James F. Crow

# *Unequal by nature: a geneticist's perspective on human differences*

In February of 2001, Craig Venter, president of Celera Genomics, commenting on the near-completion of the human genome project, said that "we are all essentially identical twins." A news headline at the time made a similar point: *Are We All One Race? Modern Science Says So.* In the article that followed, the author quoted geneticist Kenneth Kidd: "Race is not biologically definable, we are far too similar."

Venter and Kidd are eminent scientists, so these statements must be reasonable. Based on an examination of our DNA, any two human beings are 99.9 percent identical. The genetic differences between different groups of human beings are similarly minute.

*James F. Crow, a Fellow of the American Academy of Arts and Sciences since 1966, is professor emeritus of medical genetics at the University of Wisconsin. Over a career that has spanned more than fifty years, he and his collaborators have studied a variety of traits in Drosophila, dissected the genetics of DDT resistance, measured the effects of minor mutations on the overall fitness of populations, described the behavior of mutations that do not play the game by Mendel's rules, studied the effects of nonrandom mating, and considered the question "What good is sex?"*

Still, we only have to look around to see an astonishing variety of individual differences in sizes, shapes, and facial features. Equally clear are individual differences in susceptibility to disease – and in athletic, mathematical, and musical abilities. Individual differences extend to differences between group averages. Most of these average differences are inconspicuous, but some – such as skin color – stand out.

Why this curious discrepancy between the evidence of DNA and what we can clearly see? If not DNA, what are the causes of the differences we perceive between individuals and between groups of human beings?

DNA is a very long molecule, composed of two strands twisted around each other to produce the famous double helix. There are forty-six such DNA molecules in a human cell, each (along with some proteins) forming a chromosome. The DNA in a human chromosome, if stretched out, would be an inch or more in length. How this is compacted into a microscopic blob some 1/1000 inch long without getting hopelessly tangled is an engineering marvel that is still a puzzle.

The "business" part of the DNA, the part that carries genetic information, is the sequence of nucleotides, or bases, in the molecule. There are four of these, commonly designated as A, G, T, and C. (I could tell you what these letters stand for, but you wouldn't understand this essay any better if I did, so I won't.)

In the double helix, there are four kinds of base pairs: AT, GC, TA, and CG. The specific pairing rules – A with T and G with C – are dictated by the three-dimensional structure of the bases.

In a chromosome, the base pairs are in a precise sequence, and the orderly process of cell division assures the reproduction of this sequence with remarkably few errors. Chromosomes occur in pairs, one member of each pair from each parent, and the DNA sites in the two corresponding chromosomes match up. We have twenty-three pairs of chromosomes, or a total of forty-six, as previously mentioned, in each cell. These forty-six chromosomes contain about six billion base pairs. If we randomly choose a pair of bases from corresponding sites in two persons, 99.9 percent of the time they will be the same. This percentage depends only slightly on whether the two people are from the same or from different continents, from the same or from different population groups.

In order to make sense of how the DNA of human beings can be so similar, despite all the important visible and physiological differences among individuals and groups, it is helpful to recount our evolutionary history.

All mammals, including ourselves, are descended from an ancestral species that lived about one hundred million years ago. In our mammalian ancestry an average base has changed, say from an A to a T, at the almost unbelievably slow rate of about one change per billion years. This means that only a small fraction of the bases, one hundred million divided by one billion, or 1/10, have changed during that time. As a result, we share roughly 90 percent of our DNA with mice, dogs, cattle, and elephants.

Coming closer to home, the DNA of human beings and chimpanzees is 98 to 99 percent identical. The differences between us that we (and presumably the chimps) regard as significant depend on only 1 or 2 percent of our DNA.

Much of human DNA is very similar to even more remote ancestors: reptiles, invertebrates, and even plants. All living things share many functions (e.g., respiration) going back to a very distant past. Most of our DNA determines that we are human, rather than determining how we are different from any other person. So it is not so surprising that the DNA of any two human beings is 99.9 percent identical.

What produces variability between individual organisms – and makes possible evolutionary change – is errors in the DNA copying process. Sometimes, because of this, one base is changed to another – it mutates. Among the six billion base pairs each of us inherits from our parents, a substantial number – a hundred or more – are new mutations.

How can we reconcile this large number with the extremely slow rate of evolutionary change? The explanation is that only a tiny fraction of mutations persist over time. Some mutations survive as a matter of either luck or – if the mutation confers a biological advantage – natural selection. Even if advantageous, an individual mutation has little chance of surviving a long evolutionary trip. The slow rate of evolutionary change explains why we mammals are so similar in our DNA.

Molecular studies of DNA have been extremely fruitful in working out the evolutionary history of life. Much of what we know about human ancestry comes from DNA studies, supplemented by a rather spotty fossil record. The DNA evidence strongly supports the idea that the human species originated in Africa, and that European and Asiatic populations – indeed, all non-Africans – are descended from a small number of migrants from Africa. The strongest evidence for this is that Africans are more variable in their DNA than are other populations.

Analysis of DNA allows us to measure with some precision the genetic distance between different populations of human beings. By this criterion, Caucasians and Asians are relatively similar, whereas Asians and Africans are somewhat more different. The differences between the groups are small – but they are real.

DNA analysis has provided exciting new answers to old questions. But its findings can also be misleading. Take the case of men and women and sex chromosomes. Females have two X chromosomes, while males have an X and a Y. The Y chromosome makes up perhaps 1 percent of the DNA. But there is very little correspondence between the Y and the other chromosomes, including the X. In other words, the DNA of a human male differs as much from that of a female as either does from a chimpanzee of the same sex. What does this mean? Simply that DNA analysis, which has given us a revolutionary new understanding of genetics and evolution, doesn't give sensible answers to some contemporary questions that society is interested in.

Most of the differences that we notice are caused by a very tiny fraction of our DNA. Given six billion base pairs per cell, a tiny fraction – 1/1000 of six billion base-pairs – is still six *million* different base pairs per cell. So there is plenty of room for genetic differences among us. Although we differ from each other in a very tiny proportion of our DNA, we differ by a large number of DNA bases.

Some noteworthy evolutionary changes in human beings have occurred relatively rapidly, despite the slow overall rate of change at the DNA level. The difference between the skin color of Africans and Europeans probably evolved in less than fifty thousand years, an adaptation to differences in climate. Still more rapid were changes in genes that confer resistance to malaria in Africa and Mediterranean regions; it only took between four and eight thousand years for the new genes to evolve. What genetic analysis reveals is that some of the genetic changes that seem so significant to us depended on a very tiny fraction of our DNA.

But, as I said, this tiny fraction is still a very large number of bases. No two human beings are alike in the traits they possess. Some are tall, others are short; some are stocky, others thin; some are gifted musically, others tone deaf; some are athletic, others awkward; some are outgoing, others introverted; some are intelligent, others stupid; some can write great poetry or music, most cannot. And so on.

To understand our differences, we need to consider not just DNA, but its cellular products as well. This area of study is new, but it is progressing rapidly. The emphasis is changing from DNA sequences to genes. A gene is a stretch of DNA, usually several thousand base pairs long. The function of most genes is to produce proteins. The genome sequencing project has revealed that we humans have thirty to forty thousand genes. But since a gene often produces more than one kind of protein, sometimes producing different kinds for different body parts, the number of kinds of protein is more like one hundred thousand.

We share a number of genes with

chimpanzees, genes that make us primates rather than elephants or worms. Evolutionary scientists believe that many of the differences that we observe between ourselves and chimpanzees involve changes in the amount rather than in the nature of gene products. Human beings and chimpanzees share proteins that produce body hair and brains, but in chimpanzees these proteins produce more hair and less brains. Why this should be so is still far from being fully understood. But this is a research area that is advancing very rapidly, and there are good genetic leads to be followed up.

Of course, not every human difference has a genetic cause. Many are environmental, or are the result of interactions between genes and environment. Even genetically identical twins develop into distinct individuals.

The ability to learn a language is largely innate, built into the nervous system of all normal people, as demonstrated so beautifully in the effortless way in which young children learn to speak. But the particular language any individual learns obviously depends on the social setting. Mozart was a great composer partly because of his genes and partly because of his training. Ramanujan had a great talent for mathematics, but without his being exposed to a textbook – not a very good one, by the way – he could never have made his astounding discoveries. Michael Jordan has a talent for basketball, but it would never have developed had he grown up among the Inuits.

Just as there are great differences among individuals, there are average differences, usually much smaller, between groups. Italians and Swedes differ in hair color. Sometimes the differences are more conspicuous, such as the contrasting skin color and hair shape of Africans and Europeans. But, for the most part, group differences are small and largely overshadowed by individual differences.

Biologists think of races of animals as groups that started as one, but later split and became separated, usually by a geographical barrier. As the two groups evolve independently, they gradually diverge genetically. The divergences will occur more quickly if the separate environments differ, but they will occur in any case since different mutations will inevitably occur in the two populations, and some of them will persist. This is most apparent in island populations, where each island is separate and there is no migration between them. Each one has its own characteristic types. In much of the animal world, however, and also in the human species, complete isolation is very rare. The genetic uniformity of geographical groups is constantly being destroyed by migration between them. In particular, the major geographical groups – African, European, and Asian – are mixed, and this is especially true in the United States, which is something of a melting pot.

Because of this mixing, many anthropologists argue, quite reasonably, that there is no scientific justification for applying the word "race" to populations of human beings. But the concept itself is unambiguous, and I believe that the word has a clear meaning to most people. The difficulty is not with the concept, but with the realization that major human races are not pure races. Unlike those anthropologists who deny the usefulness of the term, I believe that the word "race" can be meaningfully applied to groups that are partially mixed.

Different diseases are demonstrably characteristic of different racial and ethnic groups. Sickle cell anemia, for example, is far more prevalent among people of African descent than among Europeans. Obesity is especially common in Pima Indians, the result of the sudden acquisition of a high-calorie diet to which Europeans have had enough time to adjust. Tay-Sachs disease is much more common in the Jewish population. There are other examples, and new ones are being discovered constantly.

The evidence indicating that some diseases disproportionately afflict specific ethnic and racial groups does not ordinarily provoke controversy. Far more contentious is the evidence that some skills and behavioral properties are differentially distributed among different racial groups. There is strong evidence that such racial differences are partly genetic, but the evidence is more indirect and has not been convincing to everyone.

To any sports observer it is obvious that among Olympic jumpers and sprinters, African Americans are far more numerous than their frequency in the population would predict. The disproportion is enormous. Yet we also know that there are many white people who are better runners and jumpers than the average black person. How can we explain this seeming inconsistency?

There is actually a simple explanation that is well known to geneticists and statisticians, but not widely understood by the general public or, for that matter, by political leaders. Consider a quantitative trait that is distributed according to the normal, bell-shaped curve. IQ can serve as an example. About one person in 750 has an IQ of 148 or higher. In a population with an average of about 108 rather than 100, hardly a noticeable difference, about 5 times as many will be in this high range. In a population averaging 8 points lower, there will be about 6 times fewer. A small difference of 8 points in the mean translates to several-fold differences in the extremes.

Asian Americans represent about 12 percent of the California population, yet they represent 45 percent of the student body at the University of California at Berkeley. Asians have only slightly higher average SAT scores than Caucasians, but the university's policy of admitting students with the highest SAT scores has yielded a much larger proportion from the group with the higher mean.

Two populations may have a large overlap and differ only slightly in their means. Still, the most outstanding individuals will tend to come from the population with the higher mean. The implication, I think, is clear: whenever an institution or society singles out individuals who are exceptional or outstanding in some way, racial differences will become more apparent. That fact may be uncomfortable, but there is no way around it.

The fact that racial differences exist does not, of course, explain their origin. The cause of the observed differences may be genetic. But it may also be environmental, the result of diet, or family structure, or schooling, or any number of other possible biological and social factors.

My conclusion, to repeat, is that whenever a society singles out individuals who are outstanding or unusual in any way, the statistical contrast between means and extremes comes to the fore. I think that recognizing this can eventually only help politicians and social policymakers.

These are times of very rapid change in our understanding of biological processes. The genome project is but one example. At the same time, we are getting much closer to a deep understanding of the nervous system and of human behavior. Medical knowledge improves, as does data collection and computer analysis. All of these tell us more about individual and group differences. What will be the impact of this new knowledge on societal issues? What are the political implications of modern biology?

We have seen that the DNA sequence similarities revealed by the genome project, valuable as these are for answering many interesting and important questions, are misleading in regard to important human differences. But this situation is rapidly changing. The current emphasis goes beyond simple DNA sequences to identifying the individual genes, their products, and their complex interactions. At the same time, not only the kinds of gene products (usually proteins) but their relative amounts are being investigated by much sharper new tools. Genes differ greatly in their productivity, including differences in activity in different parts of the body.

In the near future, biologists will be able to tell us much more than we now know about the genetic and environmental causes of human differences. The most obvious and immediate human benefits will be in medicine. We can foresee the time when many – we can hope most – of our individual susceptibilities to disease will be understood, so that the disease can be predicted in advance, allowing doctors to anticipate and tailor treatments for the particular person. Small steps in this direction have already been made. New treatments are under development. As a result of our genetic understanding, we also now better understand how to manipulate the environment in order to help prevent disease.

At the same time, the study of gene products and their regulation is being extended to normal traits. We can expect that the molecular biology of the future, perhaps the quite near future, will provide precisely the kind of information that in the past has depended on observation and statistical analysis of often vaguely defined traits. We shall be able, as individuals, to know a great deal about our own genetic makeup.

The magnificent advances in molecular biology will bring new depths of understanding of human differences, normal and pathological, and the extent to which these are genetic or environmental – or, as usually will be the case, both. Whether society will accept this knowledge willingly and use it wisely I don't know. My hope is that gradual progress, starting with small beginnings, can lead to rational individual behavior and thoughtful, humanitarian social policies.

It is important for society to do a better job than it now does in accepting differences as a fact of life. New forms of scientific knowledge will point out more and more ways in which we are diverse. I hope that differences will be welcomed, rather than accepted grudgingly. Who wants a world of identical people, even if they are Mozarts or Jordans?

A good society ought to provide the best kind of environment for each person and each population. We already do this in part. We give lessons to musically gifted children. We encourage athletes and give them special training (and sometimes dubious drugs). Students elect courses according to their abilities and interests. We have special classes for those with disabilities, and such classes are becoming more specific as the causes of the disabilities are understood.

We cannot, of course, tailor-make a special environment for every individual, but we can continue to move in this direction. Finding a genetic basis for a trait doesn't mean that environment is unimportant. Indeed, more environmental influences on the human organism are constantly being discovered, often through genetic studies.

A test of our democratic institutions will be the degree to which people can accept all our differences and find ways to fit them into a smooth-working, humanitarian society. And I argue that we should strive not only for maximum personal satisfaction but for maximum contribution; each of us owes society the fruits of our special gifts. I believe strongly that research into the genetic and environmental causes of human differences should continue and be supported. The newer procedures brought about by molecular advances and computers will greatly accelerate discoveries.

I believe that knowledge, even unpleasant knowledge, is far preferable to ignorance. I hope that American society can be less fearful of learning the truth about biological inequalities and more courageous in using discoveries in ways that are humanitarian and promote human welfare.

The question of equal opportunity versus equal outcomes becomes particularly vexing in those occupations and professions for which only a small fraction of a population can qualify. I have already mentioned the gross overrepresentation of African Americans among Olympic runners. This is closer to a true meritocracy than anything else I can think of: a stopwatch is color-blind. In this case, there seems to be no social purpose in demanding equal racial representation.

In some important professions, such as physics and engineering, Asian Americans are overrepresented and African Americans underrepresented. We presumably get better research because of this. This may or may not outweigh the inequity of unequal group representation. That is a social decision.

What about physicians? There may well be social considerations, perhaps temporary ones in our society, that would make race more important than test scores in selecting students for medical schools.

To achieve political and social equality it is not necessary to maintain a fiction that important human differences do not exist. The great evolutionist Theodosius Dobzhansky said it well: "People need not be identical twins to be equal before God, before the law, and in their rights to equality of opportunity."

I have emphasized that people differ, and differ greatly. They differ not only in shapes and sizes, but also in abilities and talents. They also differ in tastes and preferences. As Shaw said, "Do not do unto others as you would that they should do unto you. Their tastes may not be the same." Society's business, I think, is not to minimize individual differences. We shouldn't try to fit people into one mold.

While I expect that science will continue to provide us with further evidence of human variability, and while I welcome such variability as a source of social enrichment, there are some kinds of human variability that we could well do without. I refer to serious, painful, debilitating diseases. Many of these are the result of an unlucky throw of the genetic dice. Already there are ways of discovering, preventing, and treating some of them. More treatments are sure to come. I hope they will be accepted willingly and used responsibly. I for one would be content if the genes for Tay-Sachs disease and Duchenne muscular dystrophy were to become extinct, along with the malaria parasite and AIDS virus. I hope the great humanitarian benefits that could come from genetic research will not be held up by fears of possible future misuse.

Let me leave the last word for Jim Watson, co-discoverer of the double helix and a major figure in the genome project:

> If the next century witnesses failure, let it be because our science is not yet up to the job, not because we don't have the courage to make less random the sometimes most unfair courses of human evolution.