, but rather fixing the defective gene itself. Leber's congenital amaurosis (LCA) is a group of inherited disorders involving retinal degeneration with severe vision loss noted in early infancy. One form, LCA2, is caused by mutations in the retinal pigment epithelium. This very recent work investigated the safety of subretinal delivery of a virus engineered to carry the necessary piece of DNA. Each patient had a modest improvement in retinal function and visual acuity. Although the follow-up was very short and normal vision was not achieved, this study is a basis for further gene therapy studies in patients with LCA.

## Slide

38



Where technology improves, enterprise will follow. Genetic testing companies are springing up, like this one in California (if you don't know why it's called 23andMe, you soon will) They offer to sequence your genome, charging about \$1000, and tell you about your personal risk of disease, embracing the latest research on how your genes may affect risk for common diseases and conditions such as heart attack, arthritis and cancers. You can explore the genetics of your food preferences, eye color, athletic ability and other traits. Three companies are now offering extensive DNA read-outs directly to the public. All three companies use DNA chips to read off hundreds of thousands of the most common single-letter variations in people's DNA, known as single nucleotide polymorphisms (SNPs). While this does not provide nearly as much information as sequencing the entire genome, these scans do offer people an unprecedented glimpse of their own genetic blueprint. The real challenge, though, is not reading the genetic information but interpreting it. As yet, we don't have any idea what, if anything, most SNPs mean for those who possess them.

## Slide 39



But it's nicely presented at least, here on breast cancer. 23andMe does not report on the rare but high-risk mutations in the BRCA1 or BRCA2 genes. Rather they analyze 2 SNPs for which the lifetime risk is said to be 12% to 22%.



The cover of Ridley's book answers the riddle of why 23andMe.

Google co-founder Sergei Brin has revealed that his 23andMe account says he carries a mutation known as G2019S in gene *LRRK2*, which *p*uts him at higher risk of contracting Parkinson's. He also disclosed that his mother, who has been diagnosed with the disease, carries the same mutation (New Scientist 19 September). However, there are some awkward facts about 23andMe and Google, which might undermine user confidence in the service.

Slide 42



Notwithstanding the cautions about race, gene and disease, some disease genes are well known to be commoner in certain racial groups. About 1 in 40 Ashkenazi Jews — with or without breast cancer has a genetic mutation in BRCA1 or BRCA2. Women with these genes have a 50% to 80% chance of getting breast cancer in their lifetime. In one study of more than 5300 Ashkenazi Jewish men and women (NEJM 1997), 120 people (2.3%) had one of the three specific abnormalities in BRCA1 and BRCA2 known to be associated with a higher risk of breast cancer. A recent study (JAMA 2007) of more than 3000 women diagnosed with breast cancer looked at the risk of abnormal BRCA1 genes in different ethnicities. The results are in the graph

(www.breastcancer.org/risk/genetic/odds\_abnrml\_g enes.jsp).

Slide 43



Masha Gessen, among others, has written a gripping account of the grim choices facing those at high risk of breast and ovarian cancer.



10

- 1. Race is generally a poor guide to genetic diseases... ... doesn't distinguish environmental and genetic causes ... will not identify many genetic causes of
- disease ... could mean wrong diagnosis or treatment 2. Revolution in genetics will lead to
- personalized medicine in some form... ... the work of the 21st century

Given the controversy over race, genes and health, Lee & colleagues have recently set out 10 guiding principles, which are a mixture of science and ethics. I would have stuck more specifically to the science. Point 7 is key, though race cannot be entirely excluded (Lee Genome Biology 2007).

To conclude, race can be an indicator of health risks, but it is generally a poor indicator. It does not distinguish environmental, social and genetic causes, and is not likely to say much about the genetic causes of disease. Because different races share the vast majority of their genes, the genetic causes of disease, and their treatments, will usually be independent of race. Genetic defects that are at first sight linked to race could actually be more widely distributed, like sickle cell anaemia. And weak associations could lead to misdiagnosis or mistreatment of patients. The new revolution in genetics will lead to personalized medicine in some form, but this will start slowly with e.g. the better identification of genes that are related to disease and the mitigation of environmental factors that interact with genes. Fixing the genetic defects is a much longer project: personalized medicine is the work of the 21<sup>st</sup> century.